



**HAL**  
open science

**Etude structurale et fonctionnelle de la reconnaissance  
et de la métabolisation de lésions puriques et  
pyrimidiques dans l'ADN par la  
Formamidopyrimidine-ADN glycosylase**

Yann-Vaï Le Bihan

► **To cite this version:**

Yann-Vaï Le Bihan. Etude structurale et fonctionnelle de la reconnaissance et de la métabolisation de lésions puriques et pyrimidiques dans l'ADN par la Formamidopyrimidine-ADN glycosylase. Sciences agricoles. Université d'Orléans, 2009. Français. NNT : 2009ORLE2004 . tel-00488816

**HAL Id: tel-00488816**

**<https://theses.hal.science/tel-00488816>**

Submitted on 3 Jun 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ÉCOLE DOCTORALE SCIENCES ET TECHNOLOGIES**  
**LABORATOIRE CENTRE DE BIOPHYSIQUE MOLÉCULAIRE**

**THÈSE** présentée par :  
**Yann-Vaï LE BIHAN**

soutenue le : **11 Mai 2009**

pour obtenir le grade de : **Docteur de l'université d'Orléans**

Discipline/ Spécialité : Biophysique Moléculaire

**Étude structurale et fonctionnelle de la  
reconnaissance et de la métabolisation de lésions  
puriques et pyrimidiques dans l'ADN par la  
Formamidopyrimidine-ADN glycosylase**

**THÈSE dirigée par :**

**M. Bertrand CASTAING**

Directeur de Recherche, CNRS Orléans

**RAPPORTEURS :**

**M. Serge BOITEUX**

Directeur de Recherche, CEA Fontenay-aux-Roses

**Mme. Marie-Hélène LEDU**

Chercheur CEA, CEA Saclay

---

**JURY :**

**M. Daniel LOCKER**

Professeur, Université d'Orléans      Président du jury

**M. Serge BOITEUX**

Directeur de Recherche, CEA Fontenay-aux-Roses

**Mme. Marie-Hélène LE DU**

Chercheur CEA, CEA Saclay

**M. Didier GASPARUTTO**

Ingénieur de Recherche, CEA Grenoble

**Mme. Evelyne SAGE**

Directeur de Recherche, CNRS Orsay

**M. Bertrand CASTAING**

Directeur de Recherche, CNRS Orléans

*Ce travail de thèse a été effectué au Centre de Biophysique Moléculaire du CNRS d'Orléans, dirigé par Monsieur Jean-Claude BÉLOËL, que je remercie pour m'avoir accueilli au sein de son laboratoire.*

*Je tiens à exprimer toute ma gratitude à Monsieur Bertrand CASTAING pour m'avoir accueilli dans son équipe « Interactions nucléoprotéiques : Approches structurales et fonctionnelles de la réparation de l'ADN », pour avoir encadré mon travail de thèse, et pour m'avoir fait découvrir l'univers de la recherche fondamentale, et notamment le monde passionnant de la réparation de l'ADN. Je le remercie pour son soutien, pour tous ses conseils avisés, et pour ses qualités de patience et de disponibilité.*

*Je remercie également Monsieur Daniel LOCKER pour avoir accepté de présider ma soutenance de thèse, ainsi que Monsieur Serge BOITEUX et Madame Marie-Hélène LEDU pour avoir accepté la lourde tâche de rapporteur. De même, je remercie Monsieur Didier GASPARRUTTO et Madame Evelyne SAGE pour leur participation en tant qu'examineur.*

*Ensuite, je tiens à exprimer toute ma reconnaissance et mon attachement aux membres de l'équipe : Madame Françoise CULARD, pour sa gentillesse indéfectible et ses conseils précieux, Madame Virginie NADAN, Monsieur Damien LE MINOUX et Mademoiselle Véronique MASSONNEAU pour leur amitié et leur chaleur humaine.*

*Je remercie aussi Monsieur Franck COSTE de l'équipe « Immunité structurale » du CBM, pour m'avoir accueilli lors de mon stage de Master II, et pour sa participation à ce travail de thèse. Merci également à Madame Mélanie SPOTHEIM-MAURIZOT et Monsieur Stéphane GOFFINONT de l'équipe « Radiobiologie moléculaire » du CBM, avec qui j'ai souvent discuté et partagé mes problèmes expérimentaux.*

*Un grand merci à tous ces amis que j'ai connus au CBM, sans qui ces années passées à Orléans auraient été bien moins agréables : Oumarou, François, Damien, Frédérique, Baptiste, Romy, Pauline, Makhlouf, Aurélien, Virginie, Véronique, Olivier, Lucie, Anthony, Ibai, Brice-loïc, et j'en oublie sûrement... Je repenserai longtemps à ces pauses café, ces midis et ces soirées passées ensemble. De même, merci aux membres passés et présents de l'ADSO, pour ces moments de décontraction essentiels dans la vie d'un petit thésard breton loin de son île...*

*A Loiez, Julien, Christelle, Anne-Laure, David et Aurélie, Stéphane, Nico, Baptiste et Virginie, Mathieu, Bastien, Redouane, les lolos, Jacky, et mon petit Naël... Je vous aime tous très fort, merci pour tout.*

*Et bien sûr à ma toute ma famille, mes grands parents, mes frères Glenn et Lomig, merci d'avoir toujours été là.*

*Enfin, à mes parents, Hubert et Mireille, sans qui rien n'aurait jamais été possible. Merci de m'avoir toujours soutenu dans la vie comme dans mes études, je vous dois tout. Cette thèse est avant tout la vôtre.*

*A ma Belle-île-en-mer : Kerprat, le Stang-Perh, Toul-Poulzado, Herlin et Kerdonis,  
ces paradis battus par les vents qui ne me quitteront jamais.*



**Table des matières**

**Abréviations**

**Index des Tableaux et  
Figures**

## Table des matières

<u>Abréviations</u> .....	4
<u>Index des Tableaux et Figures</u> .....	6
<b>I. <u>Introduction</u></b> .....	<b>11</b>
<b>I.1. <u>L'ADN</u></b> .....	<b>11</b>
<b>I.2. <u>Les agressions sur les acides nucléiques et les dommages générés</u></b> .....	<b>13</b>
I.2.1. <u>Généralités</u> .....	13
I.2.1.1. Les hydrolyses.....	15
I.2.1.2. Les alkylations.....	16
I.2.1.3. Les oxydations.....	18
<b>I.3. <u>Les conséquences des dommages sur l'ADN en l'absence de réparation</u></b> .....	<b>19</b>
I.3.1. <u>Les conséquences sur les cellules</u> .....	19
I.3.2. <u>Les conséquences sur les organismes pluricellulaires</u> .....	20
<b>I.4. <u>Les systèmes de réparation de l'ADN</u></b> .....	<b>25</b>
I.4.1. <u>La réparation par réversion directe</u> .....	26
I.4.2. <u>La réparation par excision et resynthèse</u> .....	26
I.4.3. <u>La réparation des coupures double-brins</u> .....	27
<b>I.5. <u>Le système BER</u></b> .....	<b>28</b>
I.5.1. <u>Généralités</u> .....	28
I.5.2. <u>Les ADN glycosylases</u> .....	34
I.5.2.1. Les caractéristiques conservées chez toutes les ADN glycosylases.....	34
I.5.2.2. Les deux grandes familles mono- et bi-fonctionnelles.....	35
I.5.2.3. Les superfamilles structurales d'ADN glycosylases.....	37
I.5.2.4. Spécificité de substrat et redondance des ADN glycosylases.....	39
I.5.2.5. Les stratégies permettant des études structurales sur les ADN glycosylases.....	40
I.5.2.6. Les ADN glycosylases et la variabilité des immunoglobulines.....	43
I.5.3. <u>Les protéines métabolisant les sites AP et extrémités modifiées de l'ADN</u> .....	46
I.5.4. <u>Les ADN polymérases et leur implication dans la mutagenèse induite par les lésions</u> .....	48
I.5.4.1. Caractéristiques générales et familles des ADN polymérases.....	48
I.5.4.2. Les ADN polymérases, la réparation de l'ADN, et la voie BER.....	50
I.5.4.3. Les ADN polymérases de réplication de l'ADN endommagé.....	52
I.5.4.4. Les ADN polymérases réplcatives et leur implication dans la mutagenèse associée aux lésions.....	53
I.5.5. <u>Les ADN ligases</u> .....	57
<b>I.6. <u>Le système GO</u></b> .....	<b>58</b>
I.6.1. <u>Le système GO procaryote et la 8-oxo-guanine</u> .....	58
I.6.2. <u>Le système GO étendu, ou « traitement antimutagène bipartite de la 8oxoG »</u> ....	61
I.6.3. <u>Seconde extension du système GO pour le traitement des paires 8-oxo-G/G</u> .....	62
I.6.4. <u>Le système GO chez l'Homme</u> .....	64
I.6.5. <u>Le système GO procaryote et le FapyG</u> .....	66
<b>I.7. <u>La protéine bactérienne Fpg</u></b> .....	<b>67</b>
I.7.1. <u>Généralités</u> .....	67
I.7.2. <u>Spécificités de substrat de Fpg</u> .....	73
I.7.3. <u>Le mécanisme d'action de Fpg et les structures correspondantes</u> .....	80
<b>I.8. <u>L'objectif du travail de thèse et les stratégies utilisées</u></b> .....	<b>85</b>
I.8.1. <u>Stratégies pour l'étude de complexes Fpg/ADN</u> .....	85
I.8.2. <u>Objectif du travail de thèse</u> .....	86

<b>II. Résultats.....</b>	<b>87</b>
<b>II.1. Nouveaux éléments structuraux et fonctionnels sur la reconnaissance et la métabolisation des purines oxydées par la Fpg de <i>Lactococcus lactis</i> (LIFpg).....</b>	<b>87</b>
<b>II.1.1. Reconnaissance d'une lésion encombrante par LIFpg.....</b>	<b>87</b>
II.1.1.1. Publication "L'enzyme bactérienne Fpg du système BER reconnaît le FapydG substitué en N7 par un gros adduit <i>via</i> un mode de fixation non-productif" (Coste <i>et al.</i> , 2008).....	89
II.1.1.2. Effet de la base opposée aux FapyGs sur la reconnaissance de ces lésions par LIFpg.....	102
II.1.1.3. Conclusion et discussion.....	105
<b>II.1.2. Éléments fonctionnels relatifs à la reconnaissance et à l'excision de la 8-oxoguanine par la protéine Fpg de <i>L. lactis</i>.....</b>	<b>112</b>
II.1.2.1. Analyse fonctionnelle des mutants de LIFpg dans un contexte d'extraits cellulaires bruts.....	114
II.1.2.2. Validation des résultats obtenus avec les extraits cellulaires bruts <i>via</i> l'analyse fonctionnelle de mutants de LIFpg purifiés.....	118
II.1.2.3. Conclusion et discussion.....	120
<b>II.2. Nouveaux éléments structuraux et fonctionnels sur le comportement de la Fpg de <i>Lactococcus lactis</i> (LIFpg) vis-à-vis d'une pyrimidine oxydée, la 5-Hydroxy-5-Méthyle-Hydantoïne (Hyd).....</b>	<b>130</b>
<b>II.2.1. Etude structure-fonction de la reconnaissance et de la métabolisation de l'Hyd par LIFpg.....</b>	<b>130</b>
II.2.1.1. Excision de l'Hyd par LIFpg et comparaison avec l'endonucléase III de <i>E. coli</i> ( <i>EcNth</i> ).....	131
II.2.1.2. L'analogue carbocyclique de l'Hyd (cHyd), un bon modèle pour l'étude de l'interaction de LIFpg et <i>EcNth</i> avec la lésion.....	134
II.2.1.3. Bases structurales de la reconnaissance de l'Hyd par LIFpg.....	137
II.2.1.3.1. Résolution et validation de la structure cristallographique du complexe abortif entre LIFpg et un duplexe d'ADN contenant l'analogue de substrat cHyd opposé à une cytosine.....	137
II.2.1.3.2. Description et analyse de la structure du complexe non-covalent formé entre LIFpg et le duplexe d'ADN cHyd/C.....	140
II.2.1.4. Conclusion et discussion.....	147
<b>II.2.2. Formation d'un complexe covalent abortif entre l'Hyd et LIFpg.....</b>	<b>159</b>
II.2.2.1. Mise en évidence et caractérisation préliminaire du complexe covalent..	159
II.2.2.2. Analyse comparative de la spécificité de la réaction de pontage et de l'activité Hyd-ADN glycosylase de LIFpg.....	163
II.2.2.3. La proline N-terminale de LIFpg est le nucléophile impliqué dans la réaction suicide avec le cHyd.....	167
II.2.2.3.1. Évidence structurale.....	167
II.2.2.3.2. Bases fonctionnelles et sélectivité de la réaction suicide entre LIFpg et le cHyd.....	174
II.2.2.4. Conclusion et discussion.....	177
<b>III. Conclusion générale et perspectives.....</b>	<b>183</b>
<b>III.1. Conclusion générale.....</b>	<b>183</b>
<b>III.2. Perspectives.....</b>	<b>192</b>
<b>IV. Matériel et méthodes.....</b>	<b>199</b>
<b>IV.1. Matériel biologique utilisé.....</b>	<b>199</b>
IV.1.1. Mutagenèse dirigée sur LIFpg.....	199
IV.1.2. Production de protéines recombinantes LIFpg WT et mutants.....	201
IV.1.2.1. Cultures de 250 mL pour les extraits cellulaires.....	201
IV.1.2.2. Cultures de 1L pour les purifications protéiques.....	202
IV.1.3. Lyse des cellules et obtention d'extraits cellulaires.....	202
IV.1.4. Détermination de la quantité de protéine dans les extraits cellulaires.....	202

IV.1.5. <b>Purification de <i>L/Fpg</i> WT et mutantes</b> .....	204
IV.1.5.1. Chromatographie Accell QMA (Waters).....	204
IV.1.5.2. Chromatographie Sulphopropyl Sepharose Fast Flow (SPFF, Pharmacia).....	205
IV.1.5.3. Chromatographie Sulphopropyl Sepharose High Performance (SPHP, Pharmacia).....	205
IV.1.5.4. Chromatographie Ultrogel AcA54 (LKB).....	205
IV.1.5.5. Chromatographie HS20 (POROS).....	206
IV.1.5.6. Récapitulatif du suivi de purification.....	206
IV.1.5.7. Dosage et stockage de <i>L/Fpg</i> WT et mutantes.....	207
IV.1.6. <b>Source des autres protéines utilisées</b> .....	208
IV.1.7. <b>Production des oligonucléotides utilisés</b> .....	208
IV.1.8. <b>Purification des oligonucléotides utilisés</b> .....	209
IV.1.9. <b>Marquage au <sup>32</sup>P d'oligonucléotides</b> .....	210
IV.1.10. <b>Hybridation d'oligonucléotides</b> .....	211
<b>IV.2. Etudes structurales de complexes <i>Fpg</i>/ADN contenant une lésion</b> .....	<b>212</b>
IV.2.1. <b>Cristallogénèse sur les complexes <i>Fpg</i>/ADN contenant une lésion</b> .....	212
IV.2.2. <b>Vérification de la qualité des cristaux sur générateur de rayons X à anode tournante</b> .....	214
IV.2.3. <b>Collection et traitement des données de diffraction, affinement des modèles moléculaires et vérification de leur qualité</b> .....	214
IV.2.3.1. Caractéristiques des lignes de lumière utilisées à l'ESRF.....	214
IV.2.3.2. Collection et traitement des données de diffraction.....	215
IV.2.4. <b>Vérification de la qualité et analyse des modèles moléculaires</b> .....	218
IV.2.4.1. Vérification de la qualité stéréochimique des modèles protéiques.....	218
IV.2.4.2. Analyse des structures des complexes <i>Fpg</i> /ADN.....	219
<b>IV.3. Etudes biochimiques de complexes <i>Fpg</i>/ADN contenant une lésion</b> .....	<b>222</b>
IV.3.1. <b>Etudes biochimiques dans des extraits bruts</b> .....	222
IV.3.1.1. Etude de la capacité de fixation de <i>Fpg</i> sur l'ADN lésé dans des extraits bruts.....	222
IV.3.1.2. Etude de l'activité ADN glycosylase de <i>Fpg</i> sur l'ADN lésé dans des extraits bruts.....	223
IV.3.2. <b>Etudes biochimiques avec des protéines purifiées</b> .....	226
IV.3.2.1. Etude de la capacité de fixation de <i>Fpg</i> pure sur l'ADN lésé.....	226
IV.3.2.2. Etude de l'activité d'ADN glycosylases pures sur l'ADN lésé.....	230
IV.3.2.3. Etude de l'inhibition de l'activité ADN glycosylase de <i>Fpg</i> pure sur le nHyd par son analogue cHyd.....	234
IV.3.2.4. Etude de la formation d'un complexe covalent entre <i>Fpg</i> et la 5-OH-5-Me-Hydantoïne.....	235
IV.3.2.4.1. Test SDS-PAGE-TRAP pour la mise en évidence de pontages covalents entre <i>L/Fpg</i> et l'Hyd ou le cHyd.....	235
IV.3.2.4.2. Mise en évidence et caractérisation préliminaire du complexe covalent.....	236
IV.3.2.4.3. Analyse comparative de la spécificité de la réaction de pontage et de l'activité Hyd-ADN glycosylase de <i>L/Fpg</i> .....	239
IV.3.2.4.4. Bases fonctionnelles et sélectivité de la réaction suicide entre <i>L/Fpg</i> et le cHyd.....	243
<b>Solutions tampons et milieux utilisés</b> .....	<b>245</b>
<b>Références bibliographiques</b> .....	<b>251</b>

## Abréviations

- 8-oxo-G : 7,8-dihydro-8-oxo-guanine, lésion oxydative sur une guanine.
- ADN : Acide Désoxyribo-Nucléique.
- APE : AP Endonuclease.
- BER : « Base Excision Repair », soit réparation par excision de base.
- BSA : « Bovine Serum Albumine », soit albumine de sérum bovin.
- *BstFpg* : Formamidopyrimidine-ADN glycosylase de *Bacillus stearothermophilus*.
- Bz-cFapyG : modèle de lésion FapyG substitué en N7 par un groupement benzyle (non métabolisable car greffé sur un cyclopentane).
- Bz-FapyG : modèle de lésion FapyG substitué en N7 par un groupement benzyle (métabolisable car greffé sur un désoxyribose).
- CCD : « Charge-Coupled Device », soit dispositif à transfert de charge (capteur photographique).
- cFapyG : analogue carbocyclique du FapyG, non métabolisable par les ADN glycosylases car greffé sur un cyclopentane.
- cHyd : analogue carbocyclique de l'Hyd, non métabolisable par les ADN glycosylases car greffé sur un cyclopentane.
- cps : coups par seconde (unité de mesure de radioactivité).
- dATP : désoxyAdénosine Tri Phosphate.
- dCTP : désoxyCytidine Tri Phosphate.
- dGTP : désoxyGuanosine Tri Phosphate.
- dNTP : désoxyNucléotide Tri Phosphate.
- dTTP : désoxyThymidine Tri Phosphate.
- DHU : DiHydroUracile, lésion oxydative pyrimidique.
- dRP : DésoxyRibose Phosphate.
- DO : Densité Optique.
- DSB : « Double Strand Break », soit coupure double brin de l'ADN.
- DTT : Di-Thio-Thréitol (agent réducteur fort).
- *EcFpg* : Formamidopyrimidine-ADN glycosylase de *Escherichia coli*.
- *EcNth* : Endonucléase III de *Escherichia coli*.
- EDTA : « Ethylene-Diamine-Tetraacetic-Acid », soit acide éthylène-diamine-tétraacétique (chélateur de cations divalents).
- ESRF : « European Synchrotron Ray Facility », soit installation européenne de rayonnement synchrotron.
- FapyG : 2,6-diamino-4-hydroxy-5-Formamidopyrimidine, lésion oxydative sur une guanine.
- Fpg : Formamidopyrimidine-ADN glycosylase.
- HEPES : 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique (bon tampon à pH physiologique).
- HPLC-GC/MS : couplage d'une chromatographie phase liquide haute pression à une chromatographie phase gazeuse et une spectrométrie de masse.
- Hyd : 5-Hydroxy-5-Méthyle-Hydantoïne, lésion oxydative sur une thymine.
- IPTG : Isopropyl-β-D-Thio-Galactopyranoside (analogue synthétique très stable du lactose).
- kDa : kilodalton (unité de mesure de masse moléculaire).
- $K_{D\text{ app}}$  : Constante d'équilibre de dissociation apparente.
- LB : Luria-Bertani, milieu de culture bactérienne.
- *LlFpg* : Formamidopyrimidine-ADN glycosylase de *Lactococcus lactis*.
- LP-BER : « Long Patch-Base Excision Repair », soit réparation par excision de bases à longue brèche.
- MMR : « MisMatch Repair », soit réparation des mésappariements.
- Nei : ADN glycosylase procaryote, aussi appelée Endonucléase VIII.
- NER : « Nucleotide Excision Repair », soit réparation par excision de nucléotide.
- NHEJ : « Non Homologous End Joining », soit ligation des extrémités non homologues.
- Nth : ADN glycosylase procaryote, aussi appelée Endonucléase III.
- NZY+ : milieu de culture bactérienne.
- Ogg1 : Oxo-guanine-glycosylase 1.

- PCR : « Polymerase Chain Reaction », soit réaction en chaîne par polymérase.
- PDB : « Protein Data Bank », fait référence à la banque de données de structures de macromolécules biologiques (<http://www.rcsb.org/pdb/home/home.do>).
- pH : potentiel Hydrogène (unité de mesure de l'acidité).
- PMSF : Phényle Méthane Sulfonyle Fluoride (inhibiteur de protéases à sérines).
- RH : Recombinaison Homologue.
- RMN : Résonance Magnétique Nucléaire.
- RMSD : « Root Mean Square Deviation », soit l'écart type sur la position des atomes.
- ROS : « Reactive Oxygen Species », soit espèces réactives de l'oxygène.
- rpm : rotation par minute.
- SAM : S-Adénosyle-L-méthionine.
- SDS-PAGE : « Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis », soit électrophorèse en gel de polyacrylamide en présence de sodium dodécyle sulfate.
- SHM : « Somatic HyperMutation », soit hypermutation somatique.
- SIDA : Syndrome d'ImmunoDéficiency Acquis.
- Site AP : Site APurique ou APyrimidique, équivalent à un site abasique.
- SN1 et SN2 : Substitution Nucléophile de type 1 et 2.
- SP-BER : « Short Patch-Base Excision Repair », soit réparation par excision de bases à courte brèche.
- SSB : « Single Strand Break », soit coupure simple brin de l'ADN.
- TBE : Tris Borate EDTA.
- TCEP : Tris-CarboxyÉthyle-Phosphine (agent réducteur fort).
- TE : Tris EDTA.
- TEMED : TEtra-Méthyle-Ethylène-Diamine (catalyseur pour la polymérisation des gels de polyacrylamide).
- THF : Tetra-Hydro-Furane, analogue de site abasique non métabolisable par les ADN glycosylases.
- TLS : « TransLesion Synthesis », soit synthèse translésionnelle (par une ADN polymérase).
- Tm : « melting Temperature », soit température de fusion.
- Tris : Tris-(hydroxyméthyle)-aminométhane (bon tampon à pH physiologique).
- TtFpg : Formamidopyrimidine-ADN glycosylase de *Thermus thermophilus*.
- VIH : Virus d'Immunodéficience Humaine.
- WT : « Wild Type », fait référence à une enzyme de type sauvage.

## Index des Tableaux et Figures

### Tableaux

#### I. Introduction

- Tableau 1 : Défauts dans le système de réparation par excision de base et maladies humaines (p. 21).
- Tableau 2 : Comparaison des voies de réparation nucléaires et mitochondriales chez les mammifères (p. 33).
- Tableau 3 : Principales ADN glycosylases humaines et bactériennes (p. 39).
- Tableau 4 : Protéines métabolisant les sites AP et extrémités de l'ADN (p. 47).
- Tableau 5 : Classification des ADN polymérases (p. 49).
- Tableau 6 : Activité de MutT de *E. coli* et de ses homologues fonctionnels humains (p. 65).
- Tableau 7 : Différences d'activité de *EcFpg* en fonction de la base opposée à la lésion pour différentes lésions (p. 78).

#### II. Résultats

- Tableau 8 : Constantes de dissociations apparentes de Fpg pour les différents appariements possibles du cFapyG et du Bz-cFapyG (p. 103).
- Tableau 9 : Récapitulatif des analyses fonctionnelles des mutants de *LIFpg* (p. 118).
- Tableau 10 : Constantes de dissociations apparentes des *LIFpgs* WT ou mutantes E2Q et E2D pour le duplexe c8-oxo-G/C de 14 paires de bases (p. 119).
- Tableau 11 : Statistiques de collection des données de diffraction (complexe non-covalent) (p. 137).
- Tableau 12 : Statistiques sur l'affinement et sur le modèle moléculaire (complexe non-covalent) (p. 138).
- Tableau 13 : Différences d'activité de Nth en fonction de la base opposée à la lésion pour différentes lésions (p. 156).
- Tableau 14 : Statistiques de collection des données de diffraction (complexe covalent) (p. 168).
- Tableau 15 : Statistiques sur l'affinement et sur le modèle moléculaire (complexe covalent) (p. 169).
- Tableau 16 : Angles de torsion du sucre et de la liaison *N*-glycosidique de la lésion cHyd dans les complexes covalent ou non (p. 172).
- Tableau 17 : Angles de torsion du cycle de la Proline *N*-terminale de Fpg dans les complexes covalent ou non (p. 173).

#### III. Conclusion générale et perspectives

- Tableau 18 : Différences d'activité de *EcFpg* en fonction de la base opposée à la lésion pour différentes lésions (p. 190).

#### IV. Matériel et méthodes

- Tableau 19 : Tableau récapitulatif des quantifications de Fpg et des protéines totales dans les extraits bruts (p. 203).
- Tableau 20 : Oligonucléotides utilisés pour nos études (p. 208).

### Figures

#### I. Introduction

- Figure 1 : Représentation schématique de la composition chimique d'un nucléotide (p. 11).
- Figure 2 : Les bases nucléiques de l'ADN (p. 11).
- Figure 3 : Appariements canoniques de l'ADN (p. 12).
- Figure 4 : Conformations canoniques de l'ADN (p. 12).
- Figure 5 : Principaux sites de modifications sur l'ADN (p. 13).
- Figure 6 : Principales modifications de bases nucléiques (p. 14).
- Figure 7 : Site AP sous ses deux formes en équilibre (p. 15).

- Figure 8 : Désaminations hydrolytiques spontanées des bases nucléiques (p. 16).
- Figure 9 : Différentes formes de FapyG substitué en N7 (p. 17).
- Figure 10 : Formation de FapyG et 8-oxo-G par action de radical •OH, d'oxygène singulet et d'oxydants mono-électroniques sur la guanine (p. 18).
- Figure 11 : La 5-hydroxy-5-méthyle-hydantoïne (p. 19).
- Figure 12 : Le cercle vicieux du vieillissement et de l'apparition de mutations (p. 21).
- Figure 13 : Représentation schématique du système BER procaryote et eucaryote (p. 30).
- Figure 14 : Représentation schématique de l'interactome du BER et de la réparation des coupures simple-brin (p. 31).
- Figure 15 : Mécanisme des ADN glycosylases mono- et bi-fonctionnelles (p. 36).
- Figure 16 : Représentations de structures libres ou complexées à l'ADN de différentes protéines des trois principales superfamilles structurales d'ADN glycosylases (p. 37).
- Figure 17 : Mécanismes de réaction et inhibiteurs des ADN glycosylases (p. 43).
- Figure 18 : Représentations des appariements de la 8-oxo-G et comparaison avec les appariements canoniques de l'ADN (p. 54).
- Figure 19 : Première partie du système GO de *E. coli* (p. 60).
- Figure 20 : Seconde partie du système GO de *E. coli* (p. 61).
- Figure 21 : Effet de l'appariement 8-oxo-G/G sur la mutagénèse associée au traitement de cette paire de bases par les enzymes du système GO (p. 64).
- Figure 22 : Alignement de séquence des Fpg issues de *Lactococcus lactis*, *Escherichia coli*, *Thermus thermophilus* et *Bacillus stearothermophilus* (p. 68).
- Figure 23 : Schéma de l'organisation des structures secondaires de la Fpg de *Lactococcus lactis* (p. 69).
- Figure 24 : Vue stéréoscopique de la structure du complexe formé entre Fpg et un ADN (p. 69).
- Figure 25 : Superposition du squelette des Ca de la *Tt*Fpg libre et de P1GE2Q-*Lt*Fpg liée à un ADN contenant un site THF (p. 70).
- Figure 26 : Structure de l'ADN contenant un site THF complexé à P1GE2Q-*Lt*Fpg (p. 71).
- Figure 27 : Vue stéréoscopique de la triade d'intercalation M75, R109, F111 de *Lt*Fpg dans l'ADN (p. 72).
- Figure 28 : Représentation schématique du réseau de liaisons hydrogène entre P1GE2Q-*Lt*Fpg et l'ADN contenant un site THF (p. 73).
- Figure 29 : Substrats de l'activité ADN glycosylase de Fpg dérivés de purines (p. 74).
- Figure 30 : Substrats de l'activité ADN glycosylase de Fpg dérivés de pyrimidines (p. 75).
- Figure 31 : Substrats de l'activité AP lyase de Fpg (p. 77).
- Figure 32 : Représentation schématique du mécanisme catalytique de l'activité ADN glycosylase de Fpg (p. 82).
- Figure 33 : Représentation schématique du mécanisme catalytique de l'activité AP lyase de Fpg (p. 84).

## II. Résultats

- Figure 34 : Structure plane des modèles de lésion cFapyG et Bz-cFapyG (p. 88).
- Figure 35 : Gel en conditions natives (gel-retard) correspondant au titrage de la fixation de Fpg sur un duplexe cFapyG/C (p. 102).
- Figure 36 : Evolution des fractions de ligands libres en fonction de la concentration en protéine Fpg pour les différents contextes d'appariement du cFapyG et du Bz-cFapyG (p. 103).
- Figure 37 : Comparaison des constantes de dissociation apparentes obtenues avec Fpg sur le cFapyG ou le Bz-cFapyG (p. 104).
- Figure 38 : Vue stéréoscopique d'une superposition des ADNs contenant les analogues de lésions cFapyG ou Bz-cFapyG extrahélicaux dans les complexes *Lt*Fpg/ADN (p. 105).
- Figure 39 : Reconstruction de l'adduit méthyle en N7 du FapyG (p. 107).
- Figure 40 : Vue stéréoscopique des résidus de *Lt*Fpg impliqués dans la stabilisation de la partie base nucléique du cFapyG (p. 113).
- Figure 41 : Vue stéréoscopique des résidus impliqués dans la stabilisation des deux phosphates entourant la lésion cFapyG et dans la stabilisation de la cytosine opposée à la lésion (p. 113).
- Figure 42 : Structure plane de la 8-oxo-G et de son analogue c8-oxo-G (p. 114).
- Figure 43 : Gel en conditions natives correspondant à la fixation de Fpg WT ou mutantes sur un duplexe c8-oxo-G/C ou G/C (p. 115).

- Figure 44 : Gel dénaturant l'ADN correspondant aux tests d'activité des mutants de Fpg dans les extraits bruts sur le duplexe 8-oxo-G/C (p. 117).
- Figure 45 : Evolution de la coupure en 20 min et à 37°C par *LIFpg* WT ou mutantes E2Q et E2D d'un duplexe 8-oxo-G/C de 24 paires de bases (p. 120).
- Figure 46 : Stabilisation de la forme ouverte du site AP par une liaison hydrogène avec le résidu E2 de Fpg (p. 124).
- Figure 47 : Réseaux de liaisons hydrogène respectivement établis par les résidus E2 ou Q2 des structures *LIFpg*WT/*cFapyG* et *BstFpg*E2Q/8-oxo-G (p. 125).
- Figure 48 : Vue stéréoscopique de l'extrémité N-terminale de Fpg dans la structure de Fpg complexée au *cFapyG* (p. 127).
- Figure 49 : Hypothèses quant au mécanisme permettant l'activation par déprotonation du résidu P1 de Fpg (p. 128).
- Figure 50 : Formation de la 5-Hydroxy-5-Méthyle-Hydantoïne dans l'ADN (p. 130).
- Figure 51 : Cinétique d'excision de la 8-oxo-G et de l'Hyd dans différents contextes oligonucléotidiques (p. 132).
- Figure 52 : Comparaison des activités Hyd-ADN glycosylase de *LIFpg* et *EcNth* (p. 133).
- Figure 53 : L'analogue de lésion *cHyd* (p. 134).
- Figure 54 : Inhibition de l'activité Hyd-ADN glycosylase de *LIFpg* par un oligonucléotide contenant un analogue *cHyd* ou par un oligonucléotide sans lésion (p. 135).
- Figure 55 : Analyse de l'interaction de *LIFpg* et *EcNth* avec l'analogue carbocyclique *cHyd* (p. 136).
- Figure 56 : Diagramme de Ramachandran sur la structure de Fpg dans le complexe *LIFpg/cHyd* non-covalent (p. 139).
- Figure 57 : Représentation du modèle structural non-covalent obtenu entre *LIFpg* et le duplexe *cHyd/C*, montrant la lésion 5-OH-5-Me-Hydantoïne et le résidu P1 de Fpg (p. 139).
- Figure 58 : Superposition des squelettes protéiques de *LIFpg* complexée à un *cFapyG* ou à un *cHyd* (p. 140).
- Figure 59 : Superposition des ADNs contenant les lésions *cFapyG* et *cHyd* issues des structures des complexes *LIFpg/cFapyG* et *LIFpg/cHyd* (p. 142).
- Figure 60 : Comparaison des paramètres structuraux des analogues de lésions *cFapyG* et *cHyd* dans les complexes formés avec *LIFpg* (p. 143).
- Figure 61 : Réseau de liaisons hydrogène établi entre *LIFpg* et la partie base nucléique du *cFapyG* (p. 145).
- Figure 62 : Réseau de liaisons hydrogène établi entre *LIFpg* et la partie base nucléique du *cHyd* (p. 146).
- Figure 63 : Reconnaissance par une arginine de Fpg de la de la cytosine opposée au dommage (p. 149).
- Figure 64 : Reconstruction d'une thymine en *anti* dans notre structure de *LIFpg/cHyd* (p. 150).
- Figure 65 : Représentation de la thymine en *anti* dans la structure de *BstFpg* complexée à un site AP réduit opposé à une thymine (p. 151).
- Figure 66 : Reconstruction d'une guanine ou d'une adénine en *anti* dans notre structure de *LIFpg/cHyd* : apparition d'un problème stérique (p. 152).
- Figure 67 : Représentation de la guanine en *syn* dans la structure de *BstFpg* complexée à un site AP réduit opposé à une guanine (p. 153).
- Figure 68 : Reconstruction d'une adénine en *syn* dans la structure de *BstFpg* complexée à un ADN contenant un site AP réduit opposé à une guanine (p. 154).
- Figure 69 : Reconnaissance de la base opposée au dommage par *BstNth* (p. 157).
- Figure 70 : Mise en évidence d'un complexe covalent entre *LIFpg* et le duplexe *cHyd/C* (p. 159).
- Figure 71 : Effet de la température d'incubation et de la nature mono- ou bi-caténaire de l'ADN contenant le *cHyd* sur la formation du complexe covalent avec *LIFpg* (p. 160).
- Figure 72 : Cinétiques de formation de complexe covalent entre *LIFpg* et le duplexe *cHyd/C* à 37°C (p. 161).
- Figure 73 : Effet de la présence ou non de l'étape de dénaturation thermique du SDS-PAGE-TRAP sur la stabilité du complexe covalent formé entre *LIFpg* et le duplexe *cHyd/C* ou le *cHyd* simple-brin (p. 162).
- Figure 74 : Effet du pH sur la stabilité du complexe covalent formé entre *LIFpg* et le duplexe *cHyd/C* (p. 163).
- Figure 75 : Le complexe covalent formé entre *LIFpg* et la 5-Hydroxy-5-Méthyle-Hydantoïne n'est pas un intermédiaire de type base de Schiff (p. 164).

- Figure 76 : Effet du pH sur l'activité Hyd-ADN glycosylase de *L/Fpg* pour le duplexe Hyd/C, et sur la formation de complexe covalent entre *L/Fpg* et le duplexe cHyd/C (p. 166).
- Figure 77 : Effet du contexte d'appariement et de la nature désoxyribonucléique ou carbocyclique de la 5-Hydroxy-5-Méthyle-Hydantoïne sur l'efficacité de la réaction de pontage covalent avec *L/Fpg* (p. 167).
- Figure 78 : Diagramme de Ramachandran sur la structure de *L/Fpg* dans le complexe covalent formé avec le duplexe cHyd/C (p. 170).
- Figure 79 : Superposition des structures globales de *L/Fpg* complexée covalamment ou non à un cHyd (p. 170).
- Figure 80 : Représentation du modèle structural covalent obtenu entre *L/Fpg* et le duplexe cHyd/C, montrant la lésion 5-OH-5-Me-Hydantoïne et le résidu P1 de Fpg (p. 171).
- Figure 81 : Vue stéréoscopique des différences structurales entre les complexes covalent ou non au voisinage de la lésion cHyd (p. 172).
- Figure 82 : Nomenclature arbitraire utilisée pour les angles de torsion du cycle de la Proline (p. 173).
- Figure 83 : Analyse fonctionnelle par mutagénèse de l'efficacité de la réaction suicide entre *L/Fpg* et le duplexe cHyd/C (p. 174).
- Figure 84 : Sélectivité de la réaction suicide entre *L/Fpg* et le cHyd (p. 175).
- Figure 85 : Vue stéréoscopique de la localisation du résidu P1 de Fpg par rapport au cHyd dans le complexe non-covalent (p. 178).
- Figure 86 : Représentation schématique du mécanisme d'épimérisation de la 5-OH-5-Me-Hydantoïne (p. 178).
- Figure 87 : Mode de reconnaissance stéréospécifique du cHyd par *L/Fpg* (p. 179).
- Figure 88 : Représentation schématique des mécanismes réactionnels d'attaque du cHyd en C1' ou en C5, aboutissant respectivement à la formation d'un pontage covalent entre *L/Fpg* et l'Hyd ou à l'activité Hyd-ADN glycosylase (p. 181).

### III. Conclusion générale et perspectives

- Figure 89 : Positionnement du résidu P1 de Fpg par rapport aux différentes lésions ou analogues de lésions pour lesquels on a la structure d'un complexe avec l'enzyme (p. 184).
- Figure 90 : Présence d'une liaison hydrogène entre l'azote de P1 et un oxygène de certains analogues de lésions (p. 185).
- Figure 91 : Différentes possibilités de formation d'une liaison hydrogène entre l'amine N-terminale de *L/Fpg* et un oxygène du Bz-FapyG ou de l'Hyd (p. 186).
- Figure 92 : Vue stéréoscopique montrant la présence d'une molécule d'eau à proximité du résidu E2 de Fpg dans la structure avec le cFapyG, et pas avec le cHyd (p. 188).
- Figure 93 : Boucle  $\alpha$ F- $\beta$ 9 de *BstFpgE2Q* en conformation « fermée » dans les complexes avec des ADN contenant une 8-oxo-G ou une DHU (p. 189).

### IV. Matériel et méthodes

- Figure 94 : Gel SDS-PAGE représentant l'évolution de la quantité de *L/Fpg* WT produite en fonction du temps d'induction (p. 201).
- Figure 95 : Gel SDS-PAGE pour la quantification de *L/Fpg* WT dans les extraits bruts (p. 203).
- Figure 96 : Bilan de la purification de E2D-*L/Fpg* (p. 207).
- Figure 97 : Chromatogramme représentatif de la purification d'un oligonucléotide 14-merc-C (p. 210).
- Figure 98 : Cristallogénèse par la méthode de diffusion de vapeur en goutte suspendue (p. 213).
- Figure 99 : Cristal obtenu avec le complexe Fpg/(cHyd/C) à 37°C (p. 213).
- Figure 100 : Nomenclature des angles de torsion du squelette phosphate-sucre de l'ADN (p. 219).
- Figure 101 : Cercle de pseudorotation du désoxyribose dans l'ADN (p. 220).
- Figure 102 : Formalisme utilisé pour la définition de la conformation d'une base nucléique par rapport au désoxyribose dans l'ADN (p. 220).
- Figure 103 : Normes utilisées pour la définition de liaisons hydrogène (p. 221).
- Figure 104 : Principe du test d'activité ADN glycosylase (p. 224).

# **I. Introduction**

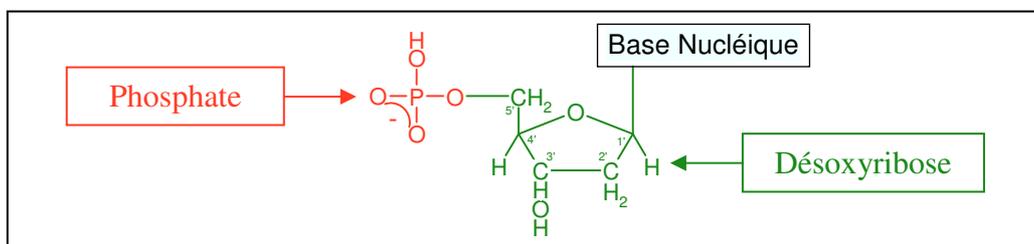
# I. Introduction

## I.1. L'ADN

L'acide désoxyribonucléique, ou ADN, est une macromolécule que l'on retrouve dans tous les organismes vivants, et qui constitue le support de l'information génétique. L'ADN est donc responsable de toutes les fonctions internes des organismes, étant donné qu'il code pour les ARN et les protéines. La pérennité de chaque espèce vivante est assurée par la stabilité et la fidélité de transmission de l'information génétique portée par l'ADN. Cependant, l'information portée par l'ADN peut se modifier au cours du temps, à cause d'erreurs lors de sa réplication ou de recombinaisons génétiques, ce qui permet l'évolution des espèces.

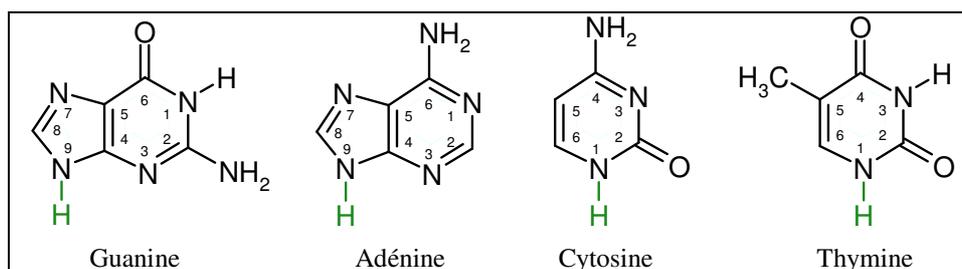
L'ADN est un polymère constitué de quatre nucléotides qui sont répétés de façon non aléatoire, donnant ainsi la séquence nucléotidique de l'ADN. Les nucléotides sont constitués de trois parties liées covalamment entre elles (Figure 1) :

- Un groupement phosphate (en rouge dans la Figure 1), qui est responsable de la charge négative de l'ADN.
- Un sucre, le 2'-désoxy-β-D-ribose ou désoxyribose (en vert dans la Figure 1).
- Une base nucléique (en noir dans la Figure 1).



**Figure 1 :** Représentation schématique de la composition chimique d'un nucléotide.

Les bases nucléiques sont présentes au nombre de quatre dans l'ADN, deux bases puriques, la guanine (G) et l'adénine (A), et deux bases pyrimidiques, la cytosine (C) et la thymine (T) (Figure 2).

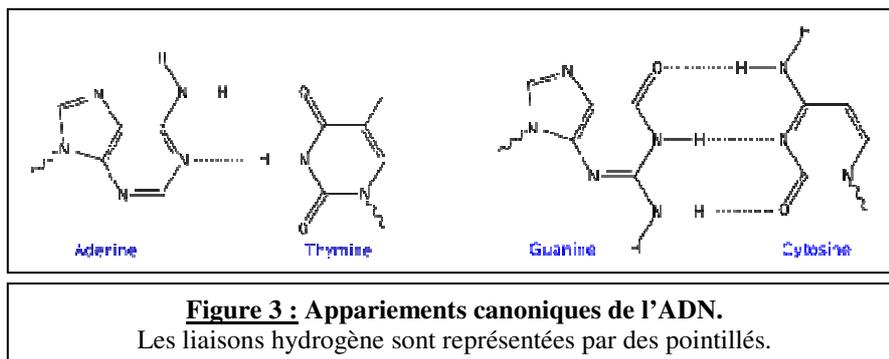


**Figure 2 :** Les bases nucléiques de l'ADN.

La position de la liaison avec le désoxyribose dans l'ADN est indiquée en vert.

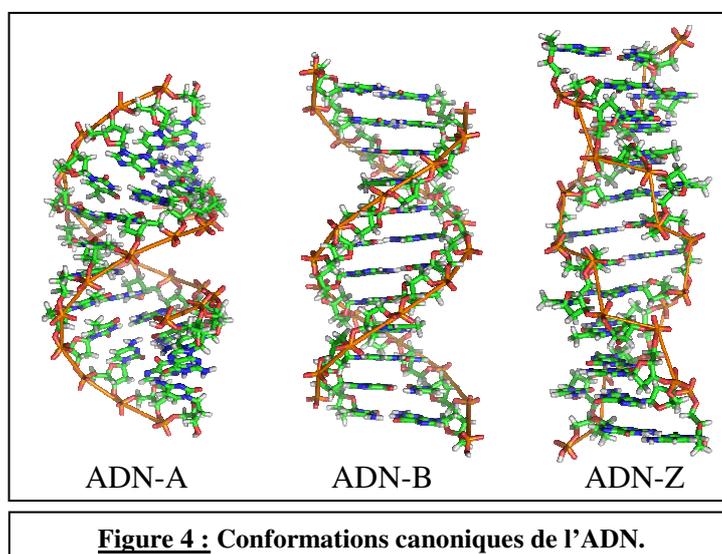
Dans l'ADN, les nucléotides sont reliés entre eux *via* une liaison covalente entre le groupement phosphate d'un nucléotide et le désoxyribose d'un autre nucléotide, la liaison 3'-5'-phosphodiester. La notation 5' et 3' fait référence à la numérotation du ribose (Figure 1).

L'ADN est en fait composé de deux brins antiparallèles se faisant face, l'extrémité 5' d'un brin faisant face à l'extrémité 3' de l'autre brin, et inversement. Ceci est possible car les nucléotides trouvés dans un brin possèdent des nucléotides complémentaires dans l'autre brin avec lesquels ils interagissent *via* des liaisons hydrogène, l'adénine étant complémentaire de la thymine, et la guanine étant complémentaire de la cytosine (Figure 3).



Les deux brins de l'ADN forment une double hélice qui peut prendre plusieurs conformations différentes (Figure 4) :

- L'ADN B, qui est la forme la plus commune de l'ADN, et qui présente un grand et un petit sillon.
- L'ADN A, qui correspond à la forme de l'ADN quand celui-ci est apparié à un ARN lors de la transcription, ou lorsque l'on déshydrate la forme B de l'ADN.
- L'ADN Z, seule forme de l'ADN à posséder une hélice gauche, relativement rare dans l'ADN et dont la fonction est encore floue.

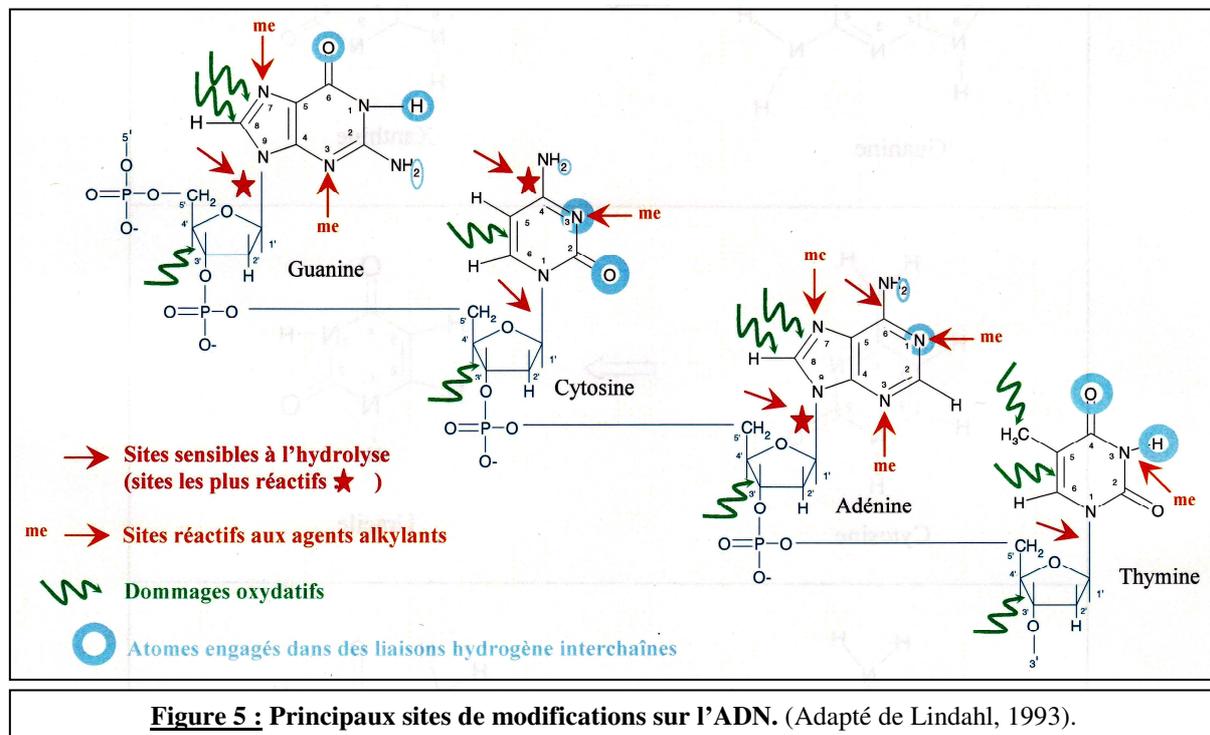


De plus, l'ADN peut prendre d'autres formes, telles que des triples hélices et des quadruplexes, mais ces structures ne seront pas détaillées ici.

## I.2. Les agressions sur les acides nucléiques et les dommages générés

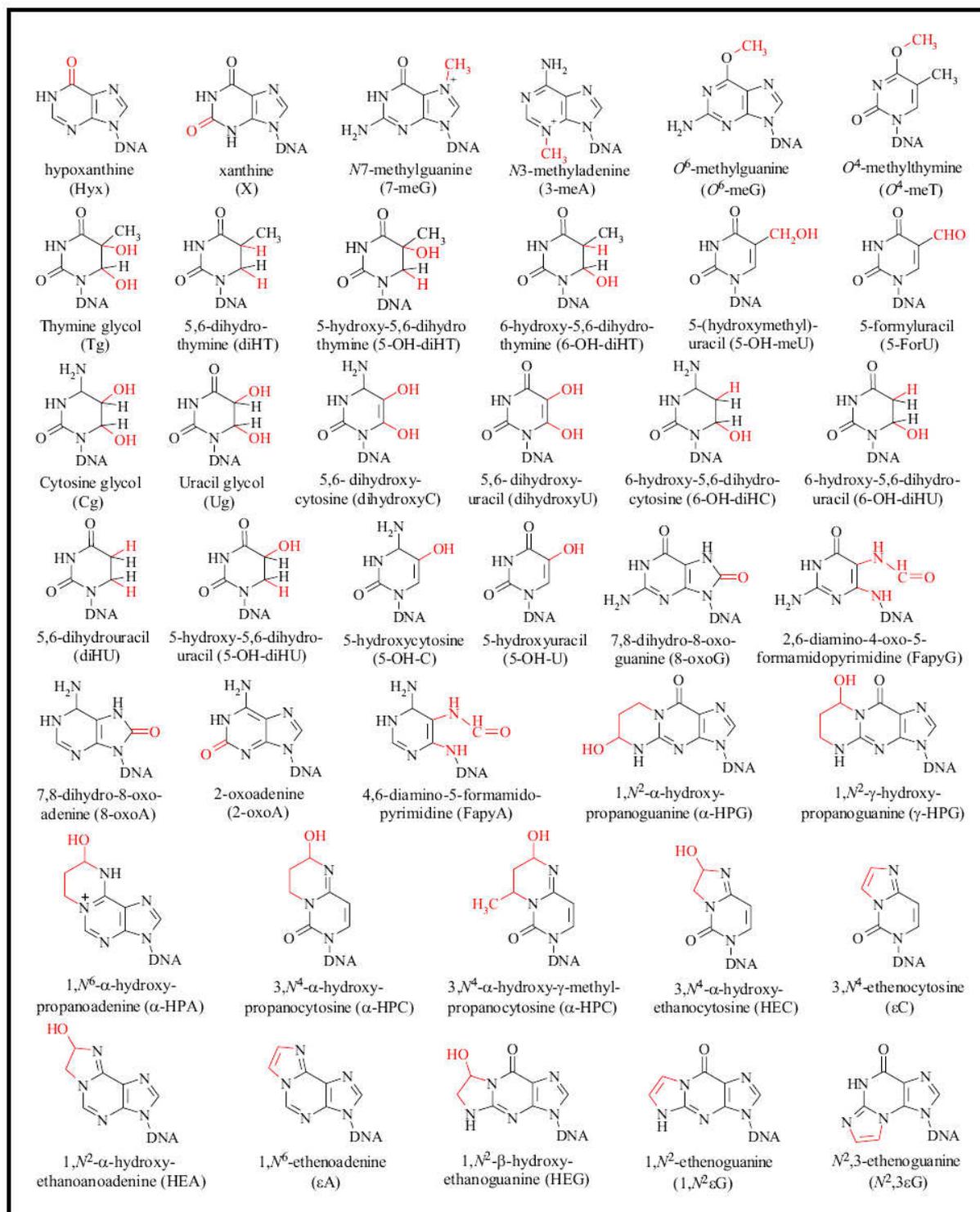
### I.2.1. Généralités

L'ADN peut être modifié par des agents physiques (rayonnements UV,  $\gamma$ ...) ou chimiques (agents alkylants, oxydants...) d'origine endogène et exogène, amenant des modifications chimiques sur sa structure. Les trois principales modifications sur l'ADN correspondent à des hydrolyses, des alkylations et des oxydations, qui peuvent intervenir à différents niveaux des bases nucléiques et du squelette phosphate-sucre de l'ADN (Figure 5).



Une pléthore de lésions peut donc être générée sur l'ADN, comprenant des lésions modifiées sur les bases nucléiques, sur le squelette phosphate-sucre, des coupures de brins, des pontages intra- et inter-brins, des pontages ADN-protéines, etc. En effet, près de 100 modifications sur l'ADN ont été identifiées, dont environ 50 modifications de bases (la Figure 6 présente les principales modifications de bases) (Krwawicz *et al.*, 2007). Cependant, beaucoup d'entre elles n'ont été détectées que *in vitro* après traitement d'ADN par des agents endommageants, les techniques de détections des dommages *in vivo* (souvent des

chromatographies couplées à de la spectrométrie de masse ou à une détection électrochimique) étant trop peu sensibles pour détecter des lésions qui sont souvent présentes en quantités extrêmement faibles dans l'ADN (Cadet *et al.*, 2002).

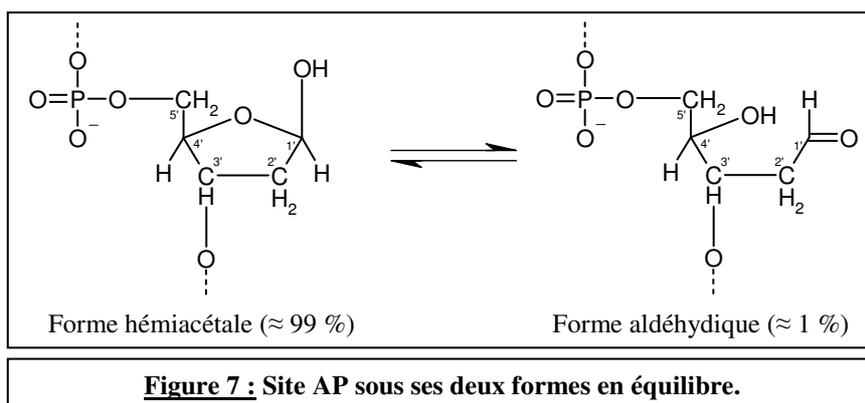


**Figure 6 : Principales modifications de bases nucléiques.**  
Issue de Krwawicz *et al.*, 2007. Les positions modifiées sont indiquées en rouge.

### I.2.1.1. Les hydrolyses

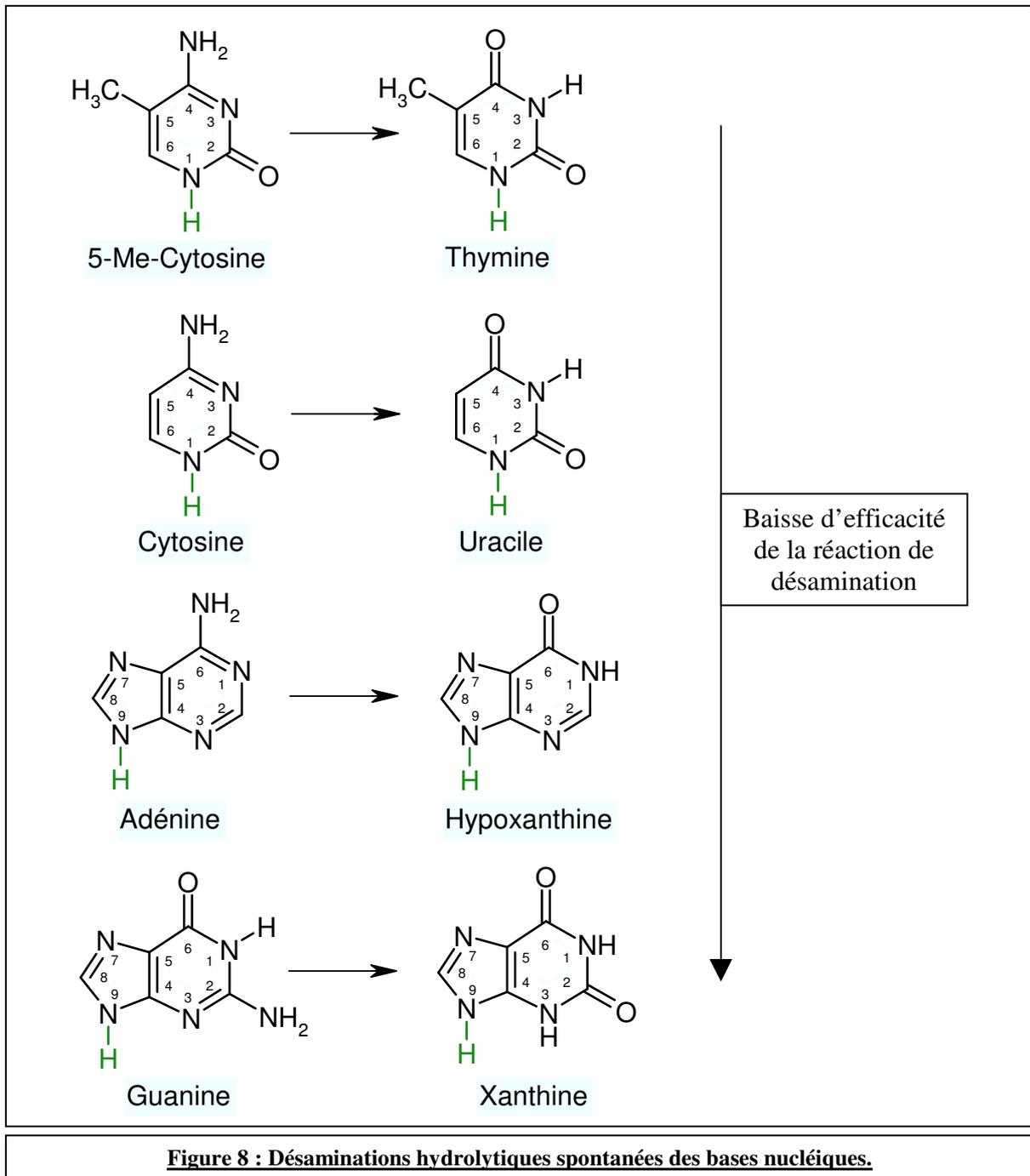
Les hydrolyses sur l'ADN peuvent être séparées en deux catégories, d'une part les hydrolyses de la liaison N-glycosidique entre le sucre et la base nucléique, et d'autre part les désaminations hydrolytiques sur les bases nucléiques (Lindahl, 1993).

Les hydrolyses de la liaison N-glycosidique entraînent des dépurinations et dépyrimidations sur les nucléotides, produisant des sites apuriques ou apyrimidiques (sites AP, voir Figure 7), qui sont des lésions très mutagènes pour les cellules en l'absence de réparation. Ces réactions, bien qu'inefficaces, sont possibles dans les conditions physiologiques de température et de pH, et sont exacerbées en cas d'augmentation de température ou d'acidification du milieu. Les bases puriques (G et A) sont les plus sensibles à ce type d'hydrolyse, avec une efficacité environ 20 fois plus importante que sur les bases pyrimidiques (C ou T). S'ils ne sont pas réparés, les sites AP générés subissent lentement une  $\beta$ -élimination sur leur forme aldéhydique hautement réactive (minoritaire, voir Figure 7), créant ainsi un autre type de lésion sur l'ADN, des coupures simple-brins (Lindahl, 1993).



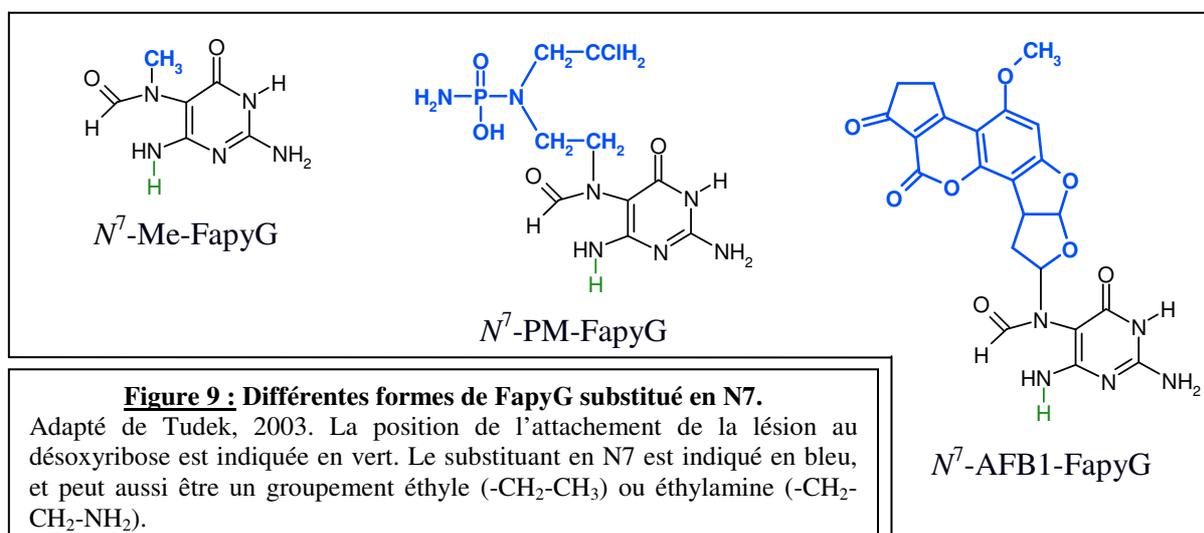
D'autre part, les désaminations hydrolytiques sur les bases nucléiques produisent des lésions de bases nucléiques ou des mésappariements dans l'ADN (Figure 8), et cela même si cette réaction est peu efficace (sauf dans le cas d'une présence significative de  $\text{NO}_2$  ou  $\text{N}_2\text{O}_3$ , dans le cas d'une inflammation par exemple) (Krwawicz *et al.*, 2007). Ainsi, la base nucléique classique la plus sensible à ce type de réaction est la cytosine, qui donne alors un uracile dans l'ADN, ce qui correspond à une lésion car seul l'ARN est sensé en contenir. En fait, la forme méthylée de la cytosine que l'on retrouve dans les génomes, la 5-Me-Cytosine, est légèrement plus susceptible de subir ce type de réaction que la cytosine, et donne une thymine (donc un mésappariement G/T). De plus, l'adénine et la thymine peuvent aussi subir ce type de désamination, donnant respectivement des lésions Hypoxanthine (Hx) et Xanthine (X) (Figure

8), mais avec une efficacité de seulement 2-3 % de celle que l'on a avec la cytosine et la 5-Me-cytosine.



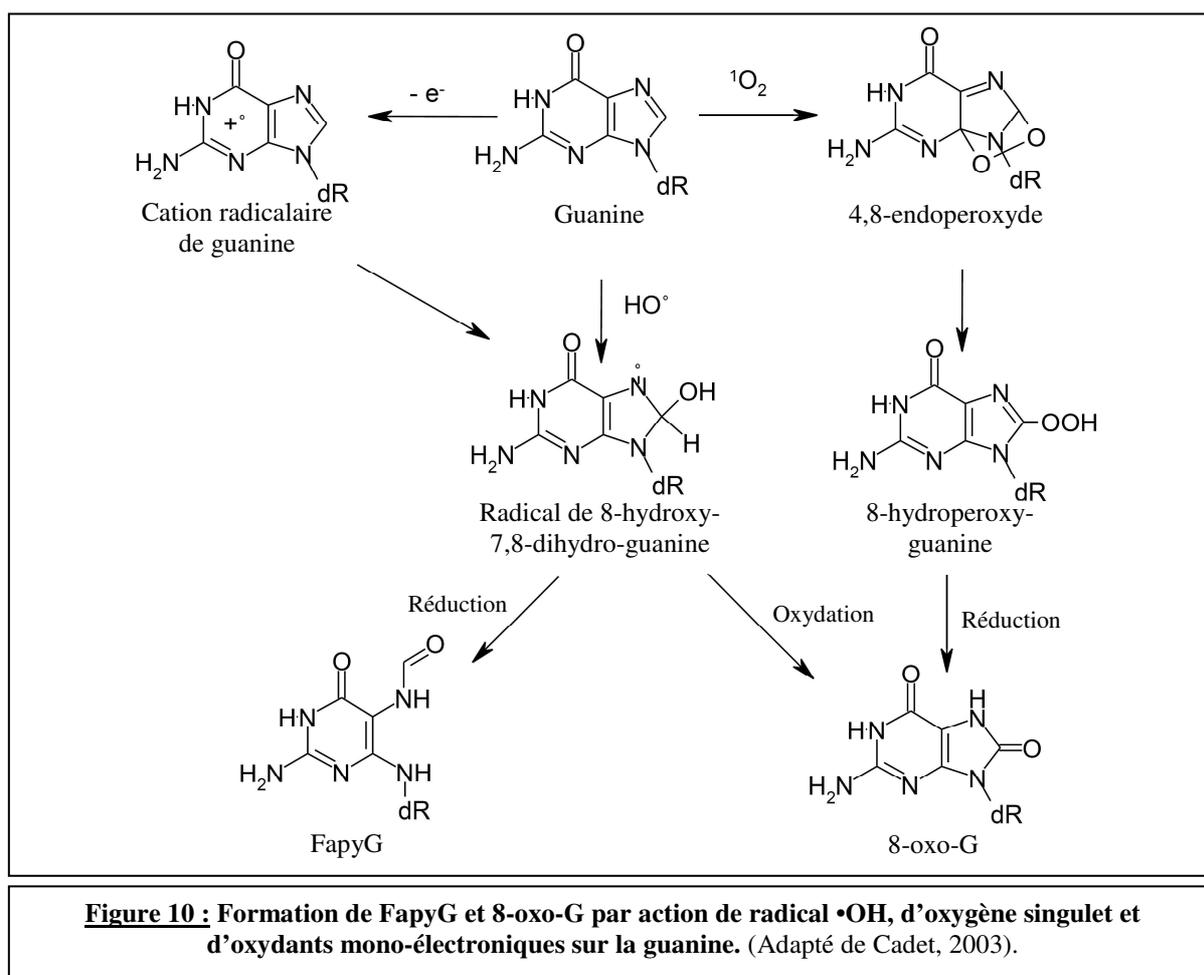
### I.2.1.2. Les alkylations

Des agents alkylants peuvent réagir avec 12 positions différentes sur les bases nucléiques, incluant tous les oxygènes exocycliques et la plupart des azotes des cycles, et peuvent aussi modifier les atomes d'oxygène des groupements phosphate du squelette de l'ADN (Krwawicz *et al.*, 2007). Dépendamment du type de substitution nucléophile utilisé pour l'alkylation, les agents alkylants peuvent être divisés en deux groupes : les agents de type SN1 (ex : *N*-méthyle-*N*-nitrosurée) qui agissent sur les oxygènes et azotes des acides nucléiques, et les agents de type SN2 (ex : méthyle méthane sulfonate, diméthyle sulfate) qui agissent surtout sur les azotes. De façon générale, les azotes sont plus sensibles que les oxygènes à ce type de réaction. Par contre, les azotes engagés dans des liaisons hydrogène sont presque non réactifs dans l'ADN double brin, et sont réactifs dans l'ADN simple brin (ex : le N3 de la cytosine ou le N1 de l'adénine). Les produits majeurs de méthylation sur les bases sont la *N*<sup>7</sup>-méthyle-guanine (*N*<sup>7</sup>-Me-G), puis la *N*<sup>3</sup>-méthyle-adénine (*N*<sup>3</sup>-Me-A) (Figure 6) (Lindahl *et al.*, 1993; Krwawicz *et al.*, 2007). On peut noter que la *N*<sup>7</sup>-Me-G est très réactive, et donne spontanément des sites AP et une forme à cycle imidazole ouvert, le *N*<sup>7</sup>-méthyle-Fapy-guanine (*N*<sup>7</sup>-Me-FapyG, le terme Fapy faisant référence à une Formamidopyrimidine, Tudek *et al.*, 1992). C'est d'ailleurs aussi le cas pour les guanines substituées en N7 par de plus gros adduits que le méthyle, donnant facilement des Fapys substitués en N7 (Tudek, 2003, et Figure 9). Des agents endogènes peuvent aussi alkyler des bases nucléiques, le plus connu étant la *S*-adénosyle-L-méthionine (SAM) (Rydberg *et al.*, 1982). Physiologiquement, le SAM est un donneur de méthyle pour de nombreuses transméthylations cellulaires, et est entre autres un cofacteur lors de la méthylation enzymatique des cytosines de l'ADN en position C5, qui participe à la régulation de l'expression des gènes chez les eucaryotes.



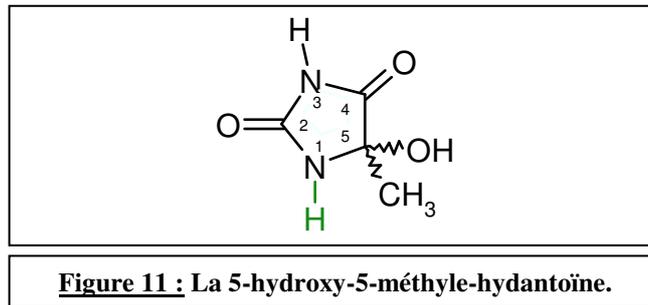
### I.2.1.3. Les oxydations

Les lésions formées par oxydations sur les acides nucléiques sont les plus nombreuses des lésions. Ces oxydations peuvent provenir non seulement de phénomènes exogènes (radiations UV et  $\gamma...$ ), mais aussi de métabolites endogènes, notamment issues de biotransformations de molécules et de la chaîne de transport d'électrons des mitochondries. Ces phénomènes ont tendance à produire des espèces réactives de l'oxygène (ROS) et de l'azote (ex :  $\bullet\text{OH}$ ,  $^1\text{O}_2$ ,  $\text{O}_3$ ,  $\text{NO}\bullet$ ,  $-\text{OONO}$ ) ainsi que des électrons libres, qui oxydent les molécules biologiques, dont les acides nucléiques (Cadet *et al.*, 2002; Krwawicz *et al.*, 2007). Ces attaques sur l'ADN donnent des coupures simple- et double-brins, des sites abasiques, et des lésions sur les sucres et les bases. La base la plus facilement oxydable est la guanine (Costa de Oliveira *et al.*, 1992; Cadet, 2003), donnant de nombreuses lésions, dont celles qui seront souvent citées dans ce manuscrit, la 7,8-dihydro-8-oxo-guanine (8-oxo-G) et la 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) (Tudek, 2003, Figure 10).



**Figure 10 :** Formation de FapyG et 8-oxo-G par action de radical  $\bullet\text{OH}$ , d'oxygène singulet et d'oxydants mono-électroniques sur la guanine. (Adapté de Cadet, 2003).

De plus, une autre lésion a été utilisée durant cette thèse, la 5-hydroxy-5-méthyle-Hydantoïne, qui est issue d'oxydation sur la thymine (Cadet *et al.*, 1978) (Figure 11).



### I.3. Les conséquences des dommages sur l'ADN en l'absence de réparation

Les différents types de lésions présentés précédemment peuvent avoir différents effets sur les cellules en l'absence de réparation. Cette absence de réparation peut venir de plusieurs phénomènes :

- une trop faible activité d'enzymes impliquées dans la réparation (ex : baisse de la quantité de protéine produite, polymorphisme amenant une baisse d'activité...).
- une production trop importante de lésions peut amener une saturation des systèmes de réparation (ex : lors d'un stress oxydant).

La non-réparation de lésions dans l'ADN peut aboutir à différents effets sur les organismes vivants, effets qui seront présentés dans ce chapitre. En l'occurrence, la variété et l'importance de ces effets illustrent bien la nécessité de mieux comprendre les mécanismes d'apparition et de réparation des lésions dans l'ADN, car cette compréhension peut amener des perspectives thérapeutiques dans de nombreux cas.

#### I.3.1. Les conséquences sur les cellules

Tout d'abord, on peut s'intéresser aux effets des dommages non réparés au niveau cellulaire. Ces effets sont majoritairement dus au comportement des ADN polymérase lorsque celles-ci rencontrent des lésions, et plusieurs cas généraux sont à prendre en compte.

Premièrement, certaines lésions ne bloquent pas la réplication, et sont correctement appariées à la base nucléique complémentaire à celle dont est issue la lésion lors de la réplication. Dans ce cas, aucun effet sur les cellules n'est détectable, étant donné que ce type de lésion n'induit pas de mortalité cellulaire ni de mutagenèse. Un exemple de ce type de lésion est la Dihydrothymine (DHT), qui est une lésion issue de thymine, et qui s'apparie uniquement

avec une adénine lors de la réplication, n'induisant donc aucune mutagénèse (Evans *et al.*, 1993).

Ensuite, certaines lésions ne bloquent pas la réplication mais y induisent un mésappariement à une ou plusieurs bases nucléiques différentes de la base complémentaire à celle dont est issue la lésion. Dans ce cas, on a apparition d'une mutagénèse après réplication des zones d'ADN contenant ces lésions, cette mutagénèse pouvant amener différentes conséquences sur les cellules (modifications d'activités cellulaires, mortalité cellulaire...). Un exemple de ce type de lésions est la 8-oxo-G, qui est une lésion issue de guanine, mais qui est connue pour pouvoir s'apparier à une adénine lors de la réplication, induisant donc possiblement des transversions G/C vers T/A lors de leur réplication (chapitre I.5.4 sur les ADN polymérase et leur implication dans la mutagénèse induite par les lésions).

Enfin, certaines lésions bloquent la réplication, ce qui aboutit à un arrêt du cycle cellulaire puis une mortalité cellulaire en l'absence de réparation. Un bon exemple de ce type de lésions correspond aux adduits interbrins de l'ADN (ex : avec le cisplatine), qui sont cytotoxiques car empêchent la réplication en rendant impossible la déshybridation des brins de l'ADN (Dronkert *et al.*, 2001).

En outre, on peut noter que beaucoup de lésions ont des effets intermédiaires à ceux évoqués précédemment. Un bon exemple de ce type de comportement est trouvé avec le  $N^7$ -Me-FapyG, qui est une lésion majoritairement bloquante pour la réplication, mais qui peut aussi amener une légère mutagénèse dans les cas où une ADN polymérase arrive à prendre en charge cette lésion (chapitre 1.5.4 sur les ADN polymérase et leur implication dans la mutagénèse induite par les lésions).

De plus, certaines lésions peuvent avoir d'autres effets que ceux présentés précédemment, par exemple en bloquant la transcription.

### **I.3.2. Les conséquences sur les organismes pluricellulaires**

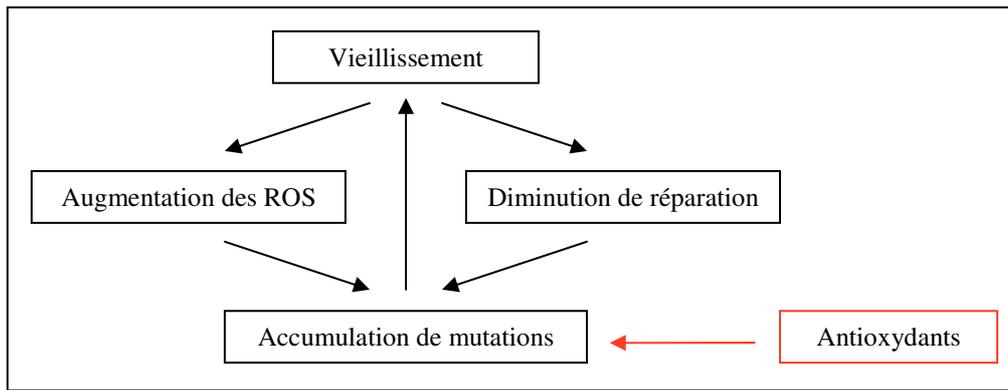
Les conséquences cellulaires évoquées précédemment sont généralement liés à des effets sur les organismes pluricellulaires tels que l'homme. Ce sous-chapitre reviendra donc sur des exemples de conséquences de dommages à l'ADN non réparés chez les organismes pluricellulaires. Un intérêt particulier sera donné aux conséquences des dommages oxydatifs ainsi qu'aux conséquences de déficiences du système de réparation par excision de base (chapitre I.5 sur le système BER), qui est le système de réparation que j'ai étudié pendant ma thèse (Tableau 1).

**Tableau 1 : Défauts dans le système de réparation par excision de base et maladies humaines.**  
(Issu de Wilson III *et al.*, 2007).

Gene	Disease connection
MYH	Reduced-function mutants linked to FAP Increased tumorigenesis in knockout mice when combined with an OGG1 deficiency
OGG1	Possible relationship of specific mutants/variants, or loss of OGG1, with cancer proneness/susceptibility
UNG	Recessive mutations linked to HIGM Deficiency in rat hippocampal neurons results in increased spontaneous apoptosis Knockout mouse embryonic fibroblasts and primary cortical neurons are more susceptible to exposure to a nitric oxide donor and oxygen and glucose-deprivation, respectively Plays an important role during brain ischemia and reperfusion in knockout mice
AAG	Null mouse neuronal cells display increased sensitivity to genotoxic agents
APE1	Putative target of lead-related carcinogenesis Haploinsufficient mice exhibit increased spontaneous mutagenesis, and elevated cancer incidence in the face of oxidative stress Reduced expression in rat hippocampal or sensory neurons results in decreased viability following hydrogen peroxide treatment
POL $\beta$	Human tumors found to express amino acid variants with reduced polymerase fidelity, or deletion or truncation mutants that may act as dominant-negative forms - reduced fidelity variants induce cellular transformation Haploinsufficient mice exhibit elevated mutation frequencies following DNA-damaging agent exposure
XRCC1	Possible relationship of amino acid variants with disease susceptibility Indirectly connected to autosomal recessive spinocerebellar ataxias via interactions with TDP1 and Aprataxin
LIG1	Missense mutations linked to a single, disease patient (46BR fibroblast strain)

### - Vieillesse :

Tout d'abord, on peut dire que la plupart des syndromes de vieillissement prématuré sont associés à des mutations de gènes codant pour des protéines de réparation de l'ADN. C'est notamment le cas pour les systèmes de réparation par excision de base et par excision de nucléotide, et pour le système de réparation des cassures double-brins (Chen *et al.*, 2007; Gorbunova *et al.*, 2007; Wilson III *et al.*, 2007). De plus, on sait désormais qu'il y a accumulation de dommages avec l'âge, avec une augmentation de la vitesse d'accumulation de ces dommages. Ceci est dû à une diminution de la capacité de réparation des organismes avec l'âge, notamment au niveau de la réparation par excision de base (BER), de la réparation des mésappariements (MMR), de la réparation des cassures double-brins (Gorbunova *et al.*, 2007). En outre, une augmentation de production d'espèces réactives de l'oxygène (ROS) avec l'âge a aussi été décrite, ce qui amplifie le phénomène d'accumulation de dommages (Chen *et al.*, 2007). Inversement, tout ce qui lutte contre les stress oxydants ou augmente la capacité de réparation est décrit comme augmentant la longévité (Chen *et al.*, 2007). L'hypothèse que le vieillissement normal des organismes serait causé (au moins en partie) par une baisse de la capacité de réparation des organismes a donc été émise. Ainsi, un schéma général de la relation entre vieillissement et dommages sur l'ADN a été présenté par Gorbunova et collaborateurs en 2007, indiquant que cette relation pourrait probablement être un « cercle vicieux », le vieillissement diminuant la capacité de réparation, d'où l'apparition d'une plus grande quantité de dommages, qui participeraient à leur tour à l'évolution du vieillissement (Figure 12).



**Figure 12 : Le cercle vicieux du vieillissement et de l'apparition de mutations.**  
 Adapté de Gorbunova *et al.*, 2007 et Chen *et al.*, 2007. Les flèches noires indiquent une conséquence. La flèche rouge indique une inhibition de l'accumulation de mutations

### - Cancers

Depuis de nombreuses années, une corrélation entre la présence de dommages oxydatifs à l'ADN, la diminution de la capacité de réparation, et la cancérogénèse a été établie (Tableau 1) (Evans *et al.*, 2004). Cependant, une baisse de l'efficacité de la réparation accompagnée d'une augmentation des dommages oxydatifs peut être liée au processus tumoral lui-même (inflammation, stress oxydatifs, changements des voies de signalisation affectant la réparation et d'autres mécanismes de défense) (Olinski *et al.*, 2003).

Par contre, des études comparatives ont montré qu'il existait une bonne corrélation entre certains polymorphismes d'enzymes impliquées uniquement dans la réparation de l'ADN et des risques de cancer. Par exemple, des variants d'activité réduite de la protéine de réparation Ogg1 (S326C, un allèle fréquent, et R154H, un allèle rare) peuvent constituer des facteurs de risques pour des cancers du poumon, colorectal, œsophagien, orolaryngéal, gastrique, cervical, de la vésicule biliaire et de la vessie (Weiss *et al.*, 2005; Dizdaroglu *et al.*, 2008). De plus, certains polymorphismes de la protéine de réparation MYH semblent aussi être un facteur de risque pour certains cancers héréditaires du colon (Huffman *et al.*, 2005; Weitz *et al.*, 2005). Ces données très intéressantes indiquent sans aucun doute un lien de cause à effet entre les enzymes de réparation et le risque d'apparition de certains cancers. De plus, ceci indique que l'augmentation des dommages oxydatifs sur l'ADN est un facteur de risque important dans de nombreuses formes de cancers, et que l'accumulation des dommages peut avoir une importance dans la prédiction de l'apparition de cancers (Evans *et al.*, 2004).

Finalement, on ne parle que de facteur de risque, car il faut garder à l'esprit que le risque d'apparition de cancer est le produit de la composition génétique complète d'un individu

et de son style de vie (exposition ou non à des agents cancérigènes, alimentation...) (Wilson III *et al.*, 2007).

D'autre part, des agents qui induisent des dommages sur l'ADN et inhibent la réparation de l'ADN sont utilisés pour traiter des tumeurs (agents chimiques tels que la bléomycine, la néocarzinostatine ou la témozolomine, agents physiques tels que la radiothérapie...), jouant sur un effet plus prononcé sur les cellules cancéreuses que sur les cellules saines (Krwawicz *et al.*, 2007). Dans ce cadre, une augmentation de capacité de réparation des cellules traitées est associée à une résistance aux agents anticancéreux, alors qu'une diminution est associée à une hypersensibilité à ces agents (Damia *et al.*, 2007). Ainsi, les mutations dans des gènes de réparation de l'ADN dans certaines tumeurs peuvent être utilisées pour obtenir une plus grande sensibilité aux agents antitumoraux. Par contre, l'utilisation d'agents endommageant l'ADN pour traiter des cancers est susceptible d'induire le développement de cancers secondaires dans les tissus sains, soit un effet secondaire désastreux (Olinski *et al.*, 2002).

### **- Maladies neurodégénératives**

Tout d'abord, la perte de fonction mitochondriale a été clairement démontrée dans plusieurs neuropathologies, notamment celles liées à l'âge telles que les maladies de Parkinson, d'Alzheimer et de Huntington (Fukae *et al.*, 2007; De Souza-Pinto *et al.*, 2008). Il y a néanmoins toujours discussion sur le fait que le dysfonctionnement des mitochondries soit ou non la cause principale de la neurodégénération, et cela même si leur bon fonctionnement est essentiel pour la survie des neurones, et le maintien de leur intégrité retarde l'apparition de la neurodégénération (De Souza-Pinto *et al.*, 2008).

Dans ce cadre, des niveaux plus élevés de dommages oxydatifs ont été détectés chez les patients atteints de maladies neurodégénératives que chez les individus sains, notamment dans l'ADN mitochondrial. C'est le cas chez les malades d'Alzheimer et de Parkinson (et cela même à un stade précoce pour Alzheimer), ce qui suggère que ces dommages pourraient avoir un rôle causatif dans l'apparition de ces maladies (Fukae *et al.*, 2007; Lovell *et al.*, 2007; De Souza-Pinto *et al.*, 2008). Cette accumulation n'est pas étonnante car l'ADN mitochondrial est plus vulnérable au stress oxydatif que l'ADN nucléaire, car il est directement exposé aux espèces réactives de l'oxygène venant de la chaîne de transport d'électrons.

La cause de cette augmentation de dommages sur l'ADN est cependant encore peu claire. Ainsi, une augmentation de production d'espèces réactives de l'oxygène (ROS) avec l'âge a été évoquée (De Souza-Pinto *et al.*, 2008). Un argument plus convainquant correspond

au déclin d'activité d'enzymes de réparation de l'ADN, qui a été corrélé à une neurodégénération (Wilson III *et al.*, 2007; De Souza-Pinto *et al.*, 2008). Notamment, une diminution d'activité et de production de protéines du système de réparation par excision de base a été observé pour les phases précoces et tardives d'Alzheimer (Lovell *et al.*, 2007; De Souza-Pinto *et al.*, 2008; Dizdaroglu *et al.*, 2008). En outre, le déclin d'activité d'enzymes de réparation est une caractéristique qui est liée à l'âge, et qui semble être plus importante dans les mitochondries que dans le noyau. Ainsi, une décroissance âge-dépendante de l'activité des enzymes du BER Ogg1, UNG et NTH1 dans différentes parties du cerveau de souris a été démontrée, cet effet étant nettement plus important dans les mitochondries que dans le noyau (Wilson III *et al.*, 2007). Tout ceci fait que l'on ne sait pas vraiment si le niveau de dommages oxydatifs augmente avec l'âge car il y a augmentation de production de ROS, car il y a baisse de la capacité de réparation, ou car ces deux phénomènes interviennent (De Souza-Pinto *et al.*, 2008).

Enfin, l'implication dans le développement de neuropathologies de polymorphismes sur des protéines de réparation a aussi été évoquée. Notamment, il a été rapporté qu'un polymorphisme sur la protéine Ogg1 était un évènement fréquent chez les malades d'Alzheimer, induisant un haut niveau de dommages oxydatifs, et constituant donc un facteur de risque dans le développement de la maladie (Mao *et al.*, 2007).

### - Athérosclérose

L'athérosclérose correspond à un dépôt de plaques d'athérome sur la paroi interne des artères, dépôts qui finissent par boucher celles-ci, amenant de graves problèmes cardiovasculaires. Il y a des preuves que les ROS jouent un rôle dans l'apparition de l'athérosclérose, et il a été suggéré que des altérations dans l'ADN cellulaire pourraient contribuer à la formation des plaques d'athérome (Olinski *et al.*, 2002). En effet, des dommages sur l'ADN sont présents en grande quantité dans les tissus athérosclérotiques, notamment la 8-oxo-G, qui a été détectée à des niveaux élevés dans les lésions de la paroi de l'aorte des patients atteints d'athérosclérose, ainsi que dans les lymphocytes de ces patients (les plaques d'athérosclérose contiennent entre autre des lymphocytes). Cela pourrait être l'un des évènements directement impliqués dans l'apparition de la maladie.

Ces données sont en accord avec le premier facteur de risque associé à l'athérosclérose, qui est l'âge. En effet, on sait que le niveau de dommages à l'ADN augmente avec l'âge, ce qui pourrait expliquer la corrélation entre le vieillissement et le développement de l'athérosclérose.

En outre, des LDLs (Low Density Lipoprotein) oxydées pourraient aussi jouer un rôle dans l'apparition de la maladie. De façon intéressante, il a été découvert que les LDLs oxydées (mais pas les formes non modifiées des LDL) amènent une diminution du niveau d'enzymes du système de réparation par excision de base (BER), ce qui pourrait expliquer les niveaux élevés de 8-OH-G observés. De plus, la vitamine C, qui est un antioxydant, prévient cette diminution de niveau d'enzymes du BER (Olinski *et al.*, 2002).

## **- VIH et SIDA**

L'hypothèse a été émise que l'augmentation du niveau de dommages sur l'ADN pourrait influencer la progression de l'infection par le virus d'immunodéficience humaine (VIH) en syndrome d'immunodéficience acquise (SIDA) (Olinski *et al.*, 2002). En effet, les patients atteints de HIV asymptomatique montrent un niveau élevé de dommages oxydatifs sur l'ADN de leurs lymphocytes, notamment de 8-OH-G et 5-OH-U. Ceci viendrait d'une création accrue de ROS et d'une diminution des antioxydants cellulaires. Il a ainsi été proposé que l'apoptose induite par les stress oxydatifs était l'une des causes principales de la perte de lymphocytes chez les patients infectés par le VIH, d'où une participation au développement du SIDA. De plus, comme évoqué précédemment, les dommages sur l'ADN pourraient aussi participer au développement de cancers associé au SIDA.

### **I.4. Les systèmes de réparation de l'ADN**

La grande diversité de lésions produites sur l'ADN a amené les organismes à développer différents systèmes de réparation qui soient capables de prendre en charge tous les types de dommages. Ce chapitre reviendra donc sur les différents systèmes de réparation connus, par contre seule la partie sur la réparation par excision de bases sera réellement développée dans le chapitre suivant car elle correspond au sujet de ce manuscrit. Les autres systèmes de réparation ne seront donc que « survolés », afin de montrer qu'ils existent, mais sans entrer dans des détails qui ne seraient pas pertinents pour la compréhension de mon travail de thèse.

#### **I.4.1. La réparation par réversion directe**

Cette voie de réparation correspond à une réparation en une seule étape enzymatique, qui permet de restituer l'information initiale sur l'ADN (Wood *et al.*, 2005).

Des exemples d'une telle réparation ont été trouvés avec la protéine AlkB bactérienne et ses homologues fonctionnels eucaryotes ABH2 et ABH3, qui enlèvent directement les méthyles des lésions 1-Me-Adénine et 3-Me-Cytosine.

D'autre part, la monomérisation par l'enzyme Photolyase des dimères de pyrimidines induits par les rayonnements UV est aussi un mécanisme de réversion directe, et est très répandu dans la nature même si il est absent de chez les mammifères.

#### **I.4.2. La réparation par excision et resynthèse**

Cette stratégie de réparation de l'ADN comprend trois systèmes de réparation distincts :

- Le système de réparation des mésappariement (MMR).
- Le système de réparation par excision de nucléotides (NER).
- Le système de réparation par excision de base (BER, ne sera pas décrit ici, car il est le sujet du chapitre suivant).

Tout d'abord, le MMR permet de réparer les mésappariements de bases et les insertions/délétions pouvant apparaître lors de la réplication ou de la recombinaison de l'ADN (Li, 2008). Après détection par des enzymes propres au MMR, un morceau d'ADN simple brin contenant le mésappariement est éliminé *via* l'action d'une exonucléase, puis il y a resynthèse par complémentarité avec l'autre brin. Il y a deux types de MMR, les « long-patch » (LP-MMR) et « very-short-patch » (VSP-MMR) suivant la taille de la partie d'ADN enlevée (Nimesh *et al.*, 2006). Le VSP-MMR enlève environ 10 nucléotides, et n'intervient que sur les mésappariements spécifiques issus de bases endommagées. Le LP-MMR peut réparer tous les types de mésappariements et insertions/délétions, et peut exciser des régions d'ADN allant jusqu'à plus d'un kilobase. Le MMR comprend trois étapes :

- La reconnaissance d'un mésappariement.
- La détermination de la base incorrecte dans ce mésappariement.
- La correction du mésappariement par excision et resynthèse.

Une caractéristique importante du MMR est que cette voie de réparation est brin-spécifique, n'agissant que sur le brin nouvellement synthétisé lors de la réplication afin d'éviter

de fixer des mutations. La voie de reconnaissance des brins est cependant encore obscure pour la plupart des organismes, hormis pour les bactéries GRAM négatives où la différence de méthylation entre le brin parent et le brin néosynthétisé est impliquée (Nimesh *et al.*, 2006; Li, 2008). De plus, on peut dire que la plupart des ADN polymérase répliquatives possèdent une activité exonucléase 3'→5' permettant de corriger les erreurs d'insertion qu'elles pourraient commettre (chapitre I.5.4 sur les ADN polymérase et leur implication dans la mutagenèse induite par les lésions). Cette activité dite d'édition correspond donc à un système précoce de réparation des mésappariements, qui ne nécessite pas l'intervention des protéines du système MMR à proprement parler.

Ensuite, le NER permet de réparer une très large gamme de lésions modifiant la forme de la double hélice d'ADN, telles que les dimères de pyrimidines, les pontages covalents intrabrans de l'ADN et les pontages ADN-protéines, ainsi qu'une grande variété de d'adduits chimiques encombrants de l'ADN (Nousspikel, 2008). La reconnaissance de ces lésions ne se fait pas directement, mais plutôt *via* les distorsions qu'elles induisent sur la double hélice d'ADN, d'où la large gamme de lésions réparables par le NER. La lésion est éliminée au sein d'un court fragment d'ADN simple brin (une douzaine de bases chez *E. coli*, 25-30 bases chez les eucaryotes), puis il y a resynthèse de l'information correcte par complémentarité avec l'autre brin. Le NER peut être divisé en deux voies, la voie NER sur le génome entier (GG-NER) et la voie NER couplée à la transcription (TC-NER), qui ne diffèrent que par les enzymes impliquées dans la reconnaissance des distorsions de l'ADN, et donc dans le type de distorsions reconnues. Le GG-NER répare les deux brins de l'ADN (transcrit et non transcrit) dans tout le génome, alors que le TC-NER n'est activé que sur le brin transcrit de l'ADN après une pause de l'ARN polymérase II (Frit *et al.*, 2002).

### **I.4.3. La réparation des coupures double-brins**

#### **- La recombinaison homologue**

Ce sous-chapitre est directement adapté de l'article de Lopez en 2006. La recombinaison homologue (RH) tire parti de la possibilité d'échange de brins entre deux séquences homologues d'ADN double brin, et cela du moment que ces séquences sont suffisamment similaires pour reformer des doubles-hélices stables. La RH peut donc intervenir dans le maintien de la stabilité des génomes, mais aussi dans la diversité et l'instabilité

génétique. La RH peut aboutir à des modifications génétiques si les zones d'ADN échangées ne sont pas totalement identiques. De plus, la RH peut apparaître sous deux formes avec et sans crossing-over, selon la voie de résolution de la jonction de Holliday.

De façon générale, la RH est impliquée dans la réparation des lésions bloquant la réplication. Les rôles les plus documentés de la RH dans la réparation de l'ADN, et donc dans le maintien de la stabilité des génomes, correspondent à son intervention dans la réparation des cassures double-brins (DSB) (et des pontages inter-brins de l'ADN). En outre, ce type de lésions bloquant la réplication, la RH peut permettre de continuer la réplication en changeant de matrice *via* l'utilisation de la chromatide sœur. Cependant d'autres lésions formées sur un seul brin peuvent arrêter la réplication (ex : produits de UV-C). Dans ce cas, la RH permet leur réparation avec moins de risques de mutagenèse que l'autre possibilité qui permet de résoudre ce type de lésion, en l'occurrence le recrutement d'ADN polymérase translésionnelles (ADN polymérase peu fidèles, chapitre I.5.4 sur les ADN polymérases et leur implication dans la mutagenèse induite par les lésions).

#### **- La ligation des extrémités non-homologues**

La voie de réparation par ligation des extrémités non-homologues (NHEJ) est une voie de réparation des DSB. Cette voie de réparation est présente chez tous les eucaryotes, mais absente chez certains procaryotes tels que *E. coli*. Le terme non-homologue du NHEJ vient du fait que les extrémités sont directement liguées sans faire appel à des régions homologues comme dans la recombinaison homologue. En fait, la NHEJ fait tout de même appel à de courtes régions homologues appelées microhomologies pour guider la réparation, qui sont souvent présentes sur de courtes extrémités simple-brins présentes au niveau des DSBs. Quand ces microhomologies sont parfaitement compatibles, la NHEJ répare généralement de façon exacte (Boulton *et al.*, 1996; Moore *et al.*, 1996; Wilson *et al.*, 1999; Budman *et al.*, 2005). Cependant si ces extrémités ne sont pas compatibles, alors on peut voir apparaître une mutagenèse lors de la NHEJ, souvent avec perte de nucléotides. De plus, un NHEJ sur des extrémités qui ne devraient pas être liguées peut amener des translocations de régions du génome, ainsi que des fusions de télomères (Espejel *et al.*, 2002).

## **I.5. Le système BER**

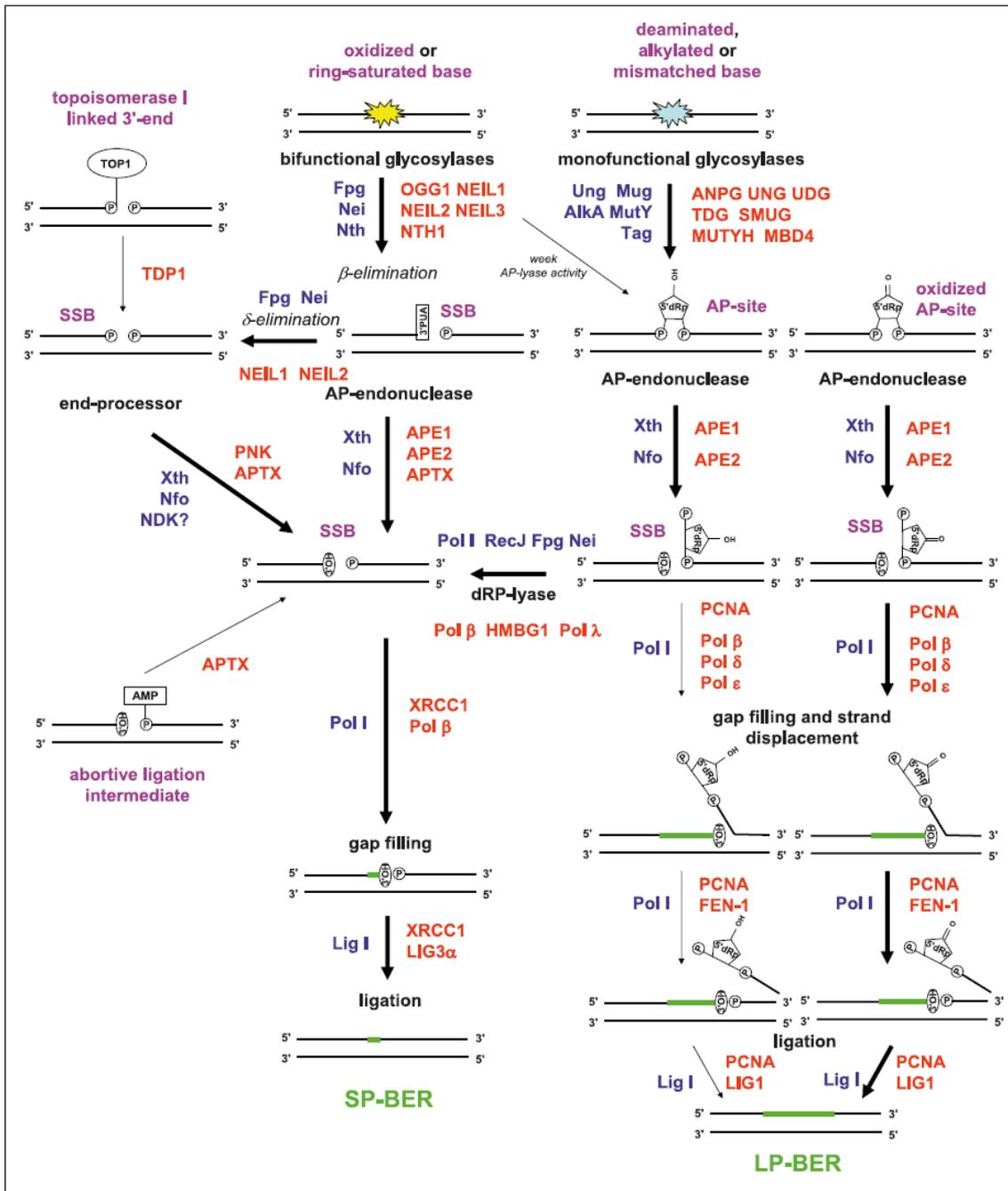
### **I.5.1. Généralités**

Comme évoqué précédemment, le choix de la voie de réparation est déterminé par le type de dommage sur l'ADN. Les modifications sur les bases de l'ADN telles que les oxydations, les alkylations et les déaminations sont principalement réparées par la réparation par excision de base (BER) (Schärer *et al.*, 2001), et correspondent aux dommages les plus fréquents sur l'ADN juste après les cassures simple-brins (SSB) et les dépurinations/dépyrimidations. Le système BER a été reconstitué grâce à des études sur des protéines purifiées de bactéries et de mammifères, et est conservé à travers l'évolution. Ce sous-chapitre sur les généralités du système BER est volontairement réduit, car les activités des différents membres du système seront détaillées dans les sous-chapitres suivants.

Le BER est initié par des ADN glycosylases, qui reconnaissent les bases endommagées, et dans certains cas les bases mésappariées, et les excisent de l'ADN en coupant la liaison N-glycosidique entre la base et le squelette phosphate-sucre de l'ADN (chapitre I.5.2 sur les ADN glycosylases).

Deux voies sont alors possibles dans le BER, le « short-patch » (SP-BER) ou le « long-patch » (LP-BER) (Figure 13).

Dans le SP-BER, les sites AP issus de l'activité des ADN glycosylases monofonctionnelles subissent un clivage du lien phosphodiester en 5' par une AP endonucléase (chapitre I.5.3 sur les protéines métabolisant les sites AP et extrémités modifiées de l'ADN), libérant des extrémités 3'-OH et 5'-désoxyribose-phosphate (5'-dRP). De plus, les AP endonucléases permettent d'enlever les extrémités 3'-PUA et 3'-P (voir légende de la Figure 13), libérant ainsi des extrémités 3'-OH sur les produits des ADN glycosylases mono- et bi-fonctionnelles. A cette étape, le choix de la voie SP-BER ou LP-BER dépend de la capacité de la Polymérase  $\beta$  humaine ou la Pol I procaryote à cliver le 5'-dRP (activité 5'-dRPase). Pour le SP-BER, la Polymérase  $\beta$  humaine ou la Pol I procaryote enlèvent les groupements 5'-dRP issus de la coupure des sites AP par les AP endonucléases (avec intervention possible d'autres enzymes ayant cette activité). Ensuite, les Pol  $\beta$  ou Pol I remplissent la lacune d'un nucléotide, et les ligases LIG3 $\alpha$  eucaryote ou ligase I bactérienne liguent le brin d'ADN.



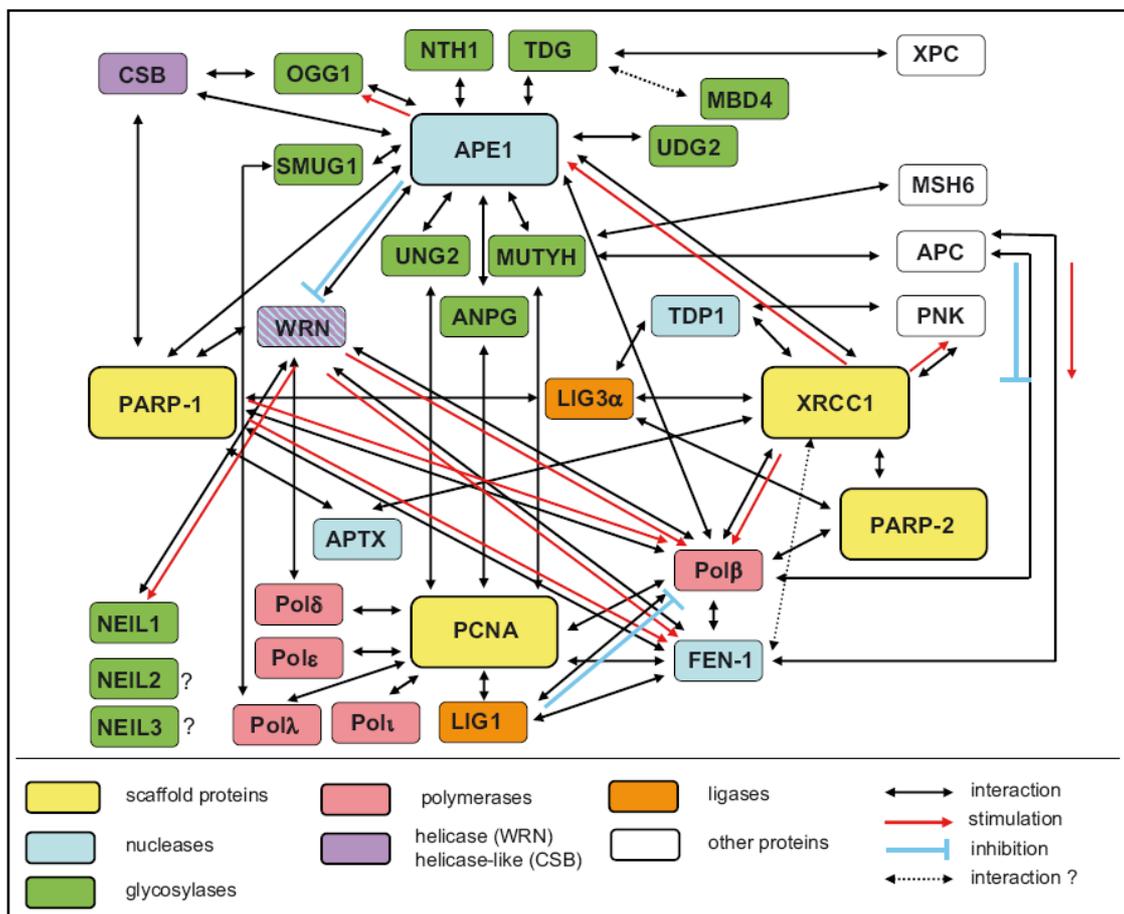
**Figure 13 :** Représentation schématique du système BER procaryote et eucaryote.

(issu de Krwawicz *et al.*, 2007). Les types de lésions réparés par une voie commune pour la réparation de coupures simple-brin et le BER sont indiqués en rose. Les enzymes de *E. coli* sont indiquées en bleu, les enzymes humaines sont indiquées en rouge. Les types d'extrémités de l'ADN sont : Phosphate (P); Groupement hydroxyle, (OH); Aldéhyde insaturé en 3' (3'-PUA), Désoxyribose phosphate en 5' (5'-dRP); Groupement adénylate (AMP); Topoisomérase I liée en 3' (TOP1).

Des modifications de l'extrémité 5'-dRP peuvent rendre cette extrémité impossible à cliver par la Pol β ou la Pol I, ce qui induit le LP-BER. Chez l'homme, le LP-BER implique que les Polymérase δ et ε soient recrutées, et celles-ci ajoutent quelques nucléotides (2 à 6) à l'extrémité 3'-OH de la cassure en déplaçant le brin qui contient l'extrémité 5'-dRP non clivable, générant ainsi une courte extrémité flottante (Flap). Cette extrémité flottante est alors excisée par la FIAP endonucléase 1 (FEN-1), et la coupure est liguée par la ligase I. Chez la

bactérie, la Pol I est responsable de la synthèse avec déplacement de brin, et son activité 5'→3' exonucléase permet d'éliminer l'extrémité flottante (2 à 8 nucléotides). Enfin, la ligase I permet la ligation de l'ADN.

On pense dorénavant que le SP-BER prédomine et qu'il est exclusivement utilisé dans certains cas, par exemple quand l'excision de base est médiée par des ADN glycosylases/AP lyases. Le LP-BER doit en fait servir de système de remplacement, et pourrait être important pour enlever les sites abasiques modifiés qui sont résistants à l'activité AP lyases de la Polymérase β (Schärer et al., 2001) (chapitre I.5.4 sur les ADN polymérase et leur implication dans la mutagenèse induite par les lésions).



**Figure 14 :** Représentation schématique de l'interactome du BER et de la réparation des coupures simple-brin. (issu de Krwawicz *et al.*, 2007).

On peut noter que de nombreuses protéines du BER humain font des interactions entre elles, ce qui indique que les voies du BER sont des processus hautement concertés, avec parfois plusieurs protéines pouvant remplir la même fonction. Un schéma représentant les différentes protéines impliquées dans ces interactions chez l'homme est donné dans la Figure 14, et indique aussi des interactions avec des protéines d'autres systèmes. On peut noter que cet

interactome comprend différents types de protéines (voir la légende de la Figure 14), dont certaines ont uniquement pour but d'être des « plateformes ou échafaudages » (scaffold) pour le recrutement ou le déplacement des protéines du système dans des complexes multiprotéiques (protéines « Proliferating Nuclear Cell Antigen » (PCNA), « X-ray cross-complementing 1 » (XRCC1), et les « Poly(ADP-ribose) polymerases » (PARP1 et 2)).

De plus, des interactions ne sont pas indiquées dans la Figure 14. En l'occurrence, la protéine de réplication A (RPA), qui interagit avec les glycosylases MUTYH et UNG2, est requise par les Pol  $\delta$  et Pol  $\epsilon$  pour la synthèse d'ADN, et stimule donc probablement le LP-BER. La Pol  $\delta$  nécessite aussi le facteur de réplication C (RF-C, qui recrute le PCNA sur la double hélice) et PCNA pour une synthèse efficace (Krwawicz *et al.*, 2007). Ces deux dernières informations lient l'interactome du BER à l'interactome cellulaire, et ont souvent été évoquées comme étant de bons indices d'un BER couplé à la réplication pour certaines ADN glycosylases.

Un autre aspect particulier du système BER qui mérite d'être discuté ici correspond au BER mitochondrial (Wilson III *et al.*, 2007; De Souza-pinto *et al.*, 2008). En effet, la plupart des études ont jusqu'ici traité du BER nucléaire, ce qui vient surtout d'une conception ancienne erronée qui disait que les mitochondries n'avaient pas de système particulier de réparation de l'ADN. Cependant, dans les dernières années, il y a eu de nombreuses avancées dans le domaine de la réparation de l'ADN mitochondrial (Tableau 2), notamment au niveau des similarités et différences avec la réparation de l'ADN nucléaire, ainsi qu'au niveau de la compréhension de leurs interactions.

Tout d'abord, on peut rappeler que l'ADN mitochondrial ne code pas pour des protéines de réparation de l'ADN. Les protéines de réparation que l'on retrouve dans les mitochondries correspondent en fait à des protéines codées par le gène nucléaire et ayant transloqué dans cet organe. Ces protéines sont souvent issues d'épissages alternatifs par rapport aux formes nucléaires, mais peuvent aussi être des produits de traduction alternatifs, ou bien des versions des protéines nucléaires ayant subi des modifications post-traductionnelles (De Souza-pinto *et al.*, 2008).

<b>Tableau 2 : Comparaison des voies de réparation nucléaires et mitochondriales chez les mammifères.</b> (Issu de Wilson III et al., 2007).			
Voie générale de réparation	Voie de réparation	Nucléaire	Mitochondriale
Excision de base	SP-BER	Oui	Oui
	LP-BER	Oui	Non (?)
Excision de nucléotide	/	Oui	Non
Mésappariement	/	Oui	Oui (?)
Recombinaison	Homologue	Oui	Oui (?)
	NHEJ	Oui	Oui

Les voies majeures de réparation sont représentées. Oui : voie établie de façon convaincante; Non : voie non observée; (?) : études non conclusives.

Pour le BER, de nombreuses ADN glycosylases possèdent une forme particulière dans les mitochondries (ex :  $\alpha$ -hOgg1 nucléaire et  $\beta$ -hOgg1 mitochondriale), et la plupart de celles-ci viennent d'un épissage alternatif du même gène produisant la forme nucléaire et la forme mitochondrial. Par contre, l'AP endonucléase présente dans les mitochondries ne semble pas être issue d'un épissage alternatif par rapport à l'APE1 nucléaire, mais correspond plutôt soit à l'APE1 nucléaire, soit au produit du clivage par une protéase du signal de localisation nucléaire en N-terminal de l'APE1 (De Souza-pinto *et al.*, 2008). D'autre part, la Polymérase  $\gamma$ , seule ADN polymérase à avoir été trouvée dans les mitochondries, est probablement responsable de toute synthèse d'ADN dans ces organites, que ce soit pour la réplication ou la réparation. Une forme mitochondriale de la LIG3 existe aussi et est issue d'un site d'initiation de la traduction alternatif sur le même ARNm qui code pour la forme nucléaire. Cette enzyme est responsable de la ligation des produits du BER, mais ne nécessite pas l'intervention de XRCC1 comme la LIG3 nucléaire. En outre, le BER des mitochondries semble être uniquement un SP-BER, avec aucune preuve de la présence d'un LP-BER.

Cette partie sur la réparation de l'ADN mitochondrial est volontairement réduite, mais des revues sur ce système peuvent être trouvées dans trois articles parus dans le même volume d'une même revue (Bogenghagen *et al.*, 2001; Dianov *et al.*, 2001; LeDoux *et al.*, 2001) et plus récemment dans De Souza-Pinto et collaborateurs en 2008.

Finalement, on ne peut que regretter que la réparation de l'ADN mitochondrial ne soit pas plus étudiée, car il y a désormais de nombreuses preuves de l'importance cruciale de l'intégrité de l'ADN mitochondrial dans la manifestation de maladies. En particulier, la réparation de l'ADN mitochondrial joue un rôle vital dans la résistance cellulaire au stress oxydatif, et il y a de plus en plus de preuves de l'intervention de mutations et dysfonctionnements des mitochondries dans l'apparition de symptômes de vieillissement et de maladies humaines spécifiques (chapitre I.3 sur les conséquences des dommages sur l'ADN en l'absence de réparation) (De Souza-Pinto *et al.*, 2008).

## **I.5.2. Les ADN glycosylases**

Les ADN glycosylases sont des enzymes présentes chez tous les organismes vivants, et dont le rôle est d'initier la voie de réparation par excision de base en clivant la liaison N-glycosidique N-C1' entre le désoxyribose et une base nucléique lésée ou mésappariée. La première de ces enzymes à avoir été découverte est l'Uracile-ADN glycosylase (UDG) par Thomas Lindahl en 1974, qui étudiait la voie de réparation des cytosines désaminées dans l'ADN (Lindahl, 1974). Jusqu'au début des années 1990, la recherche sur le BER s'est focalisée sur la caractérisation génétique et biochimique d'ADN glycosylases bactériennes, en mettant l'accent sur la détermination de la spécificité de substrat de ces enzymes. Depuis, notre compréhension des ADN glycosylases a avancé grandement à de nombreux niveaux :

- De nombreux gènes codant pour des ADN glycosylases eucaryotes ont été clonés, et les protéines correspondantes caractérisées.
- Les voies du BER correspondantes ont été reconstituées.
- De nombreux modèles de souris déficientes en une ou plusieurs ADN glycosylases ont été créés, et ont permis de commencer à comprendre les conséquences de la déficience du BER chez les mammifères.
- Des études structurales à haute résolution et des études chimiques et biochimiques sur les ADN glycosylases ont amené une meilleure compréhension des mécanismes sous-jacents à la reconnaissance des bases endommagées et à la catalyse par ces enzymes.

Ce chapitre reviendra donc sur les connaissances actuelles structurales et fonctionnelles sur ces enzymes et leurs substrats, ainsi que sur les conséquences de leur déficience.

### **I.5.2.1. Les caractéristiques conservées chez toutes les ADN glycosylases**

Les ADN glycosylases, bien qu'utilisant différents mécanismes catalytiques, possèdent des caractéristiques communes (Stivers *et al.*, 2003; Huffman *et al.*, 2005) :

- Ce sont des protéines de petite taille (20 à 45 kDa), et généralement basiques.
- Des interactions de type liaison-hydrogènes entre des groupements de la protéine et les oxygènes du squelette phosphaté de l'ADN (généralement en majorité sur le brin contenant la lésion).
- Une fixation de l'ADN glycosylase dans le petit sillon de l'ADN.
- Une courbure de l'ADN au niveau du dommage.

- Une extrusion de la lésion hors de la double hélice d'ADN *via* le grand sillon de l'ADN et sa stabilisation dans la poche du site actif de l'enzyme, permettant de rendre accessible le C1' de la lésion pour l'attaque nucléophile catalysée par l'enzyme.
- Une intercalation de résidus de l'enzyme dans l'espace créé par l'extrusion de la lésion hors de la double hélice d'ADN, forçant cette extrusion, et empêchant le retour de la lésion dans la double hélice et l'effondrement de la structure locale de l'ADN.

En fait, plusieurs de ces caractéristiques sont liées, car les ADN glycosylases utilisent la courbure de l'ADN pour faciliter l'extrusion de la lésion. En effet, il a été mis en évidence que l'ouverture d'une paire de base (et donc l'extrusion d'une base hors de la double hélice d'ADN) était plus facile si l'ADN était courbé (Ramstein *et al.*, 1988). De plus, les lésions présentes dans l'ADN lui confèrent généralement une plus grande flexibilité et diminuent la barrière d'énergétique nécessaire à l'extrusion de la base, d'où une courbure et une extrusion facilitées (Fuxreiter *et al.*, 2002; Seibert *et al.*, 2002). Ceci permet d'ailleurs aussi d'expliquer en partie la variation d'efficacité des ADN glycosylases en fonction de la séquence qui entoure le dommage, car cette séquence joue aussi sur la flexibilité de l'ADN. Un bon exemple de cette propriété peut être trouvé dans la publication de Seibert et collaborateurs en 2002, qui montre que les séquences A/T riches autour du dommage augmentent l'efficacité de l'UDG. En effet, ces séquences augmentent la flexibilité locale de l'ADN, rendant l'extrusion de la lésion plus facile, d'où une probabilité accrue d'obtenir un complexe catalytiquement actif.

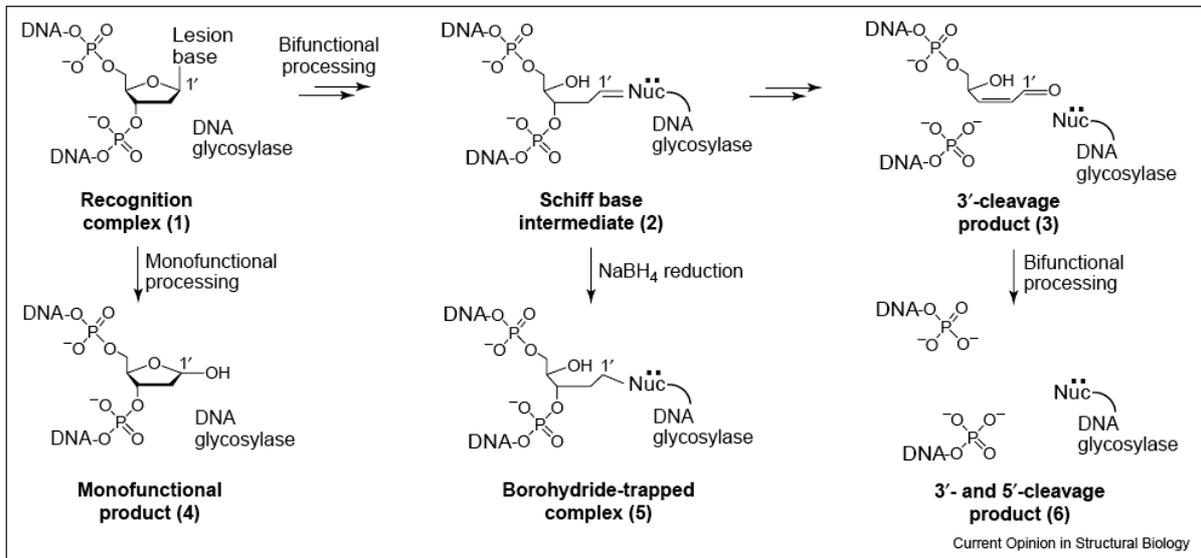
Ainsi, une étape initiale de la reconnaissance du dommage exploite la déformabilité de l'ADN au niveau de la paire de base déstabilisée par la lésion. Cependant, chaque glycosylase étant spécifique de lésions particulières, seules les bases qui peuvent être accommodées correctement dans la poche de fixation après extrusion donnent les contacts et l'orientation nécessaires pour l'excision de base.

#### **I.5.2.2. Les deux grandes familles mono- et bi-fonctionnelles**

Les ADN glycosylases peuvent être classifiées en deux grandes catégories fonctionnelles en se basant sur leur mécanisme catalytique (Shärer *et al.*, 2001; Fromme *et al.*, 2004) (Figures 15 et 17) :

- Les ADN glycosylases monofonctionnelles, qui clivent la liaison *N*-glycosidique par hydrolyse en utilisant une molécule d'eau comme nucléophile pour l'attaque de la liaison glycosidique, et qui génèrent un site AP.

- les ADN glycosylases/AP lyases bi-fonctionnelles, qui utilisent un groupement amino de la protéine comme nucléophile pour l'attaque de la liaison glycosidique. Celles-ci forment un intermédiaire de type base de Schiff, qui subit une  $\beta$ - ou  $\beta,\delta$ -élimination catalysée par l'enzyme, permettant ainsi le clivage la liaison phosphodiester en 3' ou 3' et 5' du sucre sur lequel était fixé la base lésée.



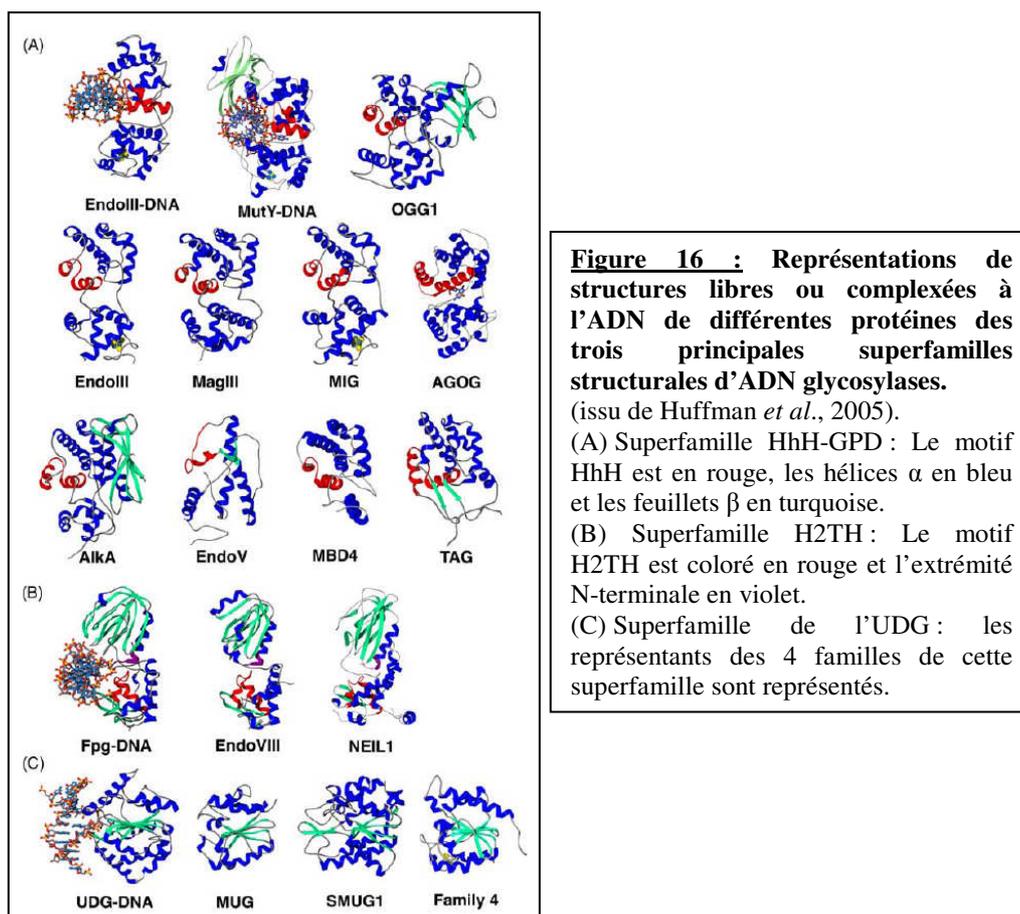
**Figure 15 : Mécanisme des ADN glycosylases mono- et bi-fonctionnelles.** (issu de Fromme *et al.*, 2004). (1) Complexe de reconnaissance de la lésion. (4) Site abasique généré par une ADN glycosylase monofonctionnelle. (2) Intermédiaire de type base de Schiff créé par une ADN glycosylase bi-fonctionnelle, donnant ensuite un clivage en 3' (3) ou 3' et 5' (6) du brin d'ADN. L'intermédiaire (2) peut être stabilisé par action d'un agent réducteur tel que le borohydrure de sodium ( $\text{NaBH}_4$ ).

On voit donc sur la figure précédente qu'une caractéristique des ADN glycosylases bi-fonctionnelles est qu'un intermédiaire réactionnel de type « imino » peut être stabilisé covalamment par réduction au borohydrure de sodium, ce qui n'est pas possible pour les enzymes monofonctionnelles qui utilisent une molécule d'eau comme nucléophile. En fait, il existe des cas particuliers comme pour l'enzyme MutY, qui est décrite comme étant monofonctionnelle, mais qui utilise l'un de ses résidus amino-acides comme nucléophile, pouvant donc être piégé par réduction (Williams *et al.*, 1998).

### I.5.2.3. Les superfamilles structurales d'ADN glycosylases

Les ADN glycosylases sont des protéines ne possédant que peu d'homologie de séquence, ce qui rend leur classification difficile. Cependant, depuis quelques années la biologie structurale a subi un essor sans précédent, ce qui a permis la résolution de nombreuses structures d'ADN glycosylases procaryotes et eucaryotes. A partir de ces données, il a alors été possible de définir des motifs structuraux communs pour certaines de ces protéines, ce qui a permis de proposer une classification par superfamilles et familles structurales d'ADN glycosylases.

A ce jour, six superfamilles structurales d'ADN glycosylases ont été définies. Ce sous-chapitre reviendra sur ces différentes superfamilles, avec une description succincte de chacune, mais sans rentrer dans le détail des familles structurales qui les constituent (Figure 16 et Tableau 3) (Lau *et al.*, 1998; Aravind *et al.*, 2000; Schärer *et al.*, 2001; Fromme *et al.*, 2004; Huffman *et al.*, 2005; Alseth *et al.*, 2006; Dalhus *et al.*, 2007; Zharkov *et al.*, 2008).



La première superfamille structurale d'ADN glycosylases est celle de l'UDG, qui comprend des enzymes qui ne partagent qu'une faible homologie de séquence entre-elles

(environ 10%), mais qui ont toutes un repliement  $\alpha/\beta$  constitué d'un feuillet- $\beta$  vrillé à quatre brins parallèles entouré d'hélices- $\alpha$ .

La seconde superfamille est celle de Nth (Endonucléase III), qui est définie grâce à la présence invariante d'un motif Hélice-épingle à cheveux-Hélice (HhH, motif de fixation non séquence-spécifique à l'ADN) suivi d'une boucle riche en prolines et glycines, ainsi que d'un aspartate important pour l'activité catalytique situé à une vingtaine de résidus en C-terminal du motif HhH, d'où le nom de superfamille HhH-GPD.

La troisième superfamille structurale d'ADN glycosylases est celle de Fpg/Nei, et est caractérisée par la présence d'une extrémité N-terminale conservée et impliquée dans l'activité catalytique (notamment la proline terminale, qui est le nucléophile catalytique), ainsi que la présence d'un motif de type doigt à zinc et d'un motif Hélice-2 Tours-Hélice (H2TH, motif de fixation non séquence-spécifique à l'ADN).

D'autres ADN glycosylases possèdent un repliement particulier et ne font donc pas partie des trois superfamilles précédentes, cependant elles ne sont pas représentées dans la Figure 16. C'est notamment le cas de la protéine AAG, qui définit à elle seule une quatrième superfamille structurale caractérisée par un feuillet- $\beta$  antiparallèle entouré d'hélices- $\alpha$ . De même, les protéines AlkC et AlkD définissent une cinquième superfamille structurale caractérisée par un repliement de type superhélice  $\alpha$ - $\alpha$ . Enfin, l'endonucléase V de bactériophage T4 définit une sixième et dernière superfamille, caractérisée par un seul domaine compact structuré en hélices- $\alpha$ .

Les enzymes des superfamilles de l'UDG, de l'AAG, de la T4 endonucléase V, et des AlkC et D sont des enzymes compactes avec un seul domaine, avec des surfaces d'interaction avec l'ADN relativement petites. Par contre, les deux autres superfamilles possèdent plusieurs domaines, avec le site actif localisé à la jonction interdomaine.

Enfin, il est important de noter que le repliement global des enzymes de chaque superfamille ne donne pas d'indication mécanistique sur les spécificités des enzymes, d'où des études structurales et fonctionnelles sur chaque membre de chaque famille. Par exemple, les protéines Fpg et Ogg1 ne font pas partie de la même superfamille structurale alors que leur spécificité de substrats est quasiment identique. De même, MutY et Ogg1 sont de la même superfamille structurale mais sont respectivement des enzymes mono- et bi-fonctionnelle ayant des substrats différents.

### I.5.2.4. Spécificité de substrat et redondance des ADN glycosylases

Le but de ce sous-chapitre est de montrer sans équivoque qu'une même lésion peut être excisée par différentes ADN glycosylase (systèmes redondants pour l'élimination des lésions), et qu'une ADN glycosylase peut généralement exciser une grande variété de lésions, ce qui est résumé dans le Tableau 3.

Super-famille structurale	Nom	Famille fonctionnelle (type d'élimination si bi-fonctionnelle)	Substrats connus
<b>ADN glycosylases humaines</b>			
UDG	UNG	Monofonctionnelle	U dans ADN sb et db; 5-OH-U.
	TDG	Monofonctionnelle	U et T appariés à G; εC
	SMUG1	Monofonctionnelle	U dans ADN sb et db; 5-OH-Me-U; 5-Formyl-U; 5-OH-U; 5-Fluoro-U; εC; T
HhH-GPD	Ogg1	Bi-fonctionnelle (β)	Me-FapyG; FapyG; 8-oxo-G; 8-oxo-A
	MYH	Monofonctionnelle	A appariés à 8-oxo-G, G et C; 2-oxo-A appariée à G
	MBD4	Monofonctionnelle	U et T; 5-OH-Me-U; 5-Formyl-U; 5-Me-C; 5-Fluoro-U
	NTH1	Bi-fonctionnelle (β)	5-OH-C; Tg; 5-OH-U; DHU; FapyG; 5-Formyl-U; Cg; DHT; 5-OH-DHT
Fpg/Nei	NEIL1	Bi-fonctionnelle (β,δ)	8-oxo-G; Tg; 5-OH-C; 5-OH-U; DHU; FapyA; FapyG; Me-FapyG; DHT
	NEIL2	Bi-fonctionnelle (β,δ)	5-OH-U; DHU; 5-OH-C; DHT; Tg; 8-oxo-G
	NEIL3	Bi-fonctionnelle (β)	FapyG
AAG	AAG	Monofonctionnelle	3-Me-A; 7-Me-A; 3-Me-G; 7-Me-G; εG; εA; EA; HX; 8-oxo-G
<b>ADN glycosylases bactériennes</b>			
UDG	Ung	Monofonctionnelle	5-OH-U; 5,6-OH-U; U dans ADN sb et db
	Mug	Monofonctionnelle	EC; HEC; PC; HPC; εG; 5-OH-C; 5-OH-U; 5-OH-Me-U; 5-Formyl-U; U et T appariés à G; εC
HhH-GPD	Nth	Bi-fonctionnelle (β)	Ug; 5-OH-C; 5-OH-U; DHU; Tg; DHT; hmh; 5-OH-DHT; 6-OH-DHT; 5-OH-Me-U; 5-Formyl-U; 6-OH-DHC; 6-OH-DHU; 5-OH-DHU; diOHC; diOHU; FapyA; εA à cycle ouvert
	MutY	Monofonctionnelle	A appariés à 8-oxo-G, G et C; 2-oxo-A appariée à G
	AlkA	Monofonctionnelle	εG; εA; EA; 3-Me-A; 3-Me-G; O2-Me-C; O2-Me-T; 7-Me-A; 7-Me-G; 8-Me-G; 5-OH-Me-U; 5-Formyl-U; HX; X
	Tag	Monofonctionnelle	3-Me-A; 3-Me-G; 7-Me-G
Fpg/Nei	Fpg	Bi-fonctionnelle (β,δ)	8-oxo-G; FapyG; Me-FapyG; oxazolone; FapyA; 5-OH-C; 5-OH-U; Tg; DHT; hmh; 5-OH-Me-U; 5-Formyl-U; Ug; εA2; εA3
	Nei	Bi-fonctionnelle (β,δ)	8-oxo-G; FapyA; Tg; 5-OH-DHT; 5-OH-DHU; hmh; 5-OH-Me-U; 5-Formyl-U; Ug; 5-OH-U; 5-OH-C; DHT; DHU; 5-OH-DHU; diOH-U; diOH-C
AlkC/AlkD	AlkC	Monofonctionnelle	3-Me-A; 3-Me-G; 7-Me-G
	AlkD	Monofonctionnelle	3-Me-A; 3-Me-G; 7-Me-G

**Tableau 3 : Principales ADN glycosylases humaines et bactériennes.**

(Adapté de Krwawicz *et al.*, 2007; Zharkov *et al.*, 2008).

La notation peu explicite de certaines lésions correspond à : 3,N<sup>4</sup>-éthéno-cytosine (εC); 1,N<sup>6</sup>-éthéno-adénine (εA); éthéno-guanines (εG); éthano-adénine (EA); éthano-cytosine (EC); hypoxanthine (HX); thymine-glycol (Tg); cytosine-glycol (Cg); uracile-glycol (Ug); 5,6-dihydro-thymine (DHT); 5,6-dihydro-uracile (DHU); 5,6-dihydro-cytosine (DHC); 3,N<sup>4</sup>-α-hydroxy-éthano-cytosine (HEC); propano-cytosine (PC); hydroxy-propano-cytosine (HPC); 5-hydroxy-5-méthyle-hydantoïne (hmh); dérivés d'oxydation de l'éthénoadénine à cycle fermé (εA2) ou ouvert (εA3).

De plus, ce tableau permet d'illustrer que les familles fonctionnelles et superfamilles structurales n'ont pas forcément de lien direct, avec des spécificités de substrats différentes pour une même famille ou superfamille.

D'autre part, la liste des substrats pour chaque enzyme n'est pas exhaustive, car celle-ci dépend surtout des molécules qui ont été testées avec une ADN glycosylase donnée.

En outre, la spécificité de substrats apparente d'une ADN glycosylase peut grandement varier selon la méthode employée pour la déterminer, notamment si on étudie l'efficacité d'une enzyme donnée sur des oligonucléotides courts ou sur des ADN longs (chapitre I.7 sur la protéine bactérienne Fpg, qui reprend plus en détail la spécificité de cette enzyme).

Enfin, la liste d'ADN glycosylases présente dans le Tableau 3 n'est pas non plus exhaustive pour les procaryotes. Cette liste reprend donc uniquement les enzymes de *E. coli* et humaines, hormis pour AlkC et AlkD, qui viennent de *Bacillus cereus* (Alseth *et al.*, 2006), mais qui ont été rajoutées dans ce tableau car elles définissent une superfamille structurale indépendante. De plus, la T4 endonucléase V (ou T4-Pdg) n'est pas montrée dans ce tableau, mais on peut dire que c'est une ADN glycosylase spécifique de l'excision des dimères de pyrimidines (Golan *et al.*, 2006).

#### I.5.2.5. Les stratégies permettant des études structurales sur les ADN glycosylases

Tout d'abord, on peut rappeler que les ADN glycosylases, hormis quelques enzymes procaryotes, font toutes plus de 200 résidus (Krwawicz *et al.*, 2007), ce qui rend difficile voire impossible leur étude structurale par résonance magnétique nucléaire (RMN), notamment si on veut étudier la structure de ces enzymes en interaction avec un ADN.

Sachant cela, les études structurales sur ces enzymes ont été réalisées, généralement par cristallographie de rayons X, ce qui donne l'avantage de pouvoir résoudre des structures de molécules de très grande taille pourvu que l'on trouve de bonnes conditions de cristallisation. Par contre, l'inconvénient de la technique est qu'elle repose sur la stabilisation de molécules dans une conformation globale donnée, ne donnant pas ou peu d'informations sur la dynamique moléculaire couplée aux activités de ces enzymes. Afin de mieux comprendre ces phénomènes dynamiques, la solution qui a été mise en œuvre consiste en la stabilisation d'intermédiaires réactionnels qui puissent apporter des informations sur les étapes clés des activités des ADN glycosylases.

Cependant, si on reprend le modèle simplifié de Michaelis et Menten pour l'activité enzymatique :



Alors on voit très rapidement que la stabilisation d'intermédiaires réactionnels sera impossible dans les conditions de mise en présence d'une enzyme active avec son substrat, car dans ces conditions le substrat est rapidement converti en produit (d'autant plus si on considère que l'on travaille à très haute concentration pour les études structurales). Les seules données structurales que l'on pourra obtenir dans ce cadre correspondent aux structures de l'enzyme libre, de son substrat, ou de son produit de réaction.

Afin de résoudre ce problème, des stratégies basées sur les connaissances mécanistiques issues de la biochimie des ADN glycosylases ont été utilisées.

La première étape de l'activité des ADN glycosylases correspond à un glissement le long de l'ADN non endommagé et/ou des sauts d'une position à une autre (Francis *et al.*, 2003; Zharkov *et al.*, 2003), permettant à l'enzyme de trouver la lésion. Une stratégie pour mettre en évidence ce type de complexe primaire avec un ADN normal a été fondée sur une technique de pontage covalent entre la protéine et un ADN non lésé (Banerjee *et al.*, 2006b).

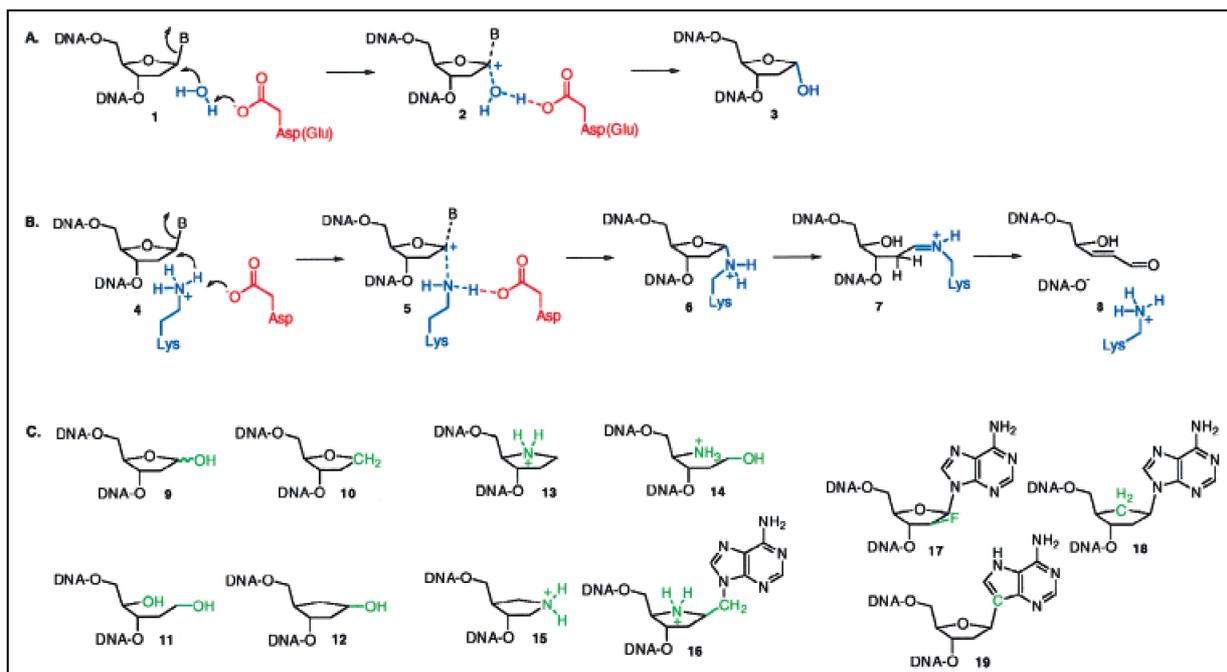
La seconde étape correspond, dès que l'enzyme rencontre une lésion, à la reconnaissance primaire de celle-ci, ce qui permet à l'enzyme de sélectionner les lésions qu'elle va extruder de la double hélice d'ADN puis placer dans la poche de son site actif. Deux stratégies différentes de pontage covalent ont été utilisées pour investiguer structurellement les types de complexes présents entre l'inspection de l'ADN endommagé et la stabilisation de la lésion dans le site actif de l'enzyme (Banerjee *et al.*, 2005; Banerjee *et al.*, 2006a; Banerjee *et al.*, 2006b; Lee *et al.*, 2008).

A ce niveau, les lésions se retrouvent stabilisées dans la poche du site actif des ADN glycosylases (Complexes 1 et 4 de la Figure 17). Pour étudier ces complexes, les études structurales ont pris avantage de deux stratégies permettant d'éviter la catalyse enzymatique. La première repose sur la mutagénèse dirigée de résidus du site actif des enzymes pour abroger l'activité glycosylase en gardant la fixation spécifique à l'ADN endommagé (Slupphaug *et al.*, 1996; Bruner *et al.*, 2000 ; Fromme *et al.*, 2003a). La seconde stratégie consiste en la définition et la synthèse d'analogues de lésions dont la liaison glycosidique est non métabolisable (Composés 17 à 19 de la Figure 17) (Schärer *et al.*, 1995; Schärer *et al.*, 1997; Chepanoske *et al.*, 1999; Haraguchi *et al.*, 2002; Wiederholt *et al.*, 2002a; Ober *et al.*, 2003; Coste *et al.*, 2004; Ober *et al.*, 2005).

Ensuite, les étapes catalytiques correspondant à l'activité glycosylase de Fpg ont aussi été étudiées. Pour ce faire, des analogues d'états de transitions ont été utilisés, dont certains pour des études structurales (Composés 13 à 16 de la Figure 17) (Deng *et al.*, 1997; Schärer *et al.*, 1998; Haushalter *et al.*, 1999; Hollis *et al.*, 2000).

Arrivé à ce stade, on a deux cas de figures : (i) Pour les ADN glycosylases monofonctionnelle, les études structurales du produit de l'activité glycosylase ne posent pas de problème, car le site abasique créé est généralement un bon inhibiteur des ADN glycosylases, formant donc des complexes stables (Parikh *et al.*, 2000). (ii) Pour les bi-fonctionnelles, il est nécessaire d'utiliser des analogues de sites abasiques ou sites abasiques stabilisés chimiquement pour pouvoir résoudre la structure de complexes avec les ADN glycosylases (Schärer *et al.*, 1998; Castaing *et al.*, 1992; Serre *et al.*, 2002; Pereira de Jésus *et al.*, 2005). De plus, pour les bi-fonctionnelles, il est possible de résoudre la structure du complexe covalent avec l'intermédiaire de type base de Schiff réduit au borohydrure de sodium (Complexe 5 de la Figure 15) (Fromme *et al.*, 2002; Gilboa *et al.*, 2002).

Par contre, à ma connaissance, aucune structure d'intermédiaires réactionnels de l'activité AP lyase n'a été résolue. Les dernières structures que l'on peut obtenir correspondent à celles des ADN glycosylases complexées à leur produit de réaction AP lyase, et ceci car ce produit de réaction est généralement un bon ligand pour les ADN glycosylases (Fromme *et al.*, 2002).



**Figure 17 : Mécanismes de réaction et inhibiteurs des ADN glycosylases.**

(A) Mécanisme réactionnel des ADN glycosylases monofonctionnelles : Une molécule d'eau est déprotonée dans le site actif par une chaîne latérale carboxylique de l'enzyme et positionnée pour attaque nucléophile sur le carbone anomérique. Dans l'état de transition (2), une charge positive apparaît sur le ribose. Le produit de réaction est un site abasique (3). (B) Mécanisme réactionnel des ADN glycosylases bi-fonctionnelles : Mécanisme similaire à celui des bi-fonctionnelles, hormis le fait que le nucléophile est un résidu aminé de l'enzyme (souvent une lysine). Le produit de clivage de la liaison glycosidique (6) subit un réarrangement, et une β-élimination permet la coupure entre le phosphate et le côté 3' du site abasique (8). (C) Inhibiteurs synthétiques des ADN glycosylases : (9) Site abasique, et ses formes réduites ou stabilisées chimiquement (10-12). (13-16) Analogues d'états de transition chargés positivement. (17-19) Analogues de nucléotides avec une liaison glycosidique non clivable par les ADN glycosylases.

Enfin, pour clore ce chapitre, on peut dire que toutes les techniques impliquant des mutants d'enzymes, des nucléotides modifiés, ou des stratégies de pontages covalents impliquent un biais qui ne peut être négligé, car ces modifications peuvent à elles-seules modifier légèrement les structures des complexes. En fait, pour valider ou invalider un modèle structural, une solution consiste à résoudre plusieurs structures pour un même complexe avec des stratégies différentes. Par exemple, on peut résoudre une structure avec une enzyme de type sauvage et un analogue de substrat non métabolisable d'une part, et avec une enzyme mutante et un substrat vrai d'autre part.

#### I.5.2.6. Les ADN glycosylases et la variabilité des immunoglobulines

Comme évoqué précédemment, le rôle principal des ADN glycosylases est le maintien de l'information génétique *via* leur participation à la réparation par excision de base. Cependant, la participation de UNG, une uracile-ADN glycosylase eucaryote, à

l'hypermutation somatique (SHM) assurant diversité et spécificité aux anticorps a récemment été évoquée (Neuberger *et al.*, 2003; Martomo *et al.*, 2006), indiquant donc qu'une enzyme sensée lutter contre la mutagénèse peut aussi participer à des effets promutagènes.

En effet, l'apparition d'uraciles dans l'ADN par désamination spontanée des cytosines en uraciles était connue (Duncan *et al.*, 1982), amenant potentiellement des transitions G/C vers A/T dans tout le génome lors de la réplication. Cependant, il a récemment été montré que des désaminations enzymatiques étaient induites spécifiquement dans les séquences codant pour les régions variables des immunoglobulines dans les lymphocytes B. Ces désaminations sont dues à la désaminase AID (Activation-Induced cytosine Deaminase), et entraînent l'apparition de paires G/U dans l'ADN (Neuberger *et al.*, 2003; Martomo *et al.*, 2006).

Ces paires G/U sont alors directement une source de transitions G/C vers A/T dans l'ADN si on a réplication des paires G/U. De plus, l'activité de UNG sur les paires G/U crée des sites AP en face de guanines, qui donnent des transitions et transversions si ces sites AP sont utilisés comme modèle lors de la réplication (Martomo *et al.*, 2006). Ces mutations correspondent surtout à des transitions G/C vers A/T, une adénine étant majoritairement insérée en face des sites AP par les ADN polymérases (Boiteux *et al.*, 1982; Simonelli *et al.*, 2005). Néanmoins, on peut avoir d'autres types de mutations avec la machinerie de réplication normale, et encore plus si des ADN polymérases de basse fidélité sont recrutées. Ce n'est normalement pas le cas car l'activité de UNG est théoriquement suivie des actions de APE1 puis de la Pol  $\beta$ , qui est une ADN polymérase fidèle (Bebenek *et al.*, 2004).

On peut aussi noter que UNG n'est pas la seule protéine de réparation qui soit impliquée dans la SHM, étant donné que les enzymes Msh2 et Msh6 (MutS-Homolog) du système de réparation des mésappariements sont aussi impliquées. Cependant, ces protéines sont sensées être couplées à des ADN polymérases réplcatives pour la resynthèse de l'ADN (Pol  $\delta$  et  $\epsilon$ ), qui sont connues pour être relativement fidèles (chapitre I.5.4 sur les ADN polymérases et leur implication dans la mutagénèse induite par les lésions) (Bebenek *et al.*, 2004).

La voie d'apparition des mutations *via* UNG dans le SHM est donc décrite comme venant majoritairement de la réplication de sites AP, ce qui ne devrait pas arriver si on avait un BER normal. En effet, dans ce cas UNG resterait fixée sur son produit de réaction jusqu'à intervention de l'APE1, justement pour éviter la réplication des sites AP (Parikh *et al.*, 1998). Le BER induit par UNG n'est donc pas conventionnel dans le cas de la SHM, ce qui peut être lié à la saturation du système BER par la création de trop d'uraciles dans l'ADN *via* l'action de l'AID. Notamment, on peut se demander si la partie du BER en aval de l'uracile-ADN

glycosylase ne serait pas plus saturée que la partie correspondant à l'action de UNG, ce qui augmenterait la quantité de sites AP non-réparés pouvant être substrats pour les ADN polymérase. Dans ce cadre, des études pourraient être menées afin de déterminer si il n'y aurait pas surproduction de UNG (ou intervention d'autres uracile ADN glycosylases potentielles) par rapport aux autres acteurs du BER, ou s'il n'y aurait pas sous-production d'AP endonucléase dans les lymphocytes B.

D'autre part, les publications évoquées précédemment indiquent que la mutagenèse du SHM viendrait aussi du recrutement d'ADN polymérase de basse fidélité (Pol  $\theta$ ,  $\eta$ ,  $\zeta$ ,  $\kappa$  et REV1). Dans ce cas, et même si on ne trouve cette information que sous forme d'hypothèse (Diaz *et al.*, 2005), l'intervention d'ADN polymérase non-fidèles pourrait participer directement aux mutations observées pour la SHM *via* la réplication de l'ADN, et cela sans avoir à passer par les sites AP, donc indépendamment de UNG. La façon dont ces polymérase seraient recrutées dans le cas de la SHM reste cependant à résoudre, étant donné que ce ne sont pas celles qui sont sensées participer aux voies de réparation induites par les protéines de réparation UNG, Msh2 et Msh6.

Une autre voie qui pourrait participer à la mutagenèse induite par UNG a été décrite par Neuberger et collaborateurs en 2005. Le dUTP, un intermédiaire lors de la synthèse *de novo* de dTTP, est incorporé lors de la réplication en donnant des paires U/A (Duncan *et al.*, 1982). L'hypothèse émise est qu'une uracile-ADN glycosylase inconnue (autre que UNG) enlèverait les U des paires U/A, induisant des mutations *via* la réplication par une ADN polymérase de faible fidélité des sites abasiques ainsi créés. Cependant, aucune donnée n'a été publiée pour étayer cette hypothèse. De plus, cette voie serait indépendante de l'AID, et serait donc possible dans tout le génome et dans tous les types de cellules en réplication.

D'autre part, il y a de multiples désaminases des acides nucléiques chez l'homme, dont beaucoup sont de fonction inconnue, si ce n'est que certaines sont capables de désaminer les cytosines dans l'ADN. Certaines pourraient donc être impliquées dans une mutagenèse UNG-dépendante dans d'autres régions du génome que celle codant pour les immunoglobulines, et dans d'autres types de tissus, pouvant ainsi participer à l'évolution des génomes, mais aussi à l'apparition de cancers.

### **I.5.3. Les protéines métabolisant les sites AP et extrémités modifiées de l'ADN**

De nombreux processus cellulaires amènent la création d'extrémités 5' et/ou 3' modifiées dans l'ADN, qui empêchent l'action des polymérases et/ou ligases, et qui nécessitent donc leur traitement pour pouvoir restaurer un ADN normal.

Parmi ces processus, les glycosylases mono- et bi-fonctionnelles génèrent respectivement des sites AP et de telles extrémités, qui sont des intermédiaires très mutagènes si ils ne sont pas réparés (Sobol *et al.*, 2003; Boiteux *et al.*, 2004; Simonelli *et al.*, 2005). De plus, des ADNs contenant des extrémités 5' et/ou 3' modifiées peuvent aussi provenir de modifications chimiques directes pendant la formation de SSB *via* l'action de ROS (Demple & DeMott, 2002). Les radiations ionisantes représentent aussi une source majeure de formation d'extrémités 3' endommagées, et des drogues anti-tumorales telles que la bléomycine et la néocarzinostatine peuvent aussi générer des SSB contenant des extrémités 3' modifiées (Dedon *et al.*, 1992). Enfin, des extrémités 3' endommagées peuvent aussi venir d'une activité abortive de l'ADN topoisomérase I (TOP1) dans les cellules humaines (Leppard & Champoux, 2005).

Pour restaurer des extrémités 3'-OH et 5'-P normales au niveau des cassures, de nombreuses enzymes procaryotes et eucaryotes différentes sont utilisées pour traiter la grande variété de sites AP et d'extrémités 3' et/ou 5' modifiées (Krwawicz *et al.*, 2007) (Tableau 4).

**Tableau 4 : Protéines métabolisant les sites AP et extrémités de l'ADN.** (issu de Krwawicz *et al.*, 2007).

Family	Name (alternative name)	Size (aa)	Activity	Known substrates	Gene
<b>human nucleases</b>					
Xth/Exo III	APE1 (APEX, HAP1, Ref-1)	318	AP endonuclease, 3'→5' exonuclease, 3' diesterase, 3' phosphatase	AP-site, 3'-PUA, 3'-P, 3'-OH	APEX1
	APE2	518	AP endonuclease, 3'→5' exonuclease, 3' diesterase, 3' phosphatase	AP-site, 3'-PUA, 3'-P, 3'-OH	APEX2
RecJ	FEN1 (MF1)	380	5'dRP-lyase, 5'→3' exonuclease	5'-flap DNA, 5'-dRP	FEN1
NM23/ NDK	PNK (PNKP)	521	5'-kinase, 3'-phosphatase	3'-P	PNKP
PLD (phospholi- pase D)	TDP1	608	tyrosyl-DNA phosphodiesterase 1	3'-TOP1-SSB-peptide	TDP1
p53	TP53 (LFS1, TRP53, p53)	393	3'→5' exonuclease	3'-OH	TP53
HIT (histidine triad)	APTX (AOA, AOAI, AXAI, EAOH, EOAHA, FHA-HIT, FLJ20157, MGC1072)	342	5' exonuclease, 3' exonuclease	5'-P-adenylate group, 3'-PUA, 3'-P	APTX
HMG-box proteins	HMGB1 (DKFZp686A04236, HMGB1, HMGB3, SBP-1)	215	5'dRP-lyase	5'-dRP	HMGB1
DnaQ	TREX1 (AGS1, AGS5, CHBL, CREN, DKFZp434J0310, DRN3)	369 314	3'→5' exonuclease	3'-OH, 3'-mismatch	TREX1
	TREX2	236	3'→5' exonuclease	3'-OH, 3'-mismatch	TREX1
	<b>E. coli nucleases</b>				
Xth/Exo III	Xth (exo III)	268	AP-endonuclease, 3'→5' exonuclease, 3' diesterase, 3'-phosphatase	AP-site, 3'-PUA, 3'-P, 3'-OH	xth
Endo IV	Nfo (endo IV)	285	AP-endonuclease, 3'→5' exonuclease, 3' diesterase, 3'-phosphatase	AP-site, 3'-PUA, 3'-P, 3'-OH	nfo
RecJ	RecJ	577	5'dRP-lyase, 5'→3' exonuclease	5'-dRP, 5'-ssDNA	recJ
Endo V	Nfi (endo V)	223	3' exonuclease	AP-site, 3'-mismatch	nfi
NM23/ NDK	NDK	143	AP-endonuclease, 3'-phosphatase, 3'→5' exonuclease	AP-site, 3'-P, 3'→5' exonuclease	ndk

Parmi ces enzymes, les AP endonucléases (APEs) sont les enzymes principales responsables du traitement des intermédiaires réactionnels du BER. Elles génèrent des extrémités 3'-OH :

- soit par clivage hydrolytique de la liaison phosphodiester en 5' d'un site AP intact, donnant des extrémités 5'-dRP et 3'-OH.
- soit en enlevant les extrémités 3'PUA ou 3'P *via* les activités 3'-Phosphoestérase et 3'-Phosphatase, respectivement.

Les APEs de *E. coli* (Xth et Nfo) excisent efficacement les sites AP et éliminent efficacement les produits de  $\beta$ - et  $\beta,\delta$ -élimination des glycosylases bi-fonctionnelles. Les APEs de mammifères (APE1 et APE2) ont une forte activité AP endonucléase et 3'-Phosphoestérase, respectivement, mais ont une activité 3'-Phosphatase réduite (Burkovics *et al.*, 2006). Les

cellules de mammifères utilisent donc une Polynucléotide Kinase (PNK) pour le traitement des extrémités 3'-P, cette enzyme possédant des activités 5'-Kinase et 3'-Phosphatase.

De plus, Xth, APE1 et APE2 possèdent une activité exonucléase 3'→5' qui leur confère une capacité de correction des erreurs des polymérase (Burkovics *et al.*, 2006).

La nucléotide diphosphate kinase (NDK) de *E. coli* possède aussi des activités 3'-Phosphoestérase, 3'-Phosphatase et 3'→5'-Exonucléase (Goswami *et al.*, 2006), mais son rôle dans le BER n'est pas encore tout à fait clair.

D'autre part, la FIAP endonucléase 1 (FEN1) humaine possède deux activités :

- exonucléase 5'→3', impliquée dans la maturation des fragments d'Okazaki pendant la réplication chez les mammifères.

- endonucléase, qui enlève les extrémités flottantes contenant un 5'-dRP durant le LP-BER.

En outre, il a été démontré que la HMGB1 de mammifère interagit spécifiquement avec des intermédiaires du BER. Elle possède une faible activité 5'-dRP-lyase et stimule les activités des APEs et de FEN1 sur les substrats du BER (Prasad *et al.*, 2007).

Enfin, on peut noter que d'autres enzymes peuvent participer à ces processus, mais leurs activités ne seront pas discutées ici.

#### **I.5.4. Les ADN polymérase et leur implication dans la mutagénèse induite par les lésions**

##### **I.5.4.1. Caractéristiques générales et familles des ADN polymérase**

Les ADN polymérase sont des enzymes que l'on retrouve chez tous les organismes vivants, jouant un rôle central dans le maintien de l'intégrité du génome et dans sa transmission d'une génération à une autre *via* la réparation de l'ADN et sa réplication.

A l'heure actuelle, on connaît cinq ADN polymérase chez *Escherichia coli*, neuf chez *Saccharomyces cerevisiae*, et seize chez l'Homme (Tableau 5) (Bebenek *et al.*, 2004). La définition des noms de chacune de ces polymérase est restée conflictuelle durant de nombreuses années, puis a été harmonisée par Burgers et collaborateurs en 2001.

**Tableau 5 : Classification des ADN polymérasés.** (issu de Bebenek *et al.*, 2004).

Name	Family	Bacterial gene	Human gene	Yeast gene	Mol. Wt. (kDa) <sup>a</sup>	3' Exo	Other activities
Ec Pol I	A	<i>pol A</i>			103	+	5' Exonuclease
γ (gamma)			<i>POLG</i>	<i>MIPI</i>	140	+	dRPlyase
θ (theta)			<i>POLQ</i>	—	290	—	ATPase, helicase
ν (nu)			<i>POLN</i>	—	100	—	
Ec Pol II	B	<i>polB</i>			89	+	
α (alpha)			<i>POLA</i>	<i>POL1 (CDC17)</i>	165	—	Primase
δ (delta)			<i>POLD1</i>	<i>POL3 (CDC3)</i>	125	+	
ε (epsilon)			<i>POLE</i>	<i>POL2</i>	225	+	
ζ (zeta)			<i>POLZ (REV3)</i>	<i>REV3</i>	353	—	
Ec Pol III	C	<i>dnaE</i>			130	(separate subunit)	
β (beta)	X		<i>POLB</i>	—	39	—	dRP lyase AP lyase
λ (lambda)			<i>POLL</i>	<i>POL4 (POLX)</i>	66	—	dRP lyase, TdT
μ (mu)			<i>POLM</i>	—	55	—	TdT
TdT				<i>TdT</i>	56	—	
σ (sigma)			<i>POLS (TRF4-1)</i>	<i>TRF4</i>	60	—	
Ec Pol IV	Y	<i>dinB</i>			40	—	
Ec PolV		<i>umuC</i>			46	—	
η (eta)			<i>POLH</i> ( <i>RAD30A, XPV</i> )	<i>RAD30</i>	78	—	
ι (iota)			<i>POL1 (RAD30B)</i>	—	80	—	dRP lyase
κ (kappa)			<i>POLK (DINB)</i>	—	76	—	
Rev1			<i>REV1</i>	<i>REV1</i>	138	—	

<sup>a</sup>Deduced from protein primary structure.

Bien que participant toutes à la synthèse de l'ADN, ces enzymes sont très variées au niveau du nombre, du type et de la taille de leurs domaines et sous-unités, et sont connues pour interagir avec une grande variété de facteurs accessoires protéiques nécessaires pour leurs activités et régulations. De plus, leurs propriétés varient fortement, notamment au niveau de l'efficacité d'incorporation de nucléotides, du nombre de nucléotides insérés par événement de fixation, et de leur fidélité. Leur spécificité de substrats varie aussi grandement, certaines préférant remplir une lacune d'un seul nucléotide, d'autres préférant copier des nucléotides endommagés, et certaines participant à la synthèse de brins entiers d'ADN au sein de complexes multiprotéiques de la machinerie de réplication. En plus de l'activité de synthèse d'ADN, certaines polymérasés ont des activités additionnelles, telles qu'une exonuclease 3'→5' ou 5'→3', une 5'-dRPase, une ATPase, ou une primase (ARN polymérase ADN dépendante, permettant la synthèse des amorces d'ARN nécessaires à la synthèse d'ADN au cours de la réplication). Ces activités sont localisées soit sur des domaines différents d'une même sous-unité, soit sur des sous-unités différentes.

En se basant sur leur structure primaire et leurs sous-unités catalytiques, elles ont été classifiées en différentes familles distinctes :

- La famille A tient son nom du gène *polA* de *Escherichia coli* qui encode la Pol I, et comprend la polymérase répllicative du bactériophage T7, la Pol γ mitochondriale eucaryote, et les polymérasés humaines θ et ν.

- La famille B comprend la Pol II de *E. coli*, produit du gène *polB*, et des homologues telles que les polymérase répliquatives de bactériophages T4 et RB69, et les polymérase eucaryotes  $\alpha$ ,  $\delta$ ,  $\epsilon$  et  $\zeta$ .
- La famille C comprend la Pol III de *E. coli*, produit du gène *polC* (*dnaE*), et les polymérase homologues présentes chez la plupart des bactéries GRAM positives.
- La famille D (absente du Tableau 5) comprend les polymérase hétérodimériques d'Archées Pol II et Pol D.
- La famille X comprend la Pol  $\beta$  de mammifère, les polymérase eucaryotes  $\lambda$ ,  $\mu$  et  $\sigma$ , et une polymérase modèle-indépendante terminale transférase (TdT).
- La famille Y comprend les polymérase IV et V de *E. coli*, les polymérase eucaryotes  $\eta$ ,  $\kappa$ ,  $\iota$  et Rev1.

#### **I.5.4.2. Les ADN polymérase, la réparation de l'ADN, et la voie BER**

Nous nous intéresserons ici uniquement aux polymérase impliquées dans la voie de réparation par excision de bases, cependant une revue sur les polymérase impliquées dans les autres voies de réparation peut être trouvée dans Bebenek *et al.*, 2004.

Plusieurs voies de réparation de l'ADN impliquent l'excision de l'ADN endommagé, suivie d'une étape de resynthèse nécessitant une polymérase. Cependant, les substrats pour la resynthèse varient suivant la voie de réparation, nécessitant différentes polymérase.

Comme évoqué précédemment dans la description du BER, après l'action d'une ADN glycosylase mono- ou bi-fonctionnelle et d'une AP lyase/AP endonucléase, on obtient une lacune d'un nucléotide dans l'ADN, ce qui nécessite l'action d'une ADN polymérase pour la resynthèse de l'information génétique correcte sur le brin qui contenait la lésion.

Chez *E. coli*, la lacune nucléotidique générée est remplie par une seule ADN polymérase, la Pol I, ce qui n'est pas le cas chez les mammifères, où différentes polymérase peuvent intervenir.

Dans la voie majeure du BER humain, le SP-BER, la Pol  $\beta$  permet l'insertion d'un nucléotide sur l'extrémité 3'-OH et permet d'enlever le groupement 5'-dRP *via* son activité 5'dRPFase, donnant ainsi un substrat correct pour la ligase. Si l'extrémité 5'-dRP est modifiée ou non clivée par l'activité 5'-dRP de Pol  $\beta$ , le LP-BER est engagé et il y a synthèse avec déplacement de brin sur 2 à 13 nucléotides, générant une partie simple brin flottante qui sera ensuite enlevée par une endonucléase spécialisée pour ce type d'extrémités, la FEN1. Les

polymérase impliquées lors de cette synthèse sont les Pol  $\beta$ ,  $\delta$  et  $\epsilon$ . Chez *S. cerevisiae*, il n'y a pas de Pol  $\beta$ , et la synthèse se fait *via* les Pol  $\delta$ ,  $\epsilon$  et  $\alpha$ .

On peut noter que comme la Pol  $\beta$ , les Pol  $\lambda$ ,  $\iota$  et  $\gamma$  humaines possèdent aussi une activité 5'-dRPase, indiquant qu'elles peuvent aussi participer aux processus de réparation qui nécessitent d'enlever un groupement 5'-dRP.

Ainsi, il a été démontré que la Pol  $\lambda$  pouvait remplacer la Pol  $\beta$  lors de la réparation *in vitro* d'un ADN contenant un uracile. Cette polymérase est donc un bon candidat pour la synthèse du BER, étant peut-être partiellement redondante par rapport à la Pol  $\beta$  ou participant à des réactions du BER spécialisées impliquant des dommages spécifiques, des cellules ou tissus spécifiques, ou des phases spécifiques du cycle cellulaire. Par exemple, la haute affinité de la Pol  $\lambda$  pour les dNTP semble indiquer que cette polymérase pourrait être impliquée dans les synthèses d'ADN des cellules pour lesquelles le pool de dNTP est bas (ex : cellules quiescentes, ce qui est en accord avec la production plus élevée de cette polymérase dans ces cellules).

En tant que membre de la famille Y, la Pol  $\iota$  a été impliquée dans le passage des lésions bloquant la réplication (voir plus loin). D'autre part, certaines propriétés de cette polymérase sont en accord avec des processus particuliers du BER, et elle pourrait notamment intervenir dans la resynthèse des lacunes suivant la réparation par le BER des dU dans l'ADN issus de l'insertion de dUMP lors de la réplication. En accord avec cette hypothèse, Pol  $\iota$  interagit avec PCNA et est localisée aux « foci » de réplication. De plus, la Pol  $\iota$  possède la capacité inattendue d'incorporer un dGMP en face d'un dT de façon plus efficace que l'incorporation correcte d'un dAMP. De plus, sur les modèles contenant plusieurs dTs consécutifs, une incorporation préférentielle de dGMP en face du premier dT est suivie d'une incorporation préférentielle de dAMP en face du second dT. Cette spécificité remarquable amène l'hypothèse que la Pol  $\iota$  participerait à un BER particulier, remplaçant des dGs dans l'ADN après qu'ils aient été excisés de façon erronée par des glycosylases à partir de paires G/T ou G/U (provenant de la désamination de 5-Me-C ou de C dans l'ADN).

La quatrième ADN polymérase humaine possédant une activité 5'-dRPase est la Pol  $\gamma$ . Étant la seule polymérase connue pour être présente dans les mitochondries, cette polymérase est proposée comme étant la seule responsable de la synthèse d'ADN dans cet organe, incluant la réplication et la réparation.

#### **I.5.4.3. Les ADN polymérases de réplication de l'ADN endommagé**

Certains dommages peuvent ne pas être réparés suffisamment rapidement pour ne pas rencontrer la machinerie de réplication. Parmi ces dommages, certains distordent la géométrie de l'hélice d'ADN et des appariements de bases, ce qui empêche la réplication normale de l'ADN. Pour contrecarrer cet effet, les cellules possèdent des ADN polymérases capables de synthèse translésionnelle (TLS) (Ohmori *et al.*, 2001). La plupart des polymérases de TLS appartiennent à la famille Y, que l'on retrouve des bactéries jusqu'à l'homme. Ces polymérases n'ont pas d'activité exonucléase 3'→5' pour la correction des erreurs (activité d'édition) et n'ont qu'une faible sélectivité pour les nucléotides dans leur site actif. Ceci leur permet d'accommoder des lésions dans leur site actif, mais implique aussi qu'elles ne possèdent qu'une faible fidélité lors de la copie d'ADN non endommagé, impliquant donc une forte augmentation de la mutagenèse. Les différentes polymérases translésionnelles de la famille Y peuvent avoir des propriétés différentes, notamment au niveau des lésions qui peuvent être contournées ou non.

Trois polymérases sont impliquées dans la TLS chez *E. coli* : les Pol IV et V de la famille Y, et la Pol II de la famille B. Toutes les trois font partie de la réponse SOS aux stress environnementaux, et modulent la mutagenèse lésion-dépendante (Pagés *et al.*, 2002) et la capacité des cellules à survivre pendant une longue période de phase stationnaire (Goodman, 2002).

Les cellules humaines possèdent cinq polymérases pour la TLS : les Pol ζ, η, ι, κ et REV1 (Bebenek *et al.*, 2004). Elles appartiennent toutes à la famille Y, sauf la Pol ζ qui appartient à la famille B. De plus, la Pol μ de la famille X est capable de réaliser la TLS *in vitro*, mais cela n'a pas été vérifié *in vivo*. De ces six polymérases, seulement trois sont présentes chez *S. cerevisiae* : les Pol ζ, η et REV1.

Le contournement des lésions se fait suivant deux voies distinctes selon le type de lésion (Bebenek *et al.*, 2004). Dans le premier cas, une seule polymérase permet l'insertion de la base et l'extension de l'extrémité résultante. Dans le second cas, deux polymérases sont utilisées pour leur contournement, l'une pour l'insertion et l'autre pour l'extension. Ainsi, la TLS peut nécessiter de multiples échanges de polymérases sur l'ADN, et probablement aussi des échanges entre polymérases et 3'-exonucléases (activité présente intrinsèquement pour Pol δ et ε). Cependant, les mécanismes responsables de ces échanges ne sont pas encore bien compris, et impliquent probablement de nombreuses protéines accessoires (incluant REV1, le

$\beta$ -clamp de *E. coli* et le PCNA eucaryote, qui interagissent avec de nombreuses Polymérase translésionnelles).

#### **I.5.4.4. Les ADN polymérase répliquatives et leur implication dans la mutagénèse associée aux lésions**

La machinerie de répliquon de l'ADN non-endommagé nécessite l'utilisation de plusieurs polymérase et protéine accessoire.

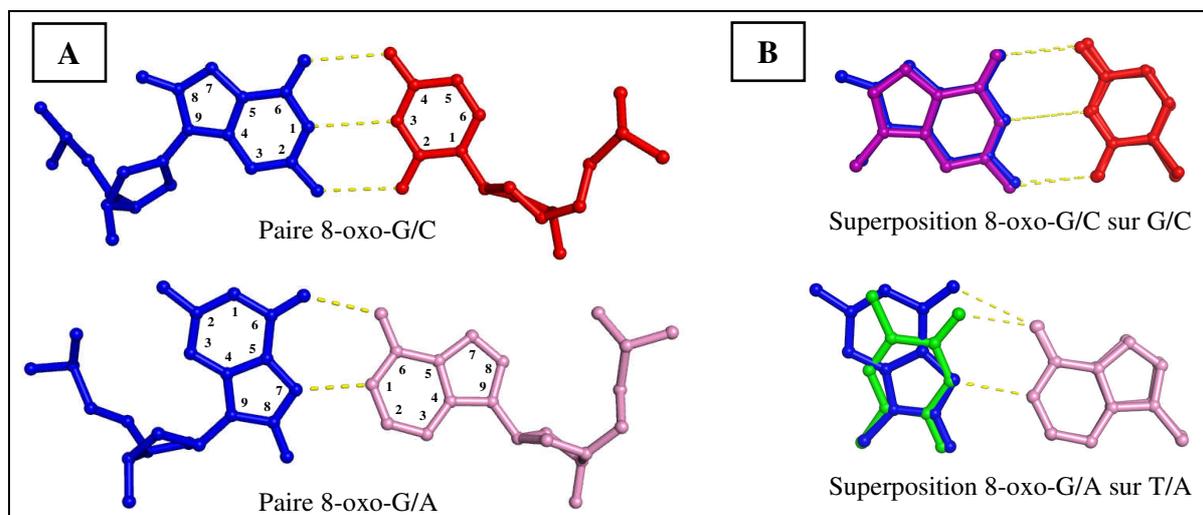
Chez *Escherichia coli*, la Pol III est responsable de la plupart de la synthèse pour les deux brins, et Pol I est impliquée dans le traitement des fragments d'Okazaki. Dans la plupart des bactéries GRAM positives, deux polymérase répliquatives sont présentes, PolC et DnaE, ayant une haute homologie de séquence avec la sous-unité alpha de Pol III.

Dans les cellule eucaryote, au moins trois polymérase participent à cette synthèse : les Pol  $\alpha$ ,  $\delta$  et  $\epsilon$ . La Pol  $\alpha$  produit les fragments d'Okazaki, fait une élongation de ces fragments pour donner de courtes amorces d'ADN, puis les Pol  $\delta$  et  $\epsilon$  prennent le relais pour le reste de la synthèse, avec une haute fidélité.

Les lésions capables de réaliser des appariements de base mimant des appariements classiques (Watson-Crick ou Hoogsteen) distordent peu l'ADN et le site actif des ADN polymérase répliquatives, et peuvent donc être prises en charge par la machinerie de répliquon de l'ADN non-endommagé (ex : Freisinger *et al.*, 2004). Cependant, certaines de ces lésions peuvent s'apparier avec différentes bases sans que l'activité d'édition des ADN polymérase ne détecte de mésappariement, notamment avec des bases qui ne sont pas complémentaires de celle dont la lésion est issue, induisant donc possiblement une mutagénèse.

L'exemple le plus étudié pour ce type de comportement est celui de la 8-oxo-dG (chapitre I.2 sur les agressions sur les acide nucléique et les dommages générés, ainsi que I.6 sur le système GO), qui est une lésion issue de guanine. Dans les cas où une polymérase répliquative procaryote ou eucaryote rencontre cette lésion, l'ADN polymérase peut insérer soit une cytosine en face de la lésion en *anti* (appariement Watson-Crick) soit une adénine en face de la lésion en *syn* (appariement de la face Watson-Crick de l'adénine avec la face Hoogsteen de la lésion) (Krahn *et al.*, 2003; Freisinger *et al.*, 2004; Hsu *et al.*, 2004) (Figure 18A). On peut d'ailleurs noter que ces appariements sont possibles car les paires de bases 8-oxo-G/C et 8-oxo-G/A miment bien les paires de bases G/C et T/A, respectivement (Figure 18B). Ceci amène une mutagénèse correspondant à des transversions d'une paire G/C vers une paire T/A

ou d'une paire T/A vers une paire G/C lors de la réplication (chapitre I.6 sur le système GO) (Michaels *et al.*, 1992b).



**Figure 18 :** (A) Représentations des appariements de la 8-oxo-G avec un C ou un A et (B) comparaison avec les appariements canoniques de l'ADN.

La 8-oxo-G est en bleu, la cytosine en rouge, l'adénine en rose, la guanine en violet et la thymine en vert. Les structures d'origine correspondent à des ADN libres résolues par cristallographie des rayons X. Les codes PDB d'origine sont : 8-oxo-G/C (183D, Lipscomb *et al.*, 1995), 8-oxo-G/A (178D, McAuley-Hecht *et al.*, 1994), G/C et T/A (150D, Leonard *et al.*, 1994). Les liaisons hydrogène sont représentées par des pointillés jaunes.

On peut noter que la préférence d'insertion de cytosine/adénine en face de la 8-oxo-G varie selon la polymérase considérée, avec une insertion privilégiée de cytosine pour la polymérase répliquatives de phage T7, les ADN polymérase de réparation Pol I de *E. coli* et Pol  $\beta$  humaine, la Pol  $\gamma$  mitochondriale, et les ADN polymérase translésionnelles humaines Pol  $\iota$  et  $\eta$ . Par contre, la polymérase de réparation Pol I de *Bacillus stearothermophilus*, les polymérase répliquatives humaine Pol  $\alpha$  et  $\delta$ , ainsi que la transcriptase inverse de HIV-1 ont le comportement inverse. Enfin, la polymérase Pol II de *E. coli* insère de façon indifférenciée les deux bases (Shibutani *et al.*, 1991; Pinz *et al.*, 1995; Lowe *et al.*, 1996; Furge *et al.*, 1997; Miller *et al.*, 2000; Krahn *et al.*, 2003; Hsu *et al.*, 2004).

De façon plus surprenante, on peut noter que ces différences d'insertion en face de la 8-oxo-G par une même ADN polymérase peuvent aussi être espèce-dépendantes, comme illustré précédemment par la différence de comportement de la Pol I de *E. coli* et *Bacillus stearothermophilus* (Lowe *et al.*, 1996; Hsu *et al.*, 2004).

De même, on peut noter que la préférence d'insertion de 8-oxo-dGMP à partir de 8-oxo-GTP en face d'une cytosine ou d'une adénine varie selon la polymérase considérée. En effet, la Pol  $\beta$  humaine préfère insérer un 8-oxo-dGMP en face d'un A, alors que la Pol III de *E. coli*

insère aussi bien le 8-oxo-dGMP en face d'un A ou d'un C (Maki *et al.*, 1992; Miller *et al.*, 2000).

En outre, pour une même polymérase, on peut noter que la préférence d'insertion d'un 8-oxo-dGMP en face d'un A ou d'un C n'est pas forcément identique à celle de l'insertion d'un A ou d'un C en face d'une 8-oxo-G dans l'ADN, comme illustré par le comportement de la Pol  $\beta$  humaine. En effet, cette polymérase préfère insérer un 8-oxo-dGMP en face d'un A plutôt que d'un C, alors qu'elle préfère insérer un C plutôt qu'un A en face d'une 8-oxo-G (Shibutani *et al.*, 1991; Miller *et al.*, 2000).

Enfin, on peut noter que l'efficacité d'incorporation d'un 8-oxo-dGMP en face d'une base normale ou d'incorporation d'une base normale en face d'une 8-oxo-G est réduite par rapport à l'efficacité d'incorporation d'une base normale en face d'une base normale complémentaire. Par exemple, l'efficacité d'insertion d'un 8-oxo-GMP par la Pol III de *E. coli* en face d'un A ou d'un C n'est que de 3-4 % de l'efficacité d'incorporation en face d'un A ou d'un C d'une base normale complémentaire (Maki *et al.*, 1992). De même, la Pol  $\beta$  humaine préfère insérer un dCMP en face d'une guanine plutôt qu'en face d'une 8-oxo-G, et préfère insérer un dGMP plutôt qu'un 8-oxo-GMP en face d'une cytosine (Miller *et al.*, 2000). Cette caractéristique n'est pas étonnante étant donné que les dNTPs normaux sont les substrats physiologiques des ADN polymérases, cependant on peut noter qu'elle permet de limiter l'effet mutagène des lésions.

Toutes ces propriétés des ADN polymérases ne tiennent compte que des possibilités d'appariement 8-oxo-G/C et 8-oxo-G/A de la 8-oxo-G, or la première publication sur les effets mutagènes de la 8-oxo-G indiquait que la Pol I de *E. coli* pouvait insérer de façon indifférenciée les quatre bases en face de cette lésion (Kuchino *et al.*, 1987). Cependant, des études plus récentes ont eu des résultats différents, avec une insertion faible ou non détectée de dGMP en face de la 8-oxo-G par différentes polymérases (Shibutani *et al.*, 1991; Korniyushyna *et al.*, 2002), ce qui a amené les études ultérieures à ne considérer que les appariements avec une adénine ou une cytosine. Une explication possible des résultats de la publication de Kuchino et collaborateurs est que la 8-oxo-G aie subit une seconde oxydation durant la manipulation, donnant ainsi les composés Imidazolone (Iz), Iminoallantoïne (Ia), Guanidinohydantoïne (Gh) et spiroiminohydantoïne (Sp) (Cadet *et al.*, 2003), qui ont été décrits comme pouvant s'apparier à une guanine et donc donner des transversions G/C vers C/G (Korniyushyna *et al.*, 2002; Kino *et al.*, 2005). Cependant, la possibilité d'appariement 8-oxo-G/G reste possible (Culp *et al.*, 1989; Thivyanathan *et al.*, 2003). En effet, les résultats de Korniyushyna et collaborateurs en 2002 indiquent une insertion 100 fois plus faible de GMP en

face de la 8-oxo-G par rapport aux insertions de CMP et AMP par la Pol I de *E. coli*, ce qui rend peu probable la possibilité d'avoir des paires 8-oxo-G/G créées *in vivo* sans toutefois l'exclure. De plus, l'incorporation de 8-oxo-GMP en face d'un G n'a pas été testée, alors que l'on sait désormais que le comportement des ADN polymérase quant à l'insertion d'un 8-oxo-GMP en face d'une base donnée est différent du comportement de ces mêmes polymérase quant à l'insertion d'une base donnée en face de la 8-oxo-G dans l'ADN.

Depuis de nombreuses années, la 8-oxo-G a été considérée comme étant la seule lésion responsable des transversions G/C vers T/A. En fait, cette conception était erronée, venant du fait que les études réalisées précédemment utilisaient le *N*<sup>7</sup>-Me-FapyG comme modèle d'étude pour le FapyG, ce dernier n'ayant été synthétisé que récemment (Haraguchi *et al.*, 2001). En l'occurrence, le *N*<sup>7</sup>-Me-FapydG est majoritairement bloquant pour la réplication, avec un effet mutagène négligeable, correspondant à une fréquence extrêmement faible de transversions G/C vers T/A et G/C vers C/G (Boiteux *et al.*, 1983; O'Connor *et al.*, 1988; Tudek *et al.*, 1992; Tudek *et al.*, 1999; Tudek, 2003; Asagoshi *et al.*, 2002). Cela a amené la pensée que le FapyG était plutôt létal que mutagène.

Cependant des études plus récentes ont montré que le FapyG donne aussi un effet promutagène non négligeable, avec une toxicité et une capacité à induire une synthèse translésionnelle variable selon la séquence entourant la lésion et selon la polymérase considérée (Graziewicz *et al.*, 2000). En effet, le FapyG peut aussi amener la Pol I de *E. coli* à incorporer un dCMP ou un dAMP en face du FapyG (aucune structure de ces appariements n'est disponible), provoquant donc le même effet mutagène que la 8-oxo-G (chapitre I.6 sur le système GO) (Wiederholt *et al.*, 2002b). Ainsi, si on considère les efficacités d'incorporation et d'extension avec un dAMP ou un dCMP en face du FapyG, on voit que la probabilité d'avoir un dC en face d'un FapyG n'est que 5 fois supérieure à celle d'avoir un dA en face d'un FapyG, soit un niveau comparable au comportement de cette polymérase avec la 8oxoG (3,4 fois) (Lowe *et al.*, 1996; Wiederholt *et al.*, 2002b). D'autre part, une étude réalisée sur des monocytes tumoraux humains a montré par HPLC/GC-MS que le FapyG était présent en quantités 2 à 3 fois supérieures que la 8-oxo-G après irradiation  $\gamma$  (Pouget *et al.*, 2000). Une étude plus récente a comparé les effets mutagènes causés par les Fapys et les 8-oxo-purines des cellules de foie de singe (cellules COS-7) (Kalam *et al.*, 2006). Une incorporation significative d'adénine opposée au FapyG a été détectée dépendamment du contexte de séquence, induisant des transversions G/C vers T/A. La 8-oxo-G a aussi amené des transversions G/C vers T/A dans le même contexte de séquence qu'avec le FapyG, mais le FapyG a montré une mutagénicité de 25 à 30% plus importante que la 8-OH-G. Contrairement aux découvertes sur

les cellules COS-7, une étude récente montre que le FapyG est cinq fois moins mutagène dans *E. coli* que la 8-oxo-G (Patro *et al.*, 2007). Enfin, il a été montré récemment que la Pol I (Fragment Klenow) de *E. coli* pouvait utiliser un Fapy-dGTP pour insérer un FapyG en face d'un C ou d'un A (25 fois moins efficace avec un A), cependant la présence *in vivo* de Fapy-dGTP n'a pas été démontrée (Imoto *et al.*, 2006).

Tout ceci montre, contrairement aux conceptions antérieures, que les effets biologiques du FapyG diffèrent quasiment totalement de ceux du FapyG méthylé en N7, et que le FapyG intervient de façon au moins aussi importante que la 8-oxo-G dans la mutagénèse associée aux dommages oxydatifs, notamment dans les cellules de mammifères.

Enfin, on peut noter que de nombreuses autres lésions créées sur l'ADN amènent des erreurs de réplication plus ou moins fréquentes, étant donc potentiellement promutagènes. C'est notamment le cas pour les lésions FapyA, N<sup>7</sup>-Me-FapyA, site AP, Thymine glycol, O<sup>6</sup>-Méthyle-guanine, O<sup>4</sup>-Méthyle-guanine, 1,N<sup>2</sup>-éthéno-guanine, N<sup>2</sup>,3-éthéno-guanine, éthéno-cytosine, éthéno-adénine, 2-hydroxyadenine (2-OH-A), 8-hydroxyadenine (8-OH-A), 5-hydroxycytosine (5-OH-C), et 5-hydroxyuracil (5-OH-U), cette liste n'étant pas exhaustive (Tudek *et al.*, 1999; Delaney *et al.*, 2002; Olinski *et al.*, 2002; De Bont *et al.*, 2004).

### **I.5.5. Les ADN ligases**

Après avoir restauré des extrémités 5'-P et 3'-OH puis remplis la lacune par une ADN polymérase, la fin du processus de réparation BER nécessite l'intervention d'une ADN ligase ATP-dépendante.

Chez la bactérie *E. coli*, la ligase I permet d'effectuer cette étape de ligation d'ADN dans les deux voies SP- et LP-BER (Krwawicz *et al.*, 2007). Par contre, les cellules de mammifères contiennent deux ligases différentes recrutées aux sites de réparation, les ADN ligases 1 et 3 $\alpha$  (Krwawicz *et al.*, 2007), dont le domaine catalytique responsable de la réaction de ligation est hautement conservé (Mortusewicz *et al.*, 2006). Deux cas de figures se présentent donc :

- Pour le SP-BER (et la réparation des SSBs), l'ADN ligase 3 $\alpha$ , qui interagit avec la polymérase  $\beta$  via la protéine XRCC1, lie la coupure pour restaurer la séquence d'ADN originelle (Schärer *et al.*, 2001; Mortusewicz *et al.*, 2006; Krwawicz *et al.*, 2007).

- Pour le LP-BER, la ligase 1 est recrutée par interaction avec le Proliferating Nuclear Cell Antigen (PCNA), et catalyse la réaction de ligation après action de FEN1 (Schärer *et al.*, 2001; Mortusewicz *et al.*, 2006; Krwawicz *et al.*, 2007).

On peut noter encore une fois que les interactions des ADN ligases avec XRCC1 et PCNA, qui eux-mêmes interagissent avec de nombreux constituants du BER, montrent que cette voie de réparation est hautement concertée (voir Figure 14).

## **I.6. Le système GO**

### **I.6.1. Le système GO procaryote et la 8-oxo-guanine**

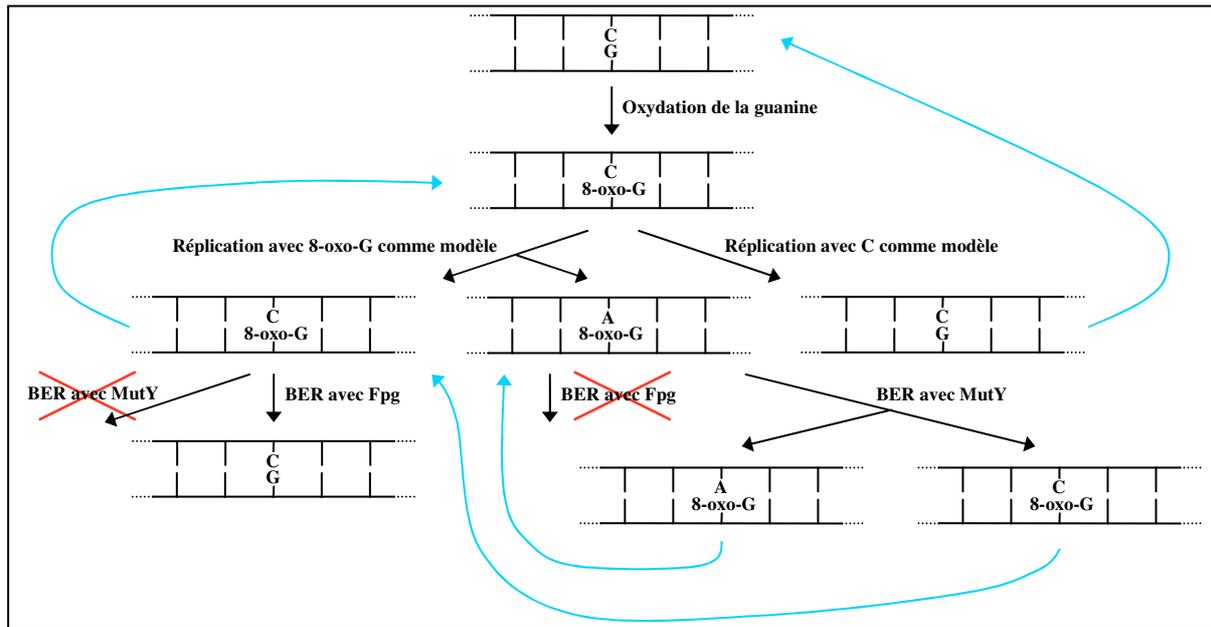
Le système GO (pour Guanine Oxydée) est une voie de réparation permettant d'éviter d'avoir à répliquer la 7,8-dihydro-8-oxo-guanine (8-oxo-G), qui est une lésion ayant des propriétés d'appariement particulières qui amènent une possible mutagénèse lors de sa réplication (chapitre I.5.4.4 sur les polymérase répliquatives et leur implication dans la mutagénèse associée aux lésions). En effet, si une guanine oxydée en 8-oxo-G est prise comme modèle lors de la réplication, on verra apparaître un taux élevé de transversion G/C vers T/A en l'absence de réparation. De même, un 8-oxo-dGMP peut être inséré par l'ADN polymérase répliquative en face d'un modèle adénine, induisant un taux élevé de transversion A/T vers C/G en l'absence de réparation (Figures 19 et 20, adaptées de Michaels *et al.*, 1992b) (chapitre I.5.4 sur les ADN polymérase et leur implication dans la mutagénèse induite par les lésions).

Le système GO procaryote est constitué d'au moins trois protéines, MutM (Fpg), MutY et MutT, et la définition de ce système est venue de plusieurs études sur l'effet mutateur de l'inactivation de ces trois protéines. En effet, il a été démontré que si le gène *mutT* est KO, on a augmentation des transversions A/T vers C/G (Yanofsky *et al.*, 1966). De même, les KO sur les gènes *mutM* et/ou *mutY* augmentent les transversions G/C vers T/A, et ce d'autant plus si on a un double KO sur ces deux gènes (Cabrera *et al.*, 1988; Nghiem *et al.*, 1988; Michaels *et al.*, 1992a). Toutes ces informations, alliées aux études sur le comportement de la 8-oxo-G lors de la réplication, ont suggéré que Fpg, MutY et MutT étaient couplées dans un système de réparation commun permettant d'éviter les effets promutagènes de la 8-oxo-G. Cependant, les activités des trois protéines dans le système sont différentes, et il est donc nécessaire de les définir (Figures 19 et 20).

L'intervention de Fpg, une ADN glycosylase/AP lyase bi-fonctionnelle de la superfamille structurale Fpg/Nei (chapitre I.5.2.3 sur les superfamilles structurales d'ADN glycosylases), dans le système GO correspond à l'excision de la 8-oxo-G dans les paires 8-oxo-G/C (Figure 19) (Michaels *et al.*, 1992b). Cette enzyme est la seule parmi les trois protéines du système à pouvoir enlever une 8-oxo-G de l'ADN. En outre, Fpg n'est pas (ou très peu) active

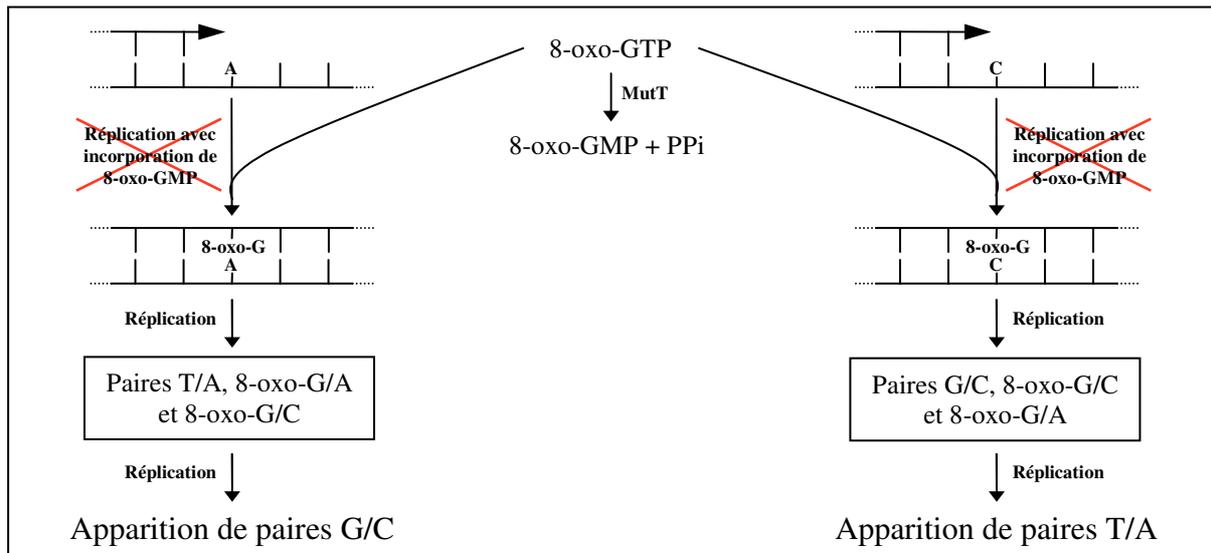
sur les paires 8-oxo-G/A (la spécificité de Fpg sera discutée plus en détail dans le chapitre suivant, qui traitera exclusivement de cette protéine). Cette spécificité de Fpg vis-à-vis de ses substrats est cruciale, car si elle était active sur les paires 8-oxo-G/A, alors elle aurait un caractère promutagène, participant à la fixation définitive de la transversion G/C vers T/A (Figure 19). Au final, l'activité de Fpg permet d'éliminer les paires 8-oxo-G/C au profit de paires G/C, éliminant donc la possibilité de mutagénèse associée à la réplication de la 8-oxo-G, d'où l'augmentation de transversions G/C vers T/A si le gène *mutM* est KO.

L'intervention de MutY, une ADN glycosylase monofonctionnelle de la superfamille structurale HhH-GPD d'ADN glycosylases (Huffman *et al.*, 2005), dans le système GO correspond à l'excision de l'adénine dans les paires 8-oxo-G/A (Figure 19). Initialement, MutY avait été décrite comme excisant les adénines des mésappariements G/A (Au *et al.*, 1989), permettant de restaurer des paires G/C. Plus récemment, il a été montré que MutY avait aussi une activité sur les paires 8-oxo-G/A, mais ne permettait pas l'excision des cytosines dans les paires 8-oxo-G/C (Michaels *et al.*, 1992a). Cette activité permet donc de convertir progressivement les paires 8-oxo-G/A en paires 8-oxo-G/C, qui peuvent être ensuite traitées par Fpg pour éliminer la 8-oxo-G. De plus, la limitation du nombre de paires 8-oxo-G/A au profit de paires 8-oxo-G/C permet de diminuer la possibilité de transversions G/C vers T/A issues de la réplication des paires 8-oxo-G/A. Une autre caractéristique de MutY est qu'elle reste fixée sur son produit, le site AP, empêchant Fpg d'exciser la 8-oxo-G en face d'un site AP. Cette caractéristique est importante, car si Fpg était amenée à exciser la 8-oxo-G dans une paire 8-oxo-G/AP, alors on aurait perte complète de l'information pour cette paire de base (Michaels *et al.*, 1992b). On peut donc noter que les activités de Fpg et MutY sont complémentaires, ce qui explique que l'effet du double mutant sur les gènes *mutM* et *MutY* est plus important que la somme des effets des simples mutants correspondants (Michaels *et al.*, 1992a).



**Figure 19 : Première partie du système GO de *E. coli*.** Permet d'éviter les transversions G/C vers T/A lors de la réplication d'une 8-oxo-G issue de l'oxydation d'une guanine dans l'ADN (cette transversion n'est pas représentée dans le schéma, mais serait le résultat d'un second cycle de réplication sur la paire 8-oxo-G/A). Les flèches noires représentent une réaction, les flèches bleues représentent des molécules équivalentes, et les croix rouges représentent une réaction inefficace.

L'intervention de MutT, dans le système GO correspond à la dégradation du 8-oxo-dGTP en 8-oxo-dGMP afin d'éviter l'incorporation de 8-oxo-dGMP lors de la réplication des adénines et cytosines (Maki *et al.*, 1992) (Figure 20). Cette enzyme n'est donc pas une ADN glycosylase mais une 8-oxo-dGTP-ase. De plus, MutT est active sur la dégradation du 8-oxo-dGDP en 8-oxo-dGMP, et sur les ribonucléosides correspondants (Tableau 6). Si on considère la Figure 20, alors on voit que l'inactivation du gène *mutT* devrait augmenter la fréquence de transversions G/C vers T/A et T/A vers G/C, or le phénotype correspondant à ce KO correspond uniquement à une augmentation de transversions T/A vers G/C. Ceci s'explique par le fait que l'incorporation de 8-oxo-dGMP en face d'un C ne pose pas de problème particulier étant donné que Fpg et MutY vont restituer une paire G/C en évitant la transversion vers une paire T/A. Par contre, si on a incorporation de 8-oxo-dGMP en face d'un A, alors MutY va être promotagène dans ce cas, induisant une transversion d'une paire T/A vers une paire G/C.



**Figure 20 :** Seconde partie du système GO de *E. coli*. Permet l'élimination du 8-oxo-GTP, et donc d'éviter les transversions T/A vers G/C et G/C vers T/A issues de l'insertion de 8-oxo-GMP par une polymérase répliquative (ces transversions sont représentées dans le schéma, mais ne sont pas sentées arriver grâce à l'activité de MutT). Les flèches noires représentent une réaction et les croix rouges indiquent une réaction inefficace.

### I.6.2. Le système GO étendu, ou « traitement antimutagène bipartite de la 8oxoG »

Si on considère le système GO décrit précédemment, on se rend vite compte que l'apparition de 8-oxo-G *in situ* dans l'ADN ou *via* l'incorporation de 8-oxo-dGMP impliquent des effets pro ou antimutagènes de MutY. En effet, cette enzyme a un effet antimutagène dans le cas de l'apparition de 8-oxo-G par oxydation d'une guanine dans l'ADN (Figure 19). Par contre, dans le cas d'un stress oxydant fort, il se peut qu'il y ait trop de 8-oxo-dGTP créé pour que MutT arrive à dégrader la totalité de ce nucléotide triphosphate modifié. Dans ce cas, on peut avoir insertion de 8-oxo-dGMP en face d'une adénine lors de la réplication d'une paire A/T, et l'activité de MutY sur la paire 8-oxo-G/A serait dans ce cas promutagène, induisant une transversion T/A vers G/C.

D'autre part, des études ont mis en évidence d'autres protéines étant impliquées dans la lutte contre les effets mutagènes de la 8-oxo-G. En l'occurrence, les protéines Nei (Endonucléase VIII) et Nth (Endonucléase III) sont capables d'exciser la 8-oxo-G dans les paires 8-oxo-G/A (Yanofsky *et al.*, 1966; Hazra *et al.*, 2000; Hazra *et al.*, 2001; Matsumoto *et al.*, 2001). Une nouvelle fois, on peut noter que l'activité de Nei et Nth d'excision de la 8-oxo-G dans une paire 8-oxo-G/A est promutagène dans le cas de l'apparition d'une 8-oxo-G *in situ*

dans l'ADN, favorisant les transversions G/C vers T/A. Par contre, dans le cas où un 8-oxo-dGMP serait inséré dans l'ADN en face d'un A, Nei et Nth préviendraient les transversions A/T vers C/G et seraient donc antimutagènes. On peut aussi noter que Nei possède une faible activité d'excision de la 8-oxo-G dans les paires 8-oxo-G/C (Hazra *et al.*, 2000), ce qui confère à la bactérie une redondance pour l'activité de Fpg sur la 8-oxo-G.

Il est donc paradoxal que des enzymes dont le rôle est d'éviter la mutagénèse puissent avoir des propriétés promutagènes selon le cas. Toutes ces informations ont amené l'hypothèse que le système GO était probablement plus compliqué que ce que l'on pensait jusqu'alors, avec un système permettant de réparer différemment les 8-oxo-G générées *in situ* dans l'ADN ou incorporées sous forme de 8-oxo-dGMP. Cette nouvelle façon de voir les choses a été nommée « traitement antimutagène bipartite de la 8oxoG » par Hazra et collaborateurs en 2001, tenant compte cette fois de l'origine des 8-oxo-G. Par contre, la façon dont les procaryotes sélectionneraient les 8-oxo-G à réparer ou non par une même protéine n'a pas été élucidée, même si l'hypothèse d'un couplage entre réplication et réparation a été émise.

### **I.6.3. Seconde extension du système GO pour le traitement des paires 8-oxo-G/G**

Le système GO initial et son évolution « traitement antimutagène bipartite de la 8-oxo-G » tiennent uniquement compte des effets des appariements 8-oxo-G/C et 8-oxo-G/A sur les transversions G/C vers T/A et T/A vers G/C. Or, comme évoqué précédemment dans le chapitre I.5.4 sur les polymérases et leur implication dans la mutagénèse associée aux lésions, la possibilité d'un appariement 8-oxo-G/G *in vivo* ne peut être écartée.

Ainsi, on sait que la guanine est la base nucléique la plus sensible aux dommages oxydatifs, et que la 8-oxo-G est la lésion créée en plus grande quantité par les stress oxydatifs. De plus, ces stress amènent une grande quantité de transversions G/C vers C/G, avec un effet antimutateur de Fpg pour ces transversions (McBride *et al.*, 1991; Costa de Oliveira *et al.*, 1992; Ono *et al.*, 1995). Tout ceci pourrait donc indiquer que l'appariement 8-oxo-G/G participe à l'apparition des transversions G/C vers C/G *in vivo*, cependant on sait que d'autres lésions induites par les stress oxydatifs peuvent donner ce type de mutations (chapitre I.5.4 sur les polymérases et leur implication dans la mutagénèse induite par les lésions), dont certaines sont substrat de Fpg.

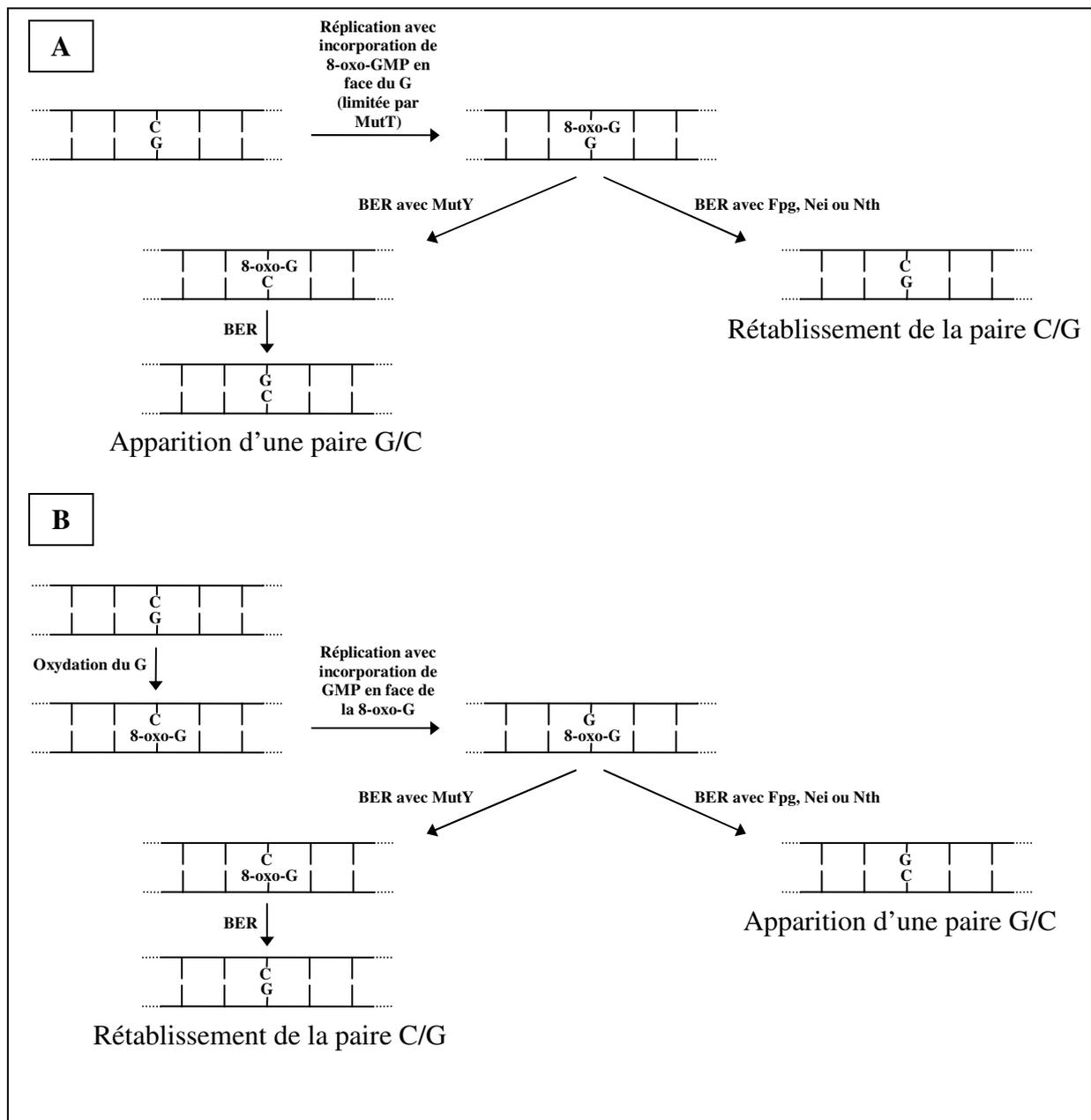
Des arguments moins contestables quant à la présence *in vivo* de paires 8-oxo-G/G sont trouvés dans l'étude de la spécificité de substrat des enzymes du système GO. Ainsi, on sait que Nei et Nth de *E. coli* excisent les 8-oxo-G appariées à une guanine avec une efficacité

dépassant celle de l'excision de cette lésion dans les paires 8-oxo-G/A (Hazra *et al.*, 2000; Hazra *et al.*, 2001; Matsumoto *et al.*, 2001). De même, on sait que Fpg de *E. coli* possède une activité d'excision de la 8-oxo-G sur les paires 8-oxo-G/G, et ceci de façon plus efficace que sur la paire 8-oxo-G/C, qui est pourtant considérée comme le substrat canonique de l'enzyme (Tchou *et al.*, 1994; Asagoshi *et al.*, 2000). Enfin, MutY possède une activité d'excision des guanines appariées aux 8-oxo-G, même si cette activité est moins efficace que celle d'excision de adénines appariées aux 8-oxo-G (Zhang *et al.*, 1998). Il paraît donc improbable que plusieurs enzymes reconnaissent la paire 8-oxo-G/G comme un substrat spécifique si cette paire de base n'existe pas *in vivo*.

Toutes ces informations tendent donc à confirmer que l'appariement 8-oxo-G/G est réellement possible *in vivo*, ce qui amène une nouvelle fois le problème de la reconnaissance de la 8-oxo-G créée *in situ* dans l'ADN, ou incorporée *via* le 8-oxo-dGMP issu du 8-oxo-dGTP.

En effet, si on considère les Figures 21A et 21B, alors on voit que l'activité de MutY serait promutagène dans le cas d'une incorporation d'un 8-oxo-dGMP en face d'un G, et antimutagène dans le cas d'une incorporation d'un dGMP en face d'une 8-oxo-G. Inversement, les activités de Fpg, Nth et Nei seraient promutagènes dans le cas d'une incorporation d'un dGMP en face d'une 8-oxo-G, et antimutagènes dans le cas d'une incorporation d'un 8-oxo-dGMP en face d'un G. Ces propriétés sont donc à nouveau en faveur du modèle du « traitement antimutagène bipartite de la 8-oxo-G ».

En outre, le triple mutant pour Fpg, Nth et Nei donne une augmentation de transversions G/C vers C/G plus importante que les simples et doubles mutants correspondants. Ces trois enzymes ont donc un effet antimutateur évitant ces transversions (Matsumoto *et al.*, 2001). D'autre part, les études sur l'insertion de nucléotides en face d'une 8-oxo-G par les ADN polymérases (chapitre I.5.4 sur les ADN polymérases et leur implication dans la mutagénèse induite par les lésions) indiquent que l'insertion de dGMP est, au mieux, peu efficace en face d'une 8-oxo-G. Tout ceci est plutôt en accord avec une insertion de 8-oxo-dGMP en face d'un G lors de la réplication d'une paire G/C (Figure 21A), et en désaccord avec l'insertion de dGMP en face d'une 8-oxo-G générée *in situ* dans l'ADN (Figure 21B). Cependant, l'absence d'étude sur l'efficacité d'insertion de 8-oxo-dGMP en face d'un G par les ADN polymérases nous empêche de conclure de façon définitive sur la voie d'apparition des paires 8-oxo-G/G, et sur leur intervention dans l'apparition de transversions G/C vers C/G.



**Figure 21 : Effet de l'appariement 8-oxo-G/G sur la mutagenèse associée au traitement de cette paire de bases par les enzymes du système GO. (A) Effet induit par l'incorporation de 8-oxo-GMP en face d'un G. (B) Effet induit par la génération de 8-oxo-G *in situ* dans l'ADN.**

#### **I.6.4. Le système GO chez l'Homme**

Le système GO de défense contre les effets mutagènes de la 8-oxo-G est conservé depuis la bactérie jusqu'à l'Homme. Cependant, des différences notables sont observables au niveau du type et du nombre de protéines impliquées dans ce système chez les procaryotes et les eucaryotes. Par exemple, les homologues fonctionnels procaryotes et humains ne sont pas forcément de la même superfamille structurale, avec des spécificités de substrat généralement

différentes, et le nombre de protéines impliquées chez l'Homme est plus important. On peut aussi noter qu'il existe aussi des différences entre les systèmes GO des différents eucaryotes, notamment entre certaines levures (*Saccharomyces cerevisiae*) et les eucaryotes supérieurs tels que l'homme, mais que ces différences ne seront pas détaillées ici. Ce sous-chapitre fera donc uniquement un parallèle succinct entre les systèmes GO bactérien et humain.

Tout d'abord, l'homologue fonctionnel humain de Fpg dans le système GO est hOgg1 (pour Oxo-guanine-glycosylase 1), et excise aussi les 8-oxo-G appariées à une cytosine. hOgg1 n'est cependant pas un homologue structural de Fpg, étant donné qu'elle fait partie de la superfamille structurale HhH-GPD d'ADN glycosylases (chapitre I.5.2.3 sur les superfamilles structurales d'ADN glycosylases) (Fromme *et al.*, 2004; Huffman *et al.*, 2005). De plus, hOgg1 n'est l'homologue fonctionnel de Fpg que pour son activité glycosylase, étant donné que cette enzyme ne catalyse qu'une  $\beta$ -élimination lors de son activité AP lyase, contre une  $\beta,\delta$ -élimination pour Fpg (Asagoshi *et al.*, 2000).

Ensuite, l'homologue fonctionnel humain de MutY dans le système GO est MYH (pour MutY-Homolog), et excise aussi les adénines appariées à une 8-oxo-G. De plus, MYH est aussi un homologue structural de MutY, appartenant à la même famille structurale HhH-GPD d'ADN glycosylases (Schärer *et al.*, 2001).

Enfin, MutT possède plusieurs homologues fonctionnels chez l'Homme, appelés MTH1, MTH2 (pour MutT-Homolog), et NUDIX5 (pour NuDIX type 5), et qui remplissent chacune une partie de l'activité de MutT. Les activités de ces trois protéines en comparaison de l'activité de MutT sont résumées dans le tableau suivant (Tableau 6, issu de Arczewska *et al.*, 2007).

**Tableau 6 : Activité de MutT de *E. coli* et de ses homologues fonctionnels humains.**

<i>E. coli</i> protein	Function	Human homologue	Function	Gene (other aliases)
MutT	8-oxodGTPase, active also on 8-oxodGDP, 8-oxoGTP and 8-oxoGDP	MTH1 (NUDT1)	8-oxodGTPase; active also on 2-oxodATP, 8-oxodATP, 8-Cl-dGTP, 8-oxoGTP, 2-oxoATP; inhibited by 8-oxodGDP and 2-oxodADP	MTH1 (NUDT1)
		MTH2 (NUDT15)	8-oxodGTPase?	MTH2 (NUDT15; FLJ10956; MGC104352; RP11-90M2.1)
		NUDT5	8-oxodGDPase; primarily identified as ADP-sugar pyrophosphatase	NUDT5 (YSA1H; hYSAH1)

En plus des enzymes évoquées précédemment, qui constituent le cœur du système GO humain par analogie avec la définition de Michaels et collaborateurs en 1992b, d'autres enzymes participent au système GO étendu chez l'homme. Ainsi, on trouve une seconde Ogg,

nommée hOgg2, qui est un homologue structural et fonctionnel de Nei dans le système GO, faisant donc aussi partie de la famille structurale HhH-GPD d'ADN glycosylases, et qui permet l'excision des 8-oxo-G appariées à des guanines ou des adénines dans l'ADN (Hazra *et al.*, 1998; Hazra *et al.*, 2001; Arczewska *et al.*, 2007). De même, il a été montré que les enzymes NEIL1 et 2 (Pour Nei-Like) et NTH1 (pour Nth-Homolog) possédaient une légère activité d'excision de la 8-oxo-G, notamment dans les paires 8-oxo-G/G, ce qui suggère leur possible participation au système (Matsumoto *et al.*, 2001; Katafuchi *et al.*, 2004; Parsons *et al.*, 2005; Hazra *et al.*, 2006).

On peut aussi noter que, comme pour le système GO procaryote, les enzymes du système GO humain sont aussi probablement capables de discriminer les 8-oxo-G créées *in situ* dans l'ADN de celles venant de l'incorporation de 8-oxo-dGMP (Hazra *et al.*, 2001). En l'occurrence, l'hypothèse d'une réparation couplée à la réplication a été émise, notamment pour MYH et NEIL1, qui ont été montrées comme étant stimulées par la protéine PCNA, qui est une protéine plateforme de la réplication (Dou *et al.*, 2008).

#### **I.6.5. Le système GO procaryote et le FapyG**

Comme décrit dans le chapitre précédent sur les polymérases et leur implication dans la mutagénèse induite par les lésions, la 8-oxo-G n'est pas la seule lésion induisant des transversions G/C vers T/A. Notamment, le FapyG a été décrit comme pouvant induire une telle mutagénèse (Wiederholt *et al.*, 2002b). Il est donc raisonnable de penser que les organismes possèdent un système comparable au système GO pour lutter contre les effets mutagènes du FapyG.

Ainsi, il est connu que Fpg excise efficacement les FapyG des paires FapyG/C, et ce de façon plus efficace que pour la paire 8-oxo-G/C (Wiederholt *et al.*, 2003). De plus, Fpg possède une sélectivité pour les FapyG dans les paires FapyG/C, avec une activité 17 fois plus importante que pour les paires FapyG/A, ce qui correspond à une spécificité comparables à celle pour la 8-oxo-G. On peut d'ailleurs déduire de ces résultats que l'augmentation des transversions G/C vers T/A dans les souches KO pour le gène *mutM* ne vient pas uniquement de la 8-oxo-G, mais aussi du FapyG.

De même, MutY excise efficacement les adénines opposées aux FapyG (Wiederholt *et al.*, 2003). Une nouvelle fois, on peut déduire de ceci que les souches KO pour *mutY* ont une augmentation de transversions G/C vers T/A aussi à cause du FapyG, et non pas uniquement à cause de la 8-oxo-G.

Par contre, l'activité de MutT sur les Fapy-dGTP a été montrée *in vitro* comme étant quasi inexistante (Imoto *et al.*, 2006), cependant la présence de Fapy-dGTP *in vivo* n'a pas encore été démontrée. Si la présence de Fapy-dGTP était avérée, alors il sera important de rechercher si les cellules possèdent une enzyme Fapy-dGTP-ase. En effet, si aucune enzyme ayant cette activité n'était trouvée, alors l'insertion de Fapy-dGMP dans l'ADN serait favorisée par rapport à celle du 8-oxo-GMP, pour lequel l'enzyme MutT diminue cette possibilité.

Enfin, les enzymes Nei et Nth ont été montrées comme pouvant exciser les FapyG dans des oligonucléotides courts, notamment en face d'un G ou d'un A, cependant avec une efficacité extrêmement faible (Wiederholt *et al.*, 2005).

Tout ceci indique qu'au moins une partie du système GO permet d'éviter aussi les effets mutagènes du FapyG. Des investigations supplémentaires sont cependant nécessaires pour vérifier s'il n'y a pas d'autres protéines impliquées dans le traitement du FapyG, et si le Fapy-dGTP est réellement présent *in vivo*.

Enfin, le traitement du FapyG par le système GO humain ne sera pas discuté ici.

## **I.7. La protéine bactérienne Fpg**

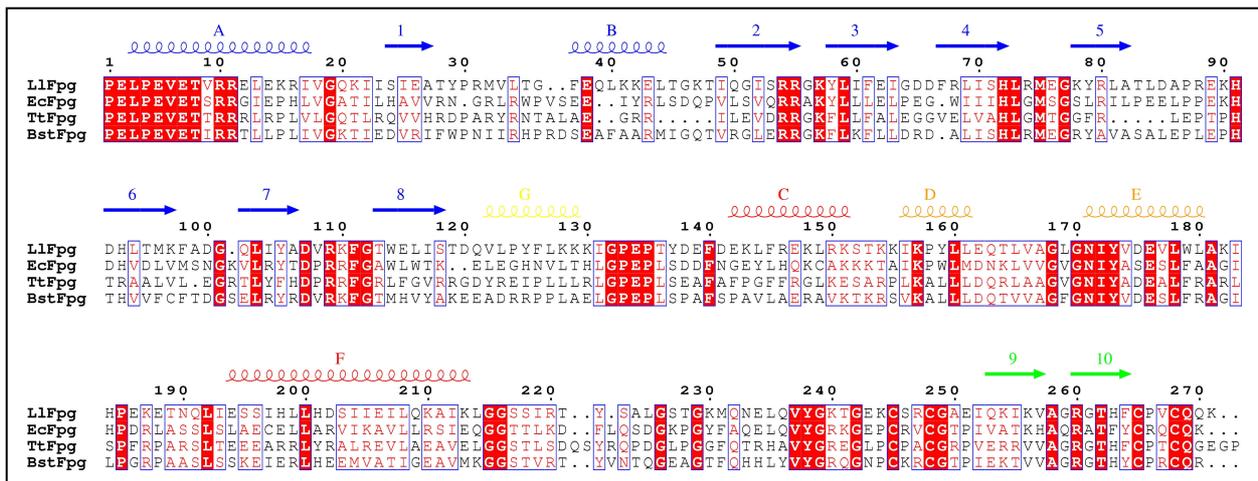
### **I.7.1. Généralités**

La Formamidopyrimidine-ADN glycosylase (Fpg) est une protéine procaryote du système BER de réparation de l'ADN. Cette protéine fut d'abord mise en évidence *via* son activité ADN glycosylase sur les résidus  $N^7$ -méthyle-FapyG de l'ADN (Chetsanga *et al.*, 1979), puis le gène correspondant de *E. coli* a été cloné (Boiteux *et al.*, 1987). L'implication de son inactivation dans l'apparition d'un phénotype mutateur correspondant à des transversions G/C vers T/A (chapitre I.6 sur le système GO) lui a ensuite valu un nom alternatif : MutM (Cabrera *et al.*, 1988).

La protéine Fpg fait environ de 31 kDa, est basique (pI entre 8,6 et 9,6 selon la souche bactérienne) (Boiteux *et al.*, 1987; Boiteux *et al.*, 1990; Castaing *et al.*, 1999), et possède une structure primaire relativement bien conservée (Figure 22) qui est caractérisée par :

- une séquence N-terminale strictement conservée (PELPEVET) et impliquée dans l'activité catalytique (Tchou *et al.*, 1995; Zharkov *et al.*, 1997; Sidorkina *et al.*, 2000).

- un motif hélice deux tours hélice (H2TH), caractéristique de la superfamille structurale Fpg/Nei (chapitre I.5.2 sur les ADN glycosylases), et représenté en orange dans les figures de ce chapitre.
- un motif de type doigt à zinc (C-terminal) non-canonique CX<sub>2</sub>CX<sub>16</sub>CX<sub>2</sub>C, ce motif étant indispensable à la fixation de la protéine sur l'ADN (Castaing *et al.*, 1993 ; Tchou *et al.*, 1993), et représenté en vert dans les figures de ce chapitre.

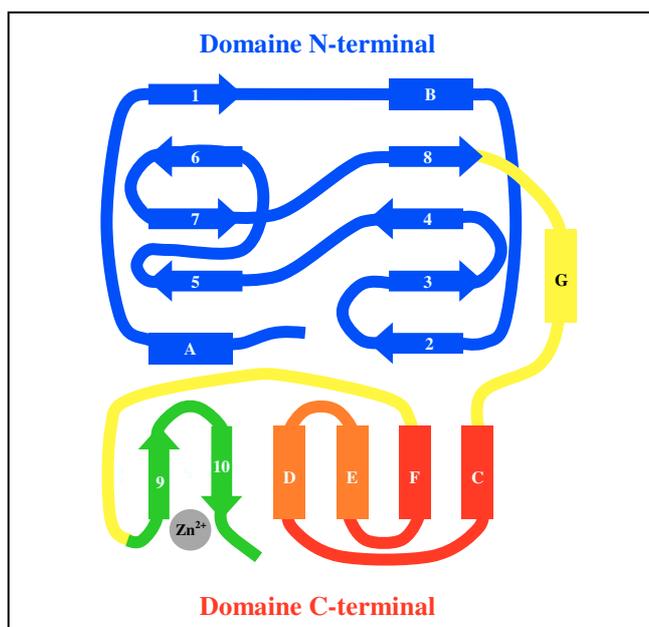


**Figure 22 :** Alignement de séquence des Fpg issues de *Lactococcus lactis*, *Escherichia coli*, *Thermus thermophilus* et *Bacillus stearothermophilus*.

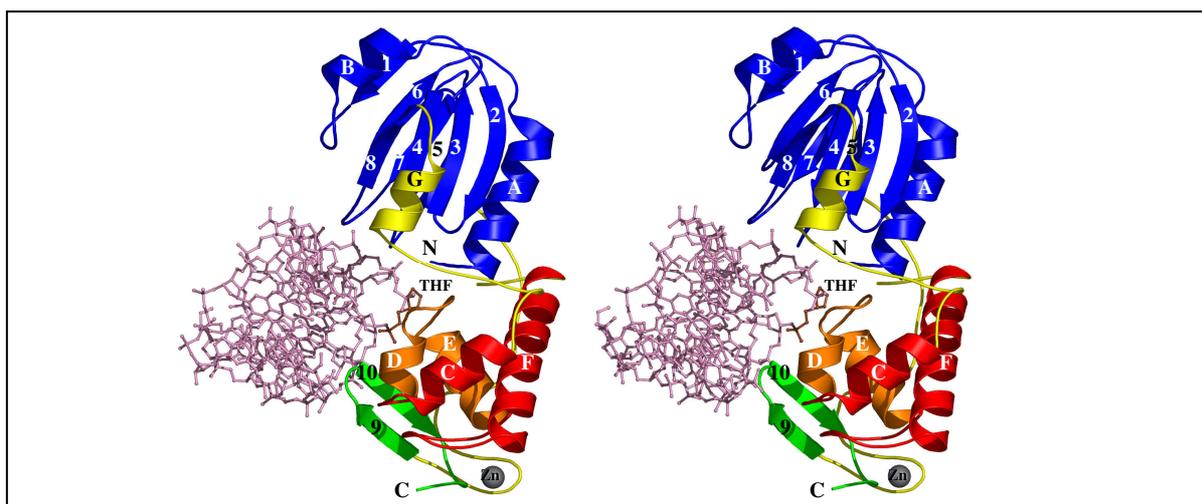
Les structures secondaires de la protéine sont représentées et sont soumises au code couleur défini par Sugahara *et al.*, 2000. Cet alignement a été effectué par CLUSTALW (Thompson *et al.*, 1994) et représenté par ESPript (Gouet *et al.*, 1999). Les résidus strictement conservés sont en blanc sur fond rouge, et les résidus similaires sont en rouge sur fond blanc.

Structuralement, Fpg est constituée de deux domaines globulaires (Figures 23 et 24) :

- Un domaine N-terminal (en bleu) organisé en sandwich  $\beta$  antiparallèle, et constitué de huit feuillets  $\beta$  notés de 1 à 8, ainsi que d'hélices  $\alpha$  notées A, B (et G pour la Fpg de *Lactococcus lactis*, cette hélice n'étant pas référencée dans la structure de *TtFpg* libre de Sugahara et collaborateurs en 2000, même si elle est présente dans cette structure).
- Un domaine C-terminal constitué de quatre hélices  $\alpha$  notées C, D, E et F, et de deux feuillets  $\beta$  notés 9 et 10. Le motif en doigt à zinc (en vert) et le motif hélice-deux tours-hélice (en orange) constituent deux sous-domaines de ce domaine C-terminal.



**Figure 23 :** Schéma de l'organisation des structures secondaires de la Fpg de *Lactococcus lactis*. Les hélices et les feuillets sont respectivement représentés par des flèches et des rectangles, et sont colorés selon la définition de Sugahara *et al.*, 2000.



**Figure 24 :** Vue stéréoscopique de la structure du complexe formé entre Fpg et un ADN. Les structures secondaires sont colorées selon la définition de Sugahara *et al.*, 2000. Le domaine Nter est en bleu, le doigt à zinc est en vert, le motif H2TH est en orange, les hélices C et F sont en rouge, l'atome de zinc est en gris, l'ADN est en rose clair, l'analogue de lésion est en marron, et une hélice supplémentaire par rapport à la définition de Sugahara est en jaune.

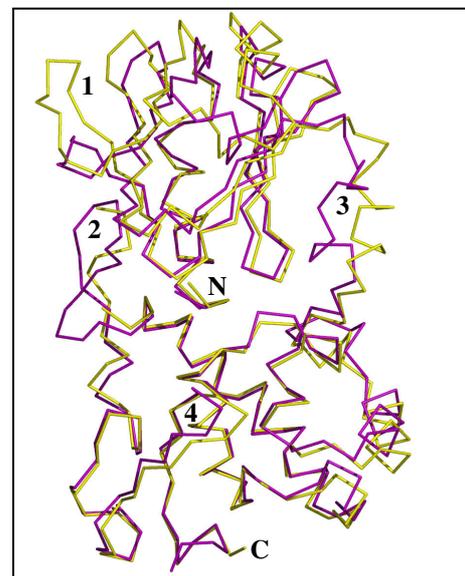
On peut noter que la Figure 24 correspond à la structure cristallographique à 1,66 Å de résolution de la Fpg de *Lactococcus lactis* (LIFpg) double mutante P1G et E2Q complexée à un ADN contenant un site THF (analogue non métabolisable de site abasique). J'ai choisi cette structure car c'est celle dont la résolution est la meilleure parmi toutes celles que nous avons obtenues au laboratoire pour un complexe Fpg/ADN lésé, et car elle possède un repliement global identique aux autres structures de Fpg. Cependant, cette structure étant peu informative

pour la compréhension du mécanisme d'action de Fpg, elle n'a pas été déposée dans la PDB et n'apparaît dans aucune des publications de notre équipe de recherche.

A ce niveau, on peut noter que les différentes Fpg venant de différentes bactéries ont toutes une structure globale très similaire, et que la fixation à l'ADN ne modifie que très peu cette structure. Afin d'illustrer ce point, la figure suivante (Figure 25) présente une superposition des structures de la Fpg de *Thermus thermophilus* (TtFpg) libre (Sugahara *et al.*, 2000) et de la LIFpg complexée à un ADN. Entre ces deux structures, l'écart type (RMSD) calculé sur la position de 230 atomes Ca de la chaîne principale, est seulement de 1,25Å, ce qui suggère que le repliement global de Fpg pour ces deux espèces, liée ou non à l'ADN, est très similaire.

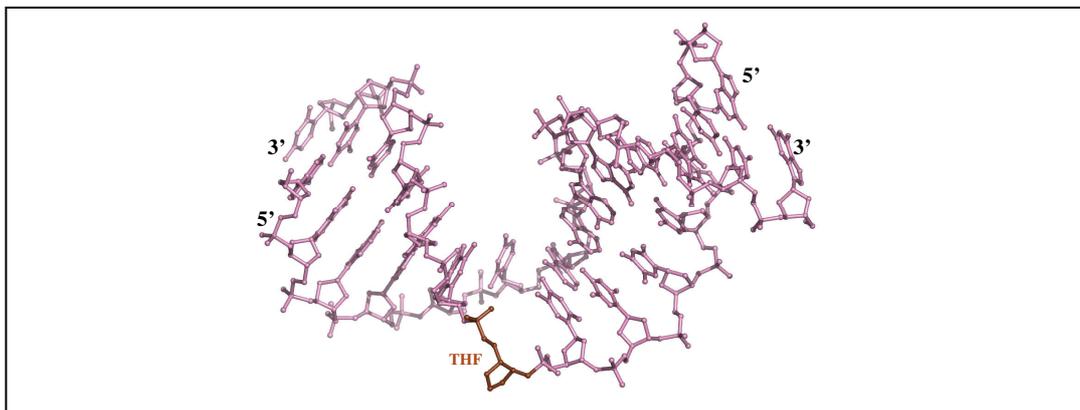
Seules quelques modifications locales de la structure de Fpg lors de la fixation à l'ADN sont visibles (Figure 25) :

- 1- Au niveau de boucles de l'extrémité N-terminale.
- 2- Au niveau de la mobilité d'une la boucle flexible, comprise entre les résidus 219 et 223, qui est absente dans la structure de LIFpg.
- 3- Dans la région variable comprise entre les résidus 119 et 130 (numérotation de LIFpg), cette variation étant probablement due à une différence de séquence plutôt qu'à la fixation à l'ADN.
- 4- Dans la boucle conservée du motif en doigt de zinc (résidus 256 à 261), qui est déplacée de 3 Å vers l'ADN dans la structure de LIFpg.



**Figure 25 :** Superposition du squelette des Ca de la TtFpg libre (en violet, code PDB 1EE8) et de P1GE2Q-LIFpg liée à un ADN contenant un site THF (en jaune). Les numéros correspondent aux informations présentes dans le texte accolé à la Figure.

On peut noter que la faible différence de structure entre la forme libre et complexée de Fpg est en bon accord avec une hypothèse concernant les ADN glycosylases, qui indique que celles-ci sont relativement rigides pour que ce soit l'ADN qui se courbe et pas l'inverse (Serre *et al.*, 2002). En effet, Fpg induit une courbure de l'ADN au niveau de la lésion allant de 60 à 75 degrés selon les structures (Figure 26).



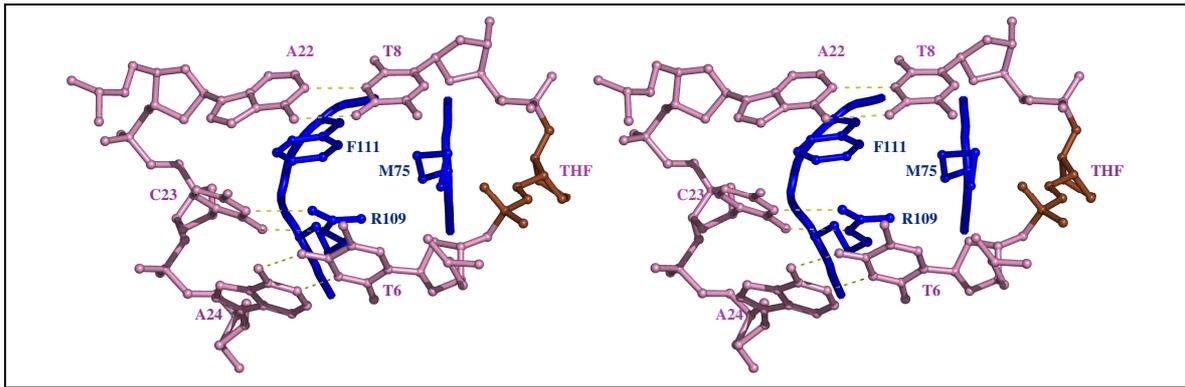
**Figure 26 : Structure de l'ADN contenant un site THF complexé à P1GE2Q-L/Fpg.**  
L'ADN est en rose clair, et la lésion est en marron.

On peut préciser qu'aucune distorsion de ce type n'a été observée pour les structures d'ADN libre contenant un site THF obtenues par RMN (Kalnik *et al.*, 1989 ; Cuniasse *et al.*, 1990 ; Goljer *et al.*, 1995), ce qui confirme que la courbure observée résulte des contacts établis avec Fpg.

En fait, la courbure de l'ADN est une caractéristique commune aux ADN glycosylases, ce qui facilite l'extrusion de la lésion hors de la double hélice d'ADN et permet sa stabilisation dans le site actif de l'enzyme (chapitre I.5.2 sur les ADN glycosylases). Ceci permet de rendre le C1' du désoxyribose de la base endommagée accessible à l'attaque nucléophile par les ADN glycosylases.

Dans le cas de Fpg, la conformation courbée de l'ADN avec la lésion extrudée est stabilisée par plusieurs facteurs :

- Un réseau de liaisons hydrogène est présent entre la protéine (relativement rigide) et l'ADN (flexible), ce qui pousse l'ADN à se courber. Ce réseau sera détaillé plus loin dans ce chapitre.
- Trois résidus strictement conservés parmi les Fpg sont insérés dans la double hélice d'ADN dans l'espace laissé vacant par l'extrusion de la lésion : la triade d'intercalation Met75, Arg109 et Phe111 selon la numérotation de L/Fpg, qui empêchent le retour de la base en position intrahélicale et préviennent l'effondrement de la structure de l'ADN au niveau de l'espace créé par le départ de la lésion (Figure 27).

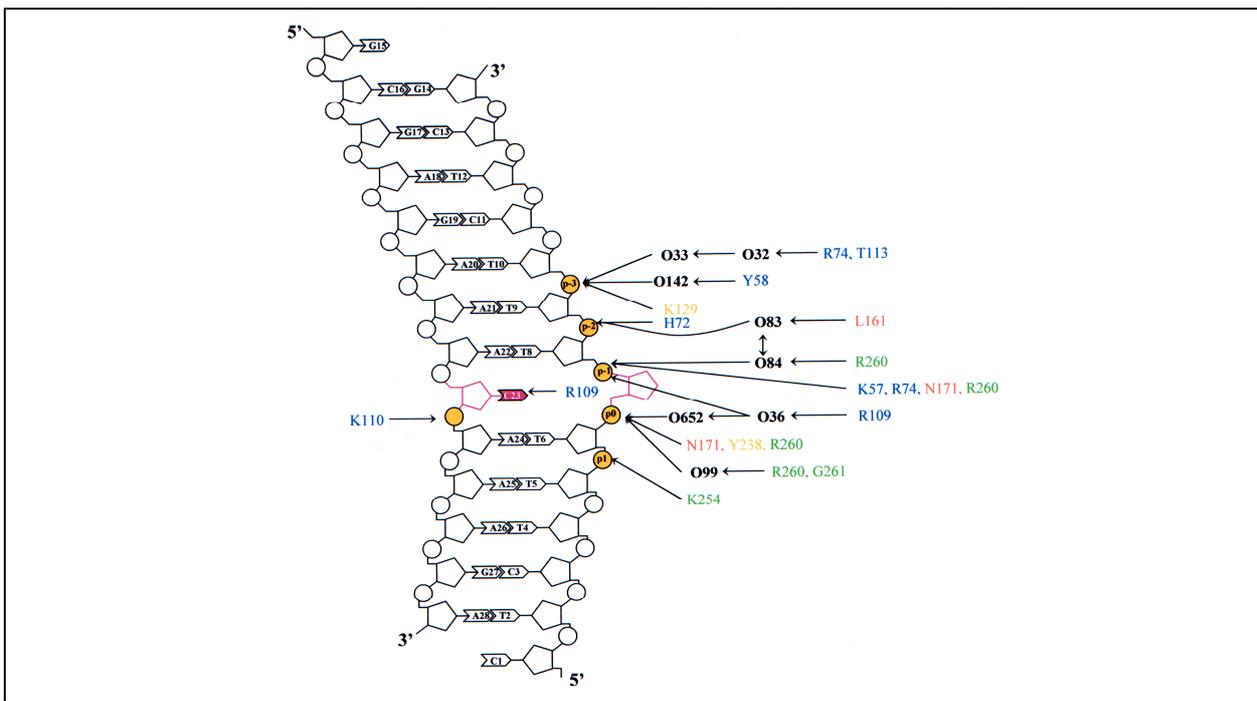


**Figure 27 :** Vue stéréoscopique de la triade d'intercalation M75, R109, F111 de LIFpg dans l'ADN. L'ADN est en rose clair, la lésion en marron, la triade d'intercalation de Fpg en bleu, et les liaisons hydrogène en jaune.

De plus, on peut noter que l'Arg109 de la triade d'intercalation permet de maintenir une conformation double-brin de l'ADN en faisant des liaisons hydrogène avec la face Watson-Crick de la base opposée à la lésion (C23, Figure 27), participant aussi à la spécificité de substrat de l'enzyme (détaillée plus loin dans ce chapitre).

D'autre part, le réseau de liaisons hydrogène établi entre Fpg et l'ADN fait majoritairement intervenir des résidus conservés de la protéine, et est surtout impliqué dans :

- La stabilisation des phosphates du brin contenant la lésion (Figure 28) (Castaing *et al.*, 1999) *via* des résidus des deux domaines de Fpg, notamment des résidus du doigt à zinc et du motif H2TH, mais aussi des résidus de boucles de l'extrémité N-terminale de la protéine, ainsi que des interactions médiées par des molécules d'eau (un exemple de ce type de réseau est présenté dans la Figure 28). Ces interactions sont relativement similaires dans toutes les structures des Fpg issues de différentes espèces bactériennes complexées à l'ADN (ces structures seront citées plus loin dans ce chapitre), et sont quasi-identiques pour les structures obtenues avec une même espèce bactérienne.
- La stabilisation de la lésion dans le site actif de l'enzyme par contact avec les groupements donneurs/accepteurs de liaisons hydrogène de celle-ci. Ce réseau d'interactions est relativement différent selon le type de lésion, expliquant la gamme de substrats extrêmement large réparée par Fpg (chapitre I.7.2 sur les spécificités de substrat de Fpg).



**Figure 28 :** Représentation schématique du réseau de liaisons hydrogène entre PIGE2Q-L/Fpg et l'ADN contenant un site THF.

Le code couleur des résidus protéiques est le même que pour la Figure 4. Les molécules d'eau sont représentées en noir avec la norme O+(numéro de la molécule dans le fichier PDB). Les liaisons hydrogène sont représentées par des flèches noires. Le site THF et la cytosine qui lui est opposée sont colorés en rose. Les phosphates du squelette de l'ADN en contact avec l'enzyme sont colorés en jaune.

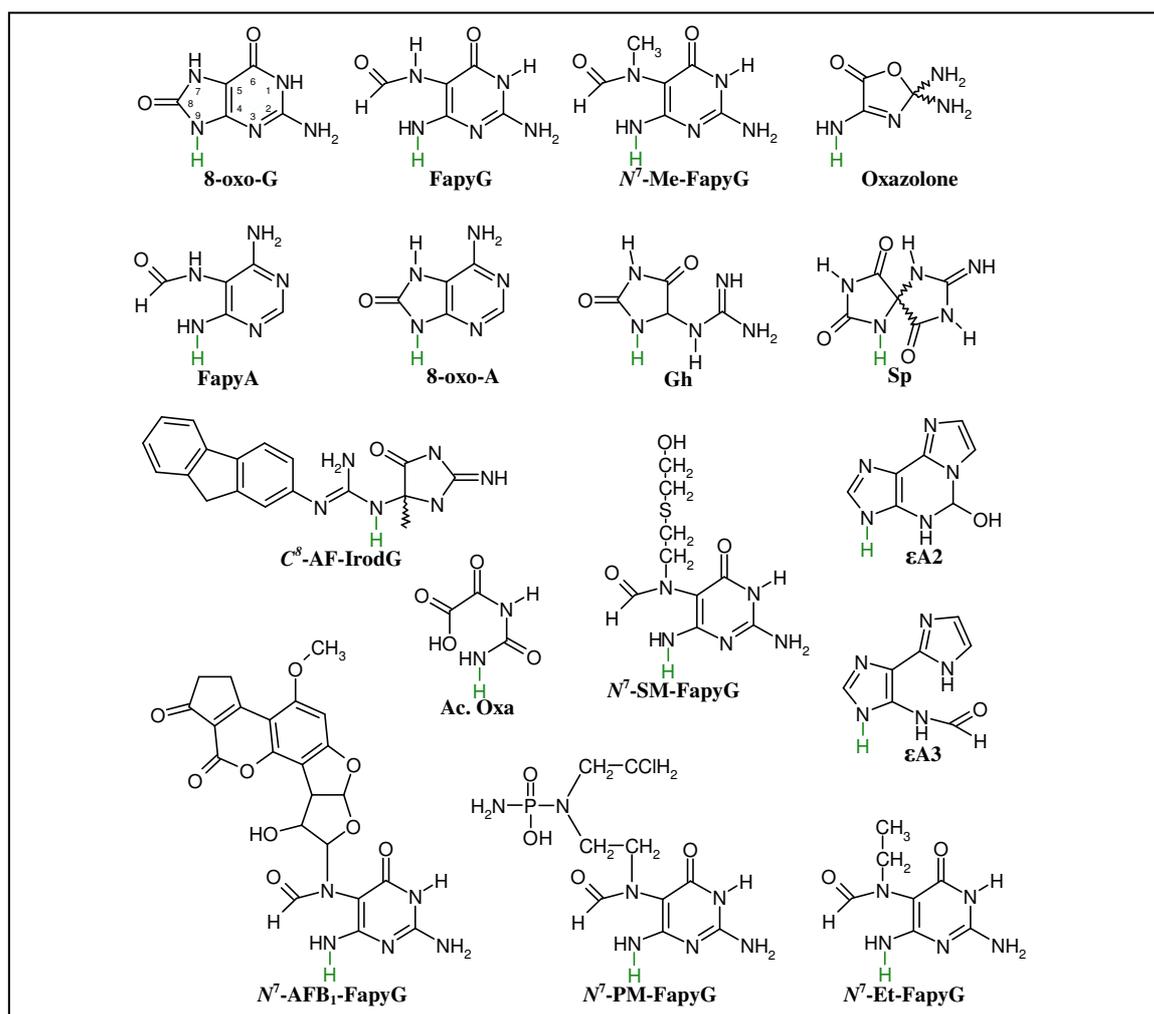
## 1.7.2. Spécificités de substrat de Fpg

### - Spécificité de substrat concernant les lésions réparées par Fpg

L'activité ADN glycosylase de Fpg a été initialement décrite comme étant spécifique de l'élimination des purines oxydées dans l'ADN double brin (Figure 29), mais cette activité a ensuite été démontrée sur des lésions pyrimidiques (Figure 30) (Chetsanga *et al.*, 1982; Chetsanga *et al.*, 1983; Tchou *et al.*, 1991; Boiteux *et al.*, 1992; Castaing *et al.*, 1992; Castaing *et al.*, 1993; Hatahet *et al.*, 1994; Tchou *et al.*, 1994; Karakaya *et al.*, 1997; Li *et al.*, 1997; Jurado *et al.*, 1998; Purmal *et al.*, 1998; Tudek *et al.*, 1998; Bourdat *et al.*, 1999; Castaing *et al.*, 1999; D'Ham *et al.*, 1999; Asagoshi *et al.*, 2000; Duarte *et al.*, 2000; Gasparutto *et al.*, 2000; Leipold *et al.*, 2000; Zhang *et al.*, 2000; Duarte *et al.*, 2001; Speina *et al.*, 2001; Fromme *et al.*, 2003a; Tudek, 2003; Wiederholt *et al.*, 2003; Alekseyev *et al.*, 2004; Simon *et al.*, 2006).

Si on considère à nouveau la spécificité de substrat de Fpg, on voit que bien que les substrats sur lesquels l'enzyme soit la plus active sont la 8-oxo-G et le FapyG, Fpg est en fait capable d'éliminer un grand nombre de bases nucléiques puriques et pyrimidiques oxydées dans l'ADN double brin.

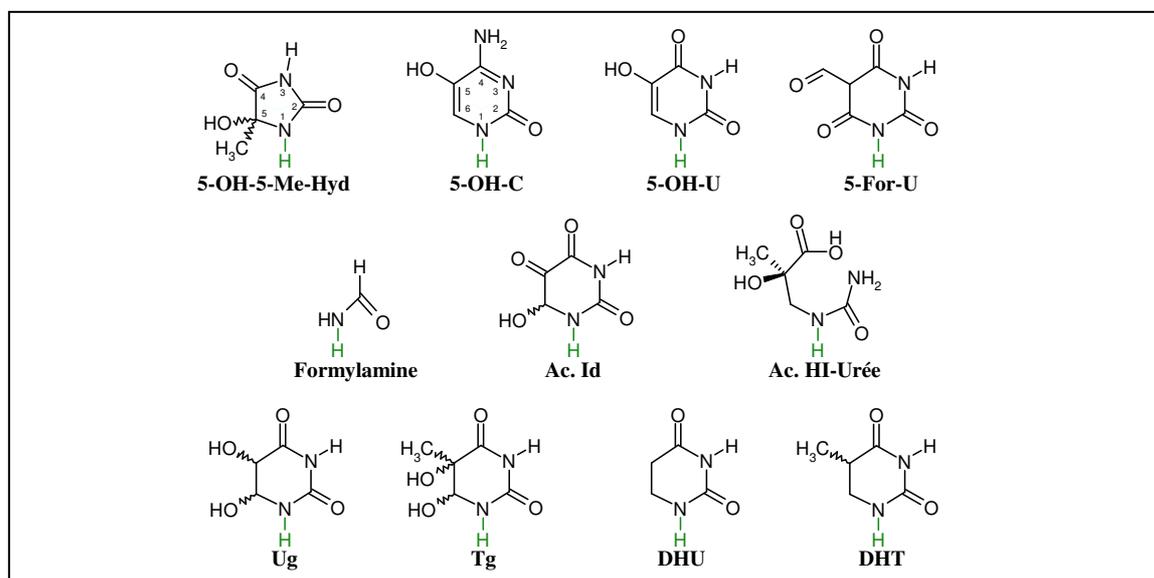
Cependant, la liste de substrats présentée dans les Figures 29 et 30 n'est pas forcément exhaustive, car elle dépend des lésions sur lesquelles l'activité de Fpg a été testée. De plus, l'activité de Fpg sur la plupart de ces lésions n'a pas été déterminée sous forme d'études comparatives (par rapport à l'activité sur l'un de ses substrats de référence, la 8-oxo-G par exemple), ce qui fait que l'on ne sait pas vraiment si certaines lésions sont de bons ou de mauvais substrats pour Fpg. En effet, les conditions de réaction lors des essais d'activité (tampon, température, salinité, etc) sont susceptibles de faire nettement varier les valeurs mesurées, ce qui rend impossible la comparaison entre les études réalisées.



**Figure 29 : Substrats de l'activité ADN glycosylase de Fpg dérivés de purines.**

La notation peu explicite de certaines lésions correspond à : Acide oxalorique (Ac. Oxa), FapyG avec adduit méthyle en  $N^7$  ( $N^7$ -Me-FapyG), Guanidinohydantoïne (Gh), Spiroiminodihydantoïne (Sp), Guanine à cycle imidazole ouvert avec adduit d'aminofluorène en  $C^8$  ( $C^8$ -AF-IrodG), FapyG avec adduit de gaz moutarde phosphoramidate en  $N^7$  ( $N^7$ -PM-FapyG), FapyG avec adduit de gaz moutarde sulfuré en  $N^7$  ( $N^7$ -SM-FapyG), FapyG avec adduit d'aflatoxine B<sub>1</sub> en  $N^7$  ( $N^7$ -AFB<sub>1</sub>-FapyG), FapyG avec adduit éthyle en  $N^7$  ( $N^7$ -Et-FapyG), dérivés d'oxydation de l'éthénoadénine à cycle fermé ( $\epsilon$ A2) ou ouvert ( $\epsilon$ A3). La position d'attachement au désoxyribose dans l'ADN est indiquée en vert.

De plus, la plupart des études qui ont conduit à la définition des substrats de Fpg ont été réalisées sur de courts oligonucléotides avec une lésion unique. Or, dans le cas où des ADN longs ont été utilisés comme modèle d'étude, la spécificité de substrat de Fpg est apparue comme plus restreinte que ce que l'on trouve dans les Figures 29 et 30 (Boiteux *et al.*, 1992; Karakaya *et al.*, 1997). Cependant, dans les études avec des ADN longs, les lésions ne sont généralement pas introduites à des positions spécifiques. En effet, dans ce cas les méthodes d'introduction de lésions (UV, rayons  $\gamma$ ...) impliquent que celle-ci soient créées à des positions aléatoires, avec des quantités différentes de chaque type de lésion (Boiteux *et al.*, 1992). Ceci, allié à la sensibilité limitée des systèmes de détection de l'activité de Fpg sur des ADN longs (généralement, des techniques de chromatographies couplées à de la spectrométrie de masse ou à une détection électrochimique), implique que ce type d'étude puisse définir certaines lésions comme non-substrat de Fpg tout simplement à cause de problèmes de sensibilité de détection. De façon générale, on peut cependant dire que l'activité de Fpg semble plus restreinte sur les ADN longs que sur de courts oligonucléotides, probablement à cause du temps de translocation supplémentaire que l'enzyme monopolise pour trouver les lésions sur de tels ADNs (Berg *et al.*, 1981 ; Winter *et al.*, 1981a et b).



**Figure 30 : Substrats de l'activité ADN glycosylase de Fpg dérivés de pyrimidines.**

La notation peu explicite de certaines lésions correspond à : 5-hydroxy-5-méthyle-hydantoïne (5-OH-5-Me-Hyd), 5-hydroxy-cytosine (5-OH-C), 5-hydroxy-uracile (5-OH-U), 5-formyle-uracile (5-For-U), Acide isodialurique (Ac. Id),  $N^3$ -(2-acide hydroxyisobutyrique)urée (Ac. HI-Urée), Uracile glycol (Ug), Thymine glycol (Tg), Dihydrouracile (DHU), Dihydrothymine (DHT). La position d'attachement au désoxyribose dans l'ADN est indiquée en vert.

Finalement, on peut tout de même indiquer que près d'une trentaine de lésions puriques et pyrimidiques ont été décrites comme pouvant être excisées par Fpg sur des oligonucléotides

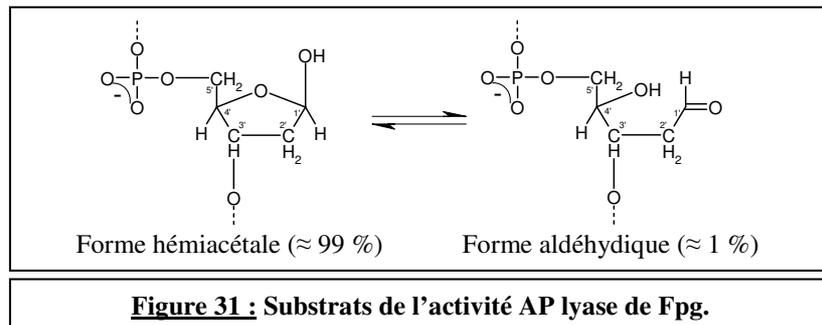
courts, ce qui fait que cette enzyme est probablement l'ADN glycosylase bactérienne dont la spécificité de substrat est la plus large.

Pour expliquer cette large spécificité de lésions réparées par l'activité ADN glycosylase de Fpg, des informations précieuses ont été apportées par la résolution de structures radiocristallographiques, notamment sur des complexes entre Fpg et des ADNs contenant une lésion. En effet, ces études ont montré que la poche de fixation de substrat de Fpg est exceptionnellement large en comparaison avec celle d'autres ADN glycosylases (Fromme *et al.*, 2003a; Zharkov *et al.*, 2003; Coste *et al.*, 2004), ce qui lui permet d'accommoder de nombreuses lésions de tailles différentes. De plus, dans cette poche, un réseau de liaisons hydrogène est créé avec la lésion, permettant sa stabilisation dans une conformation correcte pour permettre l'activité ADN glycosylase. Cependant, ce réseau de liaisons hydrogène est apparu comme relativement plastique dans les études structurales, avec des différences significatives pour les différentes lésions étudiées, en l'occurrence la 8-oxo-G, le FapyG et la DHU (Fromme *et al.*, 2003a; Coste *et al.*, 2004). En fait, la plasticité de ce réseau contribue à la large gamme de lésions excisées par Fpg, mais implique aussi que les lésions soient dans une conformation plus ou moins stable et propice pour l'activité de Fpg, d'où une différence d'activité suivant les lésions.

Outre ces informations issues de structures, d'autres paramètres sont susceptibles d'expliquer la différence d'activité ADN glycosylase de Fpg pour différentes lésions, notamment au niveau de la reconnaissance précoce des lésions *via* la stabilité des paires de bases (Banerjee *et al.*, 2006) ou la reconnaissance de groupements chimiques dans les sillons de l'ADN (Tchou *et al.*, 1994; Lipscomb *et al.*, 1995). Par exemple, il est désormais bien accepté que la stabilité des paires de bases dans l'ADN influence la capacité des ADN glycosylases à pouvoir extruder ces bases dans la poche de leur site actif. Dans ce cadre, plus l'appariement formé est instable, plus l'extrusion sera facilitée, et plus grande sera la probabilité de former un complexe catalytiquement actif avec l'enzyme, d'où une activité plus importante de l'enzyme sur ce substrat. Ainsi, on sait que les paires de bases formées avec une lésion sont généralement plus instables que celles formées entre les bases nucléiques normales correspondantes (Haraguchi *et al.*, 2002; Tudek *et al.*, 2003), ce qui explique que les ADN glycosylases (dont Fpg) soient très actives sur les paires de bases formées avec les lésions, et très peu actives sur les paires de bases canoniques.

Un autre volet de la spécificité de substrat de Fpg concernant le type de lésion qu'elle répare est que, après avoir réalisé son activité ADN glycosylase, celle-ci possède une activité

AP lyase capable de cliver le squelette phosphodiester de l'ADN en 3' et en 5' du site abasique par  $\beta,\delta$ -élimination, générant des extrémités 5' et 3' phosphates (chapitre I.7.3 sur le mécanisme d'action de Fpg et les structures correspondantes). On peut d'ailleurs préciser que l'on a des indices comme quoi la forme clivable du site AP serait probablement la forme aldéhydique (Castaing *et al.*, 1999). Cette seconde activité fait que Fpg appartient à la famille fonctionnelle des ADN glycosylases/AP lyases bi-fonctionnelles (Bailly *et al.*, 1989 ; O'connor *et al.*, 1989).



### - Spécificité de substrat concernant la base opposée à la lésion

Outre la spécificité de substrat concernant les lésions excisées par Fpg, on sait que cette enzyme possède une seconde spécificité concernant la base opposée à la lésion, que ce soit pour l'activité ADN glycosylase ou l'activité AP lyase (Tableau 7). De façon générale, Fpg est spécifique de l'élimination des lésions appariées aux bases nucléiques autres que l'adénine, ce qui permet à l'enzyme de lutter contre les effets mutagènes de ses substrats principaux, en l'occurrence la 8-oxo-G et le FapyG (chapitre I.6 sur le système GO). Cependant, bien que le principe du « tout sauf un A » soit respecté pour toutes les lésions pour lesquelles la préférence de base opposée à la lésion a été étudiée, l'ordre de préférence par rapport aux trois autres bases (T, G, et C) peut légèrement varier selon la lésion considérée, et peut même être différente pour une même lésion dans des études séparées (Tableau 7, les noms des lésions étant les mêmes que dans la légende des Figures 29 et 30).

De plus, les différences d'activité en fonction de la base en face de la lésion n'ont généralement pas la même amplitude selon la lésion considérée. Par exemple, dans la publication de Asagoshi collaborateurs en 2000, on voit que la 8-oxo-G est environ 10 fois mieux excisée en face d'un C qu'en face d'un A, alors que le  $N^7$ -Me-FapyG est seulement environ 1,7 fois mieux excisé en face d'un C qu'en face d'un A.

**Tableau 7 : Différences d'activité de *EcFpg* en fonction de la base opposée à la lésion pour différentes lésions.**

Lésion	Ordre de préférence pour la base opposée à la lésion	Publication(s) de référence
8-oxo-G	$dT \geq dG \geq dC \gg dA$	Tchou <i>et al.</i> , 1991; Tchou <i>et al.</i> , 1994; Asagoshi <i>et al.</i> , 2000
	$dG \geq dT > dC \gg dA$	Zaika <i>et al.</i> , 2004
	$dC \geq dT > dG \gg dA$	Castaing <i>et al.</i> , 1993
<i>N'</i> -Me-FapyG	$dC > dT \geq dG > dA$	Asagoshi <i>et al.</i> , 2000
FapyG	$dC \gg dA$	Wiederholt <i>et al.</i> , 2003
Gh	$dC \approx dT \approx dG \gg dA$	Leipold <i>et al.</i> , 2000
Sp	$dC \approx dT \approx dG \gg dA$	Leipold <i>et al.</i> , 2000
FapyA	$dG > dT > dA$	Wiederholt <i>et al.</i> , 2002a
$\epsilon A2$ et $\epsilon A3$	$dT > dC > dG \gg dA$	Speina <i>et al.</i> , 2001
5-OH-U	$dG \gg dA$	Hatahet <i>et al.</i> , 1994
Xanthine	Excision avec dC et pas avec dG, dT et dA	Wuenschell <i>et al.</i> , 2003
Ac.HI-urée	$dC \geq dG \geq dT > dA$	Jurado <i>et al.</i> , 1998
Site AP	$dC \gg dA$	Tchou <i>et al.</i> , 1994
DHU	Equivalente pour les quatre bases	Zaika <i>et al.</i> , 2004
5-OH-C	$dG \approx dA$	Hatahet <i>et al.</i> , 1994
Ug	$dG \geq dA$	Purmal <i>et al.</i> , 1998

A ce stade, on peut préciser que les différences d'activité en fonction de la base opposée à la lésion sont liées, au moins en partie, à une différence d'efficacité de fixation au substrat, cette fixation étant plus efficace avec une lésion appariée à une cytosine qu'avec cette même lésion appariée à une adénine (Castaing *et al.*, 1993; Tchou *et al.*, 1994; Wiederholt *et al.*, 2003).

Les données structurales sur les complexes formés entre Fpg et l'ADN ont toutes pointé le résidu R109 (numérotation de *LI*Fpg) de la triade d'intercalation comme étant le seul élément structural pouvant expliquer la discrimination quant à la base opposée à la lésion. En effet, il a été démontré dans ces structures que ce résidu était le seul qui faisait des interactions directes avec la base opposée à la lésion *via* des liaisons hydrogène. Cependant, aucune de ces structures n'a présenté de complexe avec une adénine en face de la lésion, ce qui nous empêche de savoir si cette discrimination vient bel et bien majoritairement d'une stabilisation moins efficace de l'adénine par rapport aux autres bases possibles en face de la lésion.

Un autre paramètre pouvant expliquer partiellement cette préférence de bases opposée à la lésion correspond encore une fois à la stabilité des paires de bases dans l'ADN. En effet, on sait que les paires de bases formées entre la 8-oxo-G ou le FapyG avec une guanine ou une

thymine dans l'ADN sont plus instables que les paires formées avec une adénine. Ceci participe à une meilleure extrusion de ces lésions quand elles sont opposées à une guanine ou une thymine que quand elles sont opposées à une adénine, d'où une meilleure excision (Haraguchi *et al.*, 2002). Ce paramètre n'explique cependant pas complètement cette spécificité, car les appariements formés entre la 8-oxo-G ou le FapyG et l'adénine sont moins bien réparés que ceux formés avec une cytosine, alors que les paires de base formées avec la cytosine sont les plus stables (Haraguchi *et al.*, 2002).

Enfin, le dernier paramètre que l'on retrouve souvent pour expliquer la discrimination quant à la base opposée à la lésion correspond à une reconnaissance précoce des déterminants des lésions dans l'ADN avant extrusion. L'exemple le plus étudié de ce type de phénomène est celui de la 8-oxo-G, qui est dans une conformation *anti* dans la paire 8-oxo-G/C et en *syn* dans la paire 8-oxo-G/A, donc exposant respectivement son groupement 8-oxo (et N7-H) dans le grand sillon et le petit sillon de l'ADN (chapitre I.5.4 sur les ADN polymérases et leur implication dans la mutagénèse induite par les lésions). Sachant cela, une hypothèse a été émise disant que la meilleure réparation de la 8-oxo-G opposée à une cytosine viendrait de la reconnaissance du groupement 8-oxo (et/ou N7-H) dans le grand sillon *via* un résidu de l'enzyme, reconnaissance qui serait donc impossible dans le petit sillon pour la paire 8-oxo-G/A (Lipscomb *et al.*, 1995). Cependant, ce complexe de reconnaissance précoce n'étant pas possible à piéger pour en réaliser la structure, il n'y a à l'heure actuelle que des hypothèses concernant le ou les résidus de Fpg qui seraient impliqués dans cette reconnaissance. En l'occurrence, la seule partie de Fpg qui soit intercalée dans le grand sillon de l'ADN dans les structures correspond à la boucle du doigt à zinc (Zharkov *et al.*, 2003), dont le résidu conservé Arg260 (numérotation de *Ll*Fpg, équivalent à l'Arg258 selon la numérotation de *Ec*Fpg) semble être le meilleur candidat pour cette reconnaissance.

### **- Autres spécificités de substrat de Fpg**

Outre les spécificités concernant le type de lésion et la base opposée à celle-ci, une troisième spécificité de substrat est observable chez Fpg en fonction de la séquence dans laquelle est intégrée une lésion (Graves *et al.*, 1992b; Asagoshi *et al.*, 2000), avec une préférence pour les régions riches en guanine, indiquant une possibilité d'avoir des zones d'ADN plus efficacement réparées par Fpg que d'autres.

Enfin, on a des indices comme quoi il y aurait aussi une différence de spécificité de substrat pour les Fpg provenant de différentes espèces bactériennes. Ainsi, on sait par exemple

que les Fpg de *E. coli* et de *L. lactis* ne se fixent pas avec le même ordre de préférence sur différents analogues de sites abasiques (Castaing *et al.*, 1999), ce qui suggère qu'une différence de spécificité de substrat entre ces deux formes de Fpg est possible.

### **I.7.3. Le mécanisme d'action de Fpg et les structures correspondantes**

Les nombreuses études structurales et fonctionnelles sur les ADN glycosylases, et en particulier sur Fpg, ont apporté des informations sur le mécanisme d'action de Fpg.

Tout d'abord, Fpg doit rencontrer les lésions qu'elle doit prendre en charge parmi un vaste excès de bases normales. Pour ce faire, deux mécanismes sont envisagés :

- soit l'enzyme se fixe sur l'ADN, puis glisse le long de celui-ci jusqu'à une lésion.
- soit l'enzyme fait des « sauts », se fixant et se dissociant sans cesse de l'ADN.

A ce niveau, les informations issues de la biochimie tendraient plutôt à dire que ces deux mécanismes coexistent, avec une préférence pour le « glissement » par rapport aux « sauts » (Francis *et al.*, 2003). De plus, les structures récemment résolues de la Fpg de *Bacillus stearothermophilus* (*BstFpg*) mutante E2Q interrogeant des ADNs sans lésion (codes PDB 2F5P, 2F5N et 2F5O : Banerjee *et al.*, 2006b) tendent à dire que les interactions de l'enzyme avec l'ADN normal correspondent plus majoritairement à des liaisons hydrogène médiées par des molécules d'eau que dans les complexes post-reconnaissance des lésions. Ces molécules d'eau joueraient le rôle de « lubrifiant », ce qui tendrait à confirmer l'hypothèse du « glissement » (Banerjee *et al.*, 2006b).

Ensuite, l'enzyme doit reconnaître les lésions par rapport aux bases normales. Les structures précédentes tendent à dire que les lésions seraient reconnues alors qu'elles sont encore intrahélicales, et que seules les lésions reconnues seraient extrudées de la double hélice d'ADN, et surtout pas les bases normales impliquées dans des appariements canoniques (Banerjee *et al.*, 2006b). Cependant, deux mécanismes possibles pour la reconnaissance ont été évoqués (Zharkov *et al.*, 2003) :

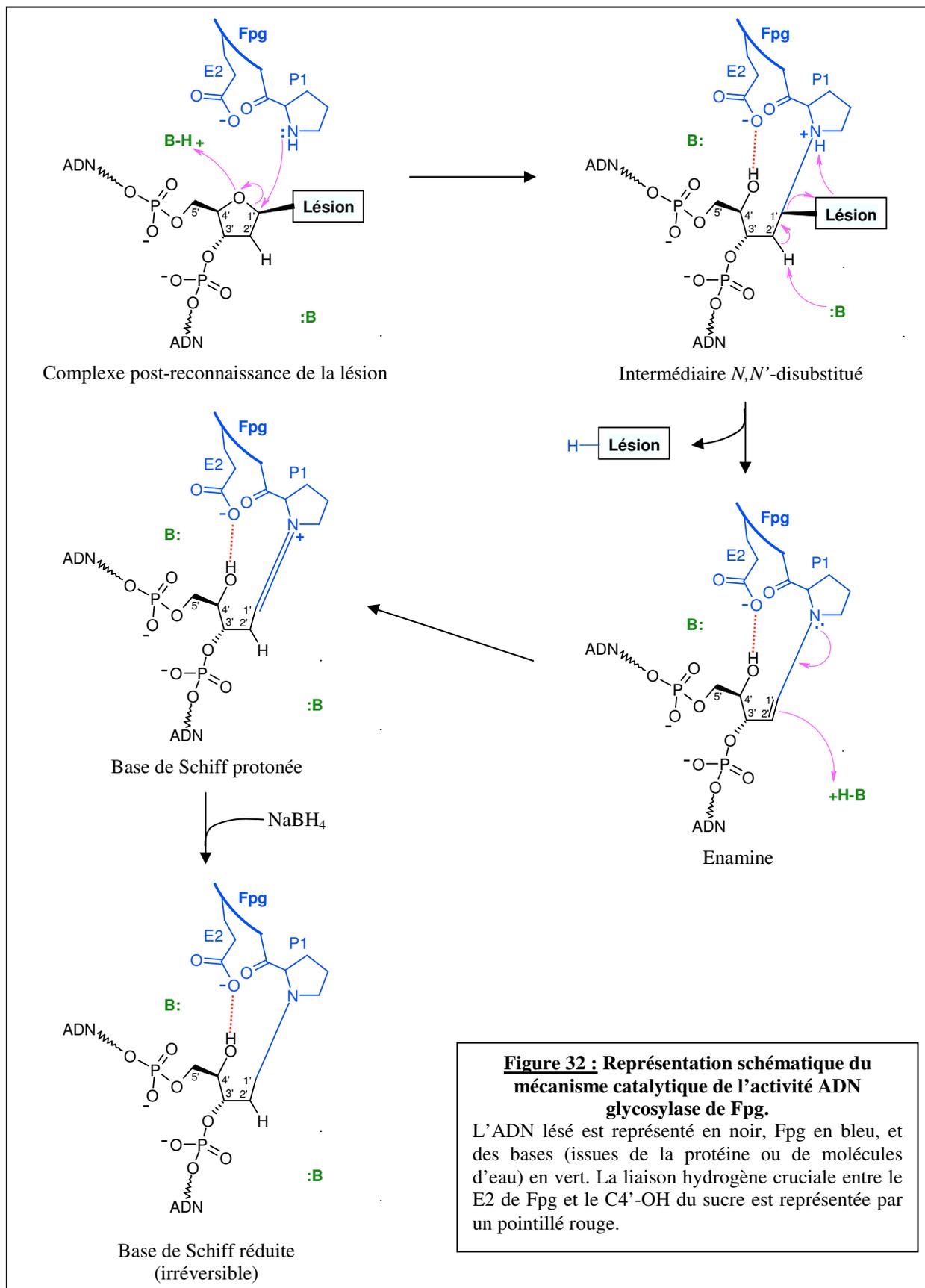
- la lecture directe, où les lésions sont reconnues par Fpg *via* leur profil de donneurs/accepteurs de protons (hypothèse de l'Arg 260 : Zharkov *et al.*, 2003).
- la lecture indirecte, où les lésions sont reconnues *via* les propriétés qu'elles apportent, en l'occurrence la déstabilisation des paires de bases qu'elles forment dans l'ADN (hypothèse de la détection *via* la courbure induite par le résidu Phe113 de la triade d'intercalation chez *BstFpg*, soit la Phe111 chez *LlFpg* : Banerjee *et al.*, 2006b).

Les informations qui ont déjà été évoquées dans le sous-chapitre précédent concernant la spécificité de substrat de Fpg tendent à indiquer que ces deux mécanismes coexistent, aucun d'entre eux ne permettant d'expliquer à lui seul la spécificité de substrat de Fpg quant à la base opposée à la lésion.

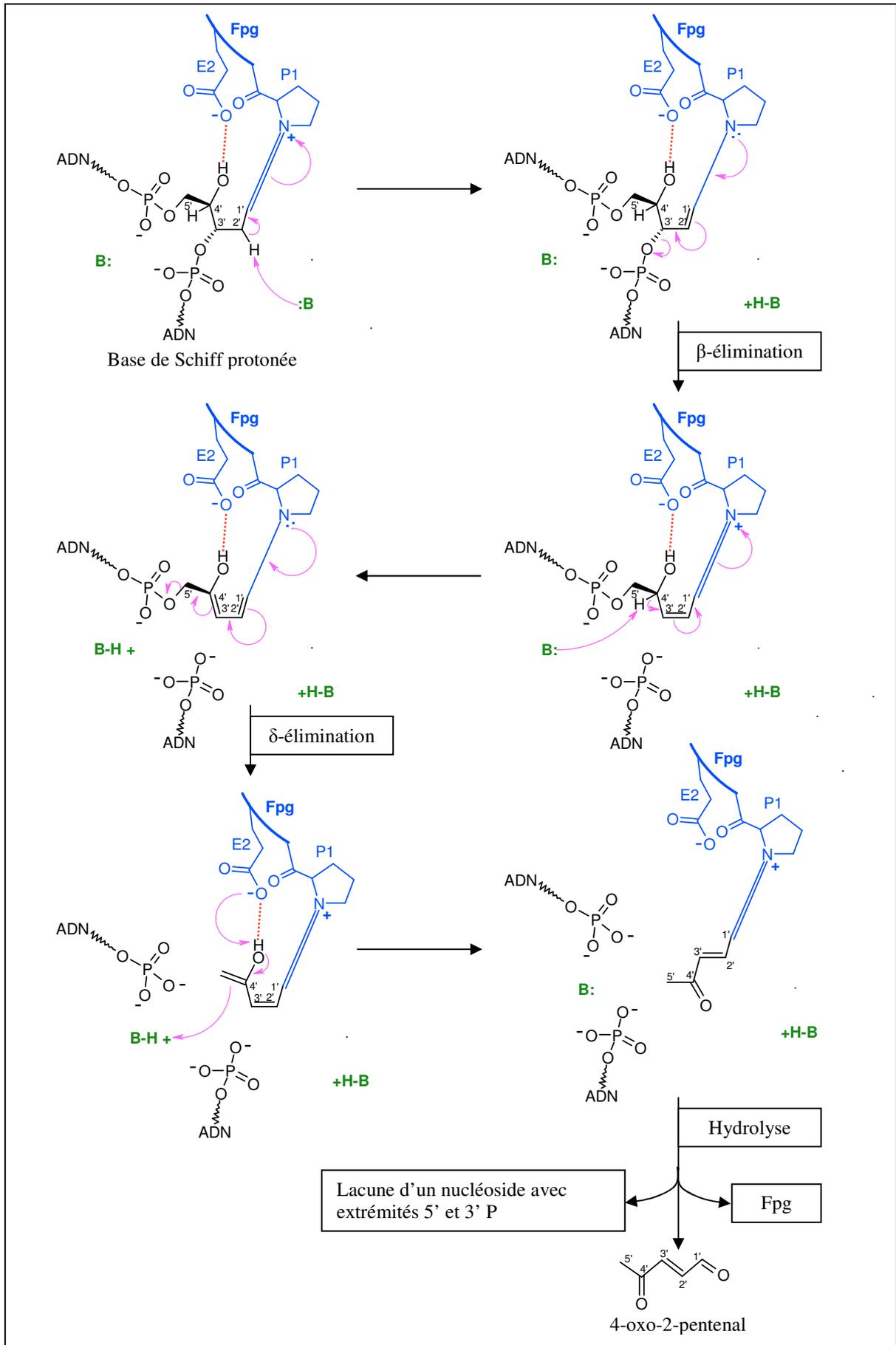
Après avoir reconnu sa cible, Fpg extrude celle-ci et la stabilise dans la poche de son site actif. Comme évoqué dans le sous-chapitre précédent, Fpg stabilise la lésion *via* des liaisons hydrogène avec des résidus tapissant la poche du site actif, et reconnaît la base opposée à la lésion. On possède plusieurs structures de tels complexes, notamment avec *Bst*Fpg mutante E2Q liée à une 8-oxo-G ou une DHU (codes PDB 1R2Y et 1R2Z, respectivement : Fromme *et al.*, 2003a), et avec *Lf*Fpg de type sauvage ou mutante  $\Delta$ P1 liée à un FapyG greffé sur un cyclopentane (codes PDB 1XC8 et 1TDZ, respectivement : Coste *et al.*, 2004). Le processus de stabilisation de la base lésée dans la poche du site actif de Fpg permet vraisemblablement une évaluation des substrats à exciser, notamment au niveau du positionnement optimal du C1' de la lésion permettant l'activité catalytique de Fpg, ainsi qu'au niveau de la discrimination de la base opposée à la lésion. C'est d'ailleurs ce qu'ont révélé les structures précédentes si on considère les différences pour le réseau de liaisons hydrogène formé avec les différentes lésions.

Fpg réalise alors son activité catalytique, qui se déroule en plusieurs étapes. Tout d'abord, l'activité ADN glycosylase se déroule comme suit (Figure 32) : L'enzyme forme un intermédiaire réactionnel covalent entre l'amine secondaire de sa proline N-terminale et le C1' du désoxyribose portant la lésion : l'intermédiaire *N,N'*-disubstitué (Tchou *et al.*, 1995). La formation de cet intermédiaire induit l'ouverture du cycle du désoxyribose, la forme ouverte étant stabilisée par une liaison hydrogène avec le résidu aspartate en position 2 de Fpg (Gilboa *et al.*, 2002). Cependant, on peut noter que cette stabilisation *via* E2 n'est pas en mesure d'expliquer pourquoi la mutation E2Q de l'enzyme induit une perte totale de l'activité ADN glycosylase de Fpg (Lavrukhin *et al.*, 2000), étant donné qu'un résidu glutamine est toujours capable de réaliser cette liaison hydrogène importante. Une base de la protéine arrache alors l'un des protons en 2' du désoxyribose, ce qui conduit au clivage de la liaison *N*-glycosidique entre la lésion et le sucre, puis à la formation d'un second intermédiaire covalent de type base de Schiff protonée entre Fpg et le site AP ouvert (Tchou *et al.*, 1995; Zharkov *et al.*, 1997; Castaing *et al.*, 1999). A ce niveau, l'activité ADN glycosylase est terminée, et l'intermédiaire de type base de Schiff peut être piégé par réduction au borohydrure (ou cyanoborohydrure) de sodium (Tchou *et al.*, 1995; Castaing *et al.*, 1999). Plusieurs structures de tels complexes ont été résolues avec *Ec*Fpg (code PDB 1K82 : Gilboa *et al.*, 2002) et avec *Bst*Fpg (code PDB

1L1Z : Fromme *et al.*, 2002), ce qui a permis de confirmer de façon définitive la présence de cet intermédiaire, ainsi que sa structure.



Après l'excision de la base lésée, Fpg réalise son activité AP lyase (Figure 33). On peut noter ici que Fpg peut aussi prendre en charge des sites abasiques créés par d'autres enzymes ou par des dépurinations/dépyrimidations hydrolytiques, la protéine se liant covalamment sur le site abasique sous forme aldéhyde en formant l'intermédiaire de type base de schiff. Dans ce cas, la protéine considère le site AP comme un nouveau substrat et s'y fixe. De nombreuses structures ont été résolues pour de tels complexes non-covalents avec des analogues non métabolisables de sites AP. En l'occurrence, on possède les structure de *L/Fpg* mutante P1G liée à un 1,3-propanediol (code PDB 1KFV : Serre *et al.*, 2002), de *BstFpg* liée à un site abasique réduit placé en face d'un C, d'un T ou d'un G (codes PDB 1L1T, 1L2C et 1L2D, respectivement : Fromme *et al.*, 2002), et de *L/Fpg* de type sauvage ou mutante P1G liées à un 1,3-propanediol ou un tétrahydrofurane (codes PDB 1NNJ, 1PJJ, 1PJI et 1PM5 : Pereira de Jésus *et al.*, 2005). On peut noter que la structure avec le site abasique réduit est d'ailleurs probablement la plus proche de la réalité à cause de son cycle désoxyribose ouvert. A ce niveau, une base arrache un proton en C2' du sucre, et l'intermédiaire de type base de Schiff subit un réarrangement électronique avec délocalisation d'électrons  $\pi$ . Cela conduit au clivage de la liaison phosphodiester en 3' du site abasique par  $\beta$ -élimination. Une seconde base arrache alors un proton en 4' du désoxyribose, ce qui produit une seconde délocalisation d'électrons  $\pi$ , induisant le clivage de la liaison phosphodiester en 5' du site abasique par une seconde  $\delta$ -élimination. A ce niveau, le complexe restant entre l'enzyme et ce qui reste du désoxyribose est hydrolysé, et l'enzyme est restituée sous forme active avec relargage de 4-oxo-2-pental. Les deux activités catalytiques de Fpg sont alors terminées, et le produit de réaction est une lacune d'un nucléoside délimitée par des extrémités 5' et 3' phosphate. On possède d'ailleurs une structure du complexe formé entre *BstFpg* et le produit final de réaction (code PDB 1L2B : Fromme *et al.*, 2002). Enfin, le complexe Fpg/ADN se dissocie, et l'enzyme peut à nouveau recommencer à chercher une lésion. On peut d'ailleurs noter ici que Fpg, en plus de ses activités ADN glycosylase et AP lyase, possède une activité 5'-dRPase (Graves *et al.*, 1992(a)). Cette activité permet de cliver par  $\beta$ -élimination en 3' d'un site abasique déjà clivé en 5' (par exemple par une AP endonucléase, chapitre I.5.3 sur les protéines métabolisant les sites AP et extrémités modifiées de l'ADN). Ce mécanisme est donc identique à une partie de l'AP lyase de Fpg, et ne sera donc pas détaillé ici.



**Figure 33 :** Représentation schématique du mécanisme catalytique de l'activité AP lyase de Fpg. Le code couleur est le même que pour la Figure 32.

Finalement, on peut noter que, bien que la cristallographie des rayons X soit une technique qui donne des informations sur un état « immobile » des macromolécules, la détermination de structures correspondant à différentes étapes des activités de Fpg a permis de mieux comprendre et de valider, en une demi douzaine d'années, les mécanismes réactionnels de cette protéine qui se déroulent en une fraction de seconde.

## **I.8. L'objectif du travail de thèse et les stratégies utilisées**

### **I.8.1. Stratégies pour l'étude de complexes Fpg/ADN**

Bien que la protéine Fpg de *Escherichia coli* (*EcFpg*) soit la forme pour laquelle le plus grand nombre d'études ont été réalisées, ce travail de thèse utilisera uniquement la protéine Fpg de *Lactococcus lactis* (*LlFpg*). En fait, il a été observé qu'à haute concentration et à 4°C, la protéine Fpg purifiée est lentement clivée en deux peptides inactifs, ce qui rend sa conservation difficile. Cependant, cette instabilité intrinsèque est plus prononcée pour *EcFpg* que pour son homologue *LlFpg* (Pereira de Jesus *et al.*, 2002). Cela nous a conduits à utiliser préférentiellement la protéine *LlFpg* afin de pouvoir mener en parallèle des études structurales et fonctionnelles.

D'autre part, les différents substrats qui seront utilisés dans nos études seront de deux types en fonction de ce que l'on veut regarder. En l'occurrence, les études d'activité de Fpg seront réalisées sur des substrats vrais, pouvant donc être traités normalement par l'enzyme. Par contre, les études de fixation de Fpg sur des ADNs lésés et les études cristallographiques de ces mêmes complexes seront faites sur des analogues de substrats afin d'éviter que Fpg ne les métabolise (chapitre I.5.2.5 sur les stratégies permettant des études structurales sur les ADN glycosylases). Les analogues utilisés au laboratoire correspondent à ceux définis par Ober et collaborateurs en 2003, où la partie base nucléique des lésions est inchangée, mais où le désoxyribose est remplacé par un cyclopentane (dérivés carbocycliques). En effet, le cyclopentane ne possède pas d'oxygène dans son cycle, ce qui rend impossible l'ouverture de ce cycle, et donc empêche l'activité ADN glycosylase de Fpg (chapitre I.7.3 sur le mécanisme d'action de Fpg et les structures correspondantes).

## **I.8.2. Objectif du travail de thèse**

L'objectif essentiel de mon travail de thèse est de préciser la spécificité de substrat de *LIFpg* via des études fonctionnelles utilisant des techniques biochimiques. Ces études s'inscrivent en complément d'études structurales par radiocristallographie des rayons X, dont les résultats seront systématiquement mis en parallèle avec ceux des études fonctionnelles. Les résultats obtenus seront organisés en deux grandes parties.

La première partie traitera des déterminants structuraux et fonctionnels relatifs à la reconnaissance et la métabolisation par Fpg des purines oxydées. L'effet de la base opposée à différentes lésions puriques sur la fixation et l'activité de *LIFpg* sera présenté. Bien que de nombreuses études aient été réalisées dans ce sens, celles-ci ont toutes été basées sur *EcFpg*. Cette étude permettra donc de voir si il y a ou non réellement des différences de spécificité de substrat entre les Fpg provenant de différentes bactéries. Au cours de ces études, un intérêt particulier sera porté sur la capacité de Fpg à réparer des lésions de grande taille. En effet, bien que la poche du site actif de Fpg soit relativement large (chapitre I.7 sur la protéine bactérienne Fpg), elle semble tout de même trop étroite pour pouvoir stabiliser certaines lésions de taille parfois très importante. En l'occurrence, le comportement de Fpg vis-à-vis de résidus FapyG (une guanine oxydée) substitués ou non en N7 par de gros adduits sera étudié (voir Figure 29). Les structures de Fpg complexée à des Fapys substitués ou non en N7 ayant été résolues dans l'équipe, celles-ci seront aussi discutées. Quelques données relatives à la reconnaissance des lésions encombrantes par la protéine Ogg1 de levure, l'homologue fonctionnel eucaryote de Fpg, seront également présentées.

La seconde partie des résultats sera consacrée à l'étude de la reconnaissance et de la métabolisation par Fpg d'une pyrimidine oxydée. Comme modèle d'étude, nous avons choisi la 5-OH-5-Me-Hydantoïne, une thymine oxydée décrite précédemment comme substrat de Fpg. Nous montrerons que, dépendamment de la base opposée au dommage, cette lésion constitue un bon ou un mauvais substrat pour Fpg. De plus, nous montrerons que, suivant ce contexte d'appariement, cette lésion est un meilleur ou un plus mauvais substrat pour l'enzyme Nth de *E. coli* (une ADN glycosylase spécifique de l'élimination des pyrimidines oxydées) que pour Fpg. Les bases structurales de la reconnaissance de cette lésion par *LIFpg* seront aussi présentées. En outre, au cours de cette dernière étude, nous avons mis en évidence la possibilité d'engager Fpg dans un complexe covalent avec la 5-OH-5-Me-Hydantoïne. Pour finir, une caractérisation structurale et fonctionnelle de cette réaction de type « suicide » sera présentée.

## **II. Résultats**

## **II. Résultats**

### **II.1. Nouveaux éléments structuraux et fonctionnels sur la reconnaissance et la métabolisation des purines oxydées par la Fpg de *Lactococcus lactis* (LIFpg)**

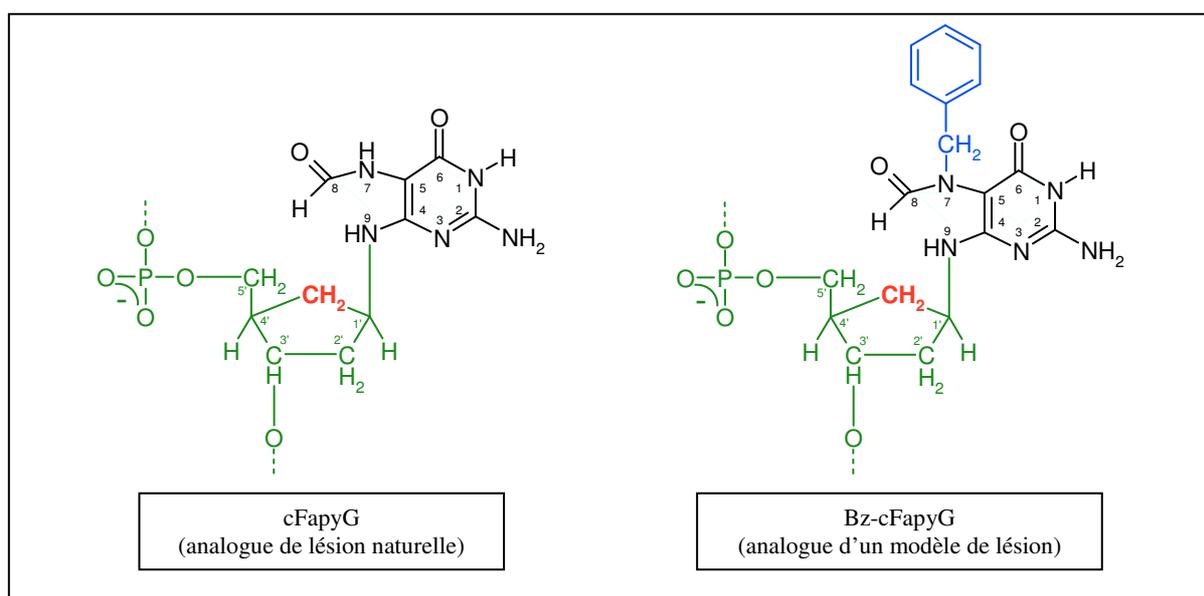
#### **II.1.1. Reconnaissance d'une lésion encombrante par LIFpg**

Comme nous l'avons déjà développé dans l'introduction du manuscrit, la 2,6-diamino-4-hydroxy-5-formamidopyrimidine ou FapyG est l'un des substrats de référence de Fpg sur lequel elle possède une forte activité (Wiederholt *et al.*, 2003). De plus, Fpg est aussi capable d'exciser des dérivés du FapyG substitués en N7 par des adduits plus ou moins encombrants, tels que le N<sup>7</sup>-méthyle-FapyG (N<sup>7</sup>-Me-FapyG; Tchou *et al.*, 1991; Tudek *et al.*, 1998; Asagoshi *et al.*, 2000), le N<sup>7</sup>-éthyle-FapyG (N<sup>7</sup>-Et-FapyG; Tudek *et al.*, 1998), le FapyG avec un adduit de gaz moutarde sulfuré (N<sup>7</sup>-SM-FapyG; Li *et al.*, 1997), le FapyG avec un adduit de gaz moutarde phosphoramidate (N<sup>7</sup>-PM-FapyG; Chetsanga *et al.*, 1982) et le FapyG avec un adduit d'aflatoxine B<sub>1</sub> (N<sup>7</sup>-AFB<sub>1</sub>-FapyG; Chetsanga *et al.*, 1983; Alekseyev *et al.*, 2004) (Figure 29, chapitre I.7.2 sur les spécificités de substrat de Fpg). Cependant, bien que la capacité de Fpg à exciser toutes ces lésions ait été démontrée, l'efficacité de Fpg sur celles-ci est apparue comme décroissante avec l'augmentation de la taille de l'adduit en N7, avec une exception pour le N<sup>7</sup>-SM-FapyG qui est a priori au moins aussi bien excisé que le N<sup>7</sup>-Me-FapyG (Li *et al.*, 1997). Il apparaît donc que, d'une façon générale, l'augmentation de l'encombrement stérique en position N7 des FapyGs soit un paramètre jouant défavorablement sur l'activité de Fpg.

Afin de mieux comprendre les déterminants structuraux de l'activité ADN glycosylase de Fpg, plusieurs structures cristallographiques de complexes formés entre l'enzyme et différents substrats ou analogues de substrats pour l'activité ADN glycosylase de l'enzyme (FapyG, 8-oxo-G, DHU) ont été réalisées à ce jour (Fromme *et al.*, 2003a; Coste *et al.*, 2004). Celles-ci ont montré que Fpg stabilise les lésions dans une conformation extrahélicale dans la poche du site actif. Cette extrusion du nucléoside endommagé a pour effet d'exposer son C1' à l'attaque par l'azote du résidu Proline N-terminal de l'enzyme. Il s'agit là de la première étape de l'activité ADN glycosylase de l'enzyme (Figure 32). De plus, ces structures ont montré que Fpg utilise différents modes de fixation pour permettre l'attaque en C1' sur différentes lésions. Ces multiples possibilités de modes de fixation des lésions sont d'ailleurs à l'origine de la capacité de Fpg à métaboliser différentes lésions (chapitre I.7.2 sur les spécificités de substrat

de Fpg), expliquant la spécificité de substrat très large de cette enzyme. Cependant, les structures évoquées précédemment se sont limitées à des lésions peu encombrantes, n'expliquant donc pas comment Fpg était capable de stabiliser des lésions parfois très grosses telles que les FapyGs substitués en N7 par des adduits de grande taille. En effet, bien que Fpg possède une poche de fixation exceptionnellement large parmi les ADN glycosylases procaryotes (Fromme *et al.*, 2003a; Zharkov *et al.*, 2003; Coste *et al.*, 2004), celle-ci semble tout de même trop petite pour accommoder des lésions très encombrantes, telles que le  $N^7$ -PM-FapyG ou le  $N^7$ -AFB<sub>1</sub>-FapyG par exemple.

Dans le but de comprendre comment Fpg arrive à stabiliser de telles lésions, nous avons réalisé une étude structurale et fonctionnelle sur Fpg complexée à des oligonucléotides de 14 paires de bases (Tableau 20, « Matériel et méthodes ») contenant un analogue carbocyclique du FapyG (cFapyG) ou du  $N^7$ -benzyle-FapyG (Bz-cFapyG) (Figure 34). De par leur nature carbocyclique, ces analogues ne sont pas métabolisables par les ADN-glycosylases (Chepanoske *et al.*, 1999). Le Bz-cFapyG n'est pas un analogue de lésion naturelle, mais est en fait un modèle moléculaire de FapyG encombré qui a été spécialement synthétisé par des collaborateurs chimistes pour cette étude (Pr. Thomas Carell, Ludwig-Maximilians Universität, Munich, Allemagne).



**Figure 34 : Structure plane des modèles de lésion cFapyG et Bz-cFapyG.**

La partie FapyG des deux molécules est représentée en noir, et la partie squelette phosphate/cyclopentane de ces analogues de lésions sont représentées en vert. Le carbone remplaçant l'oxygène du désoxyribose est représenté en rouge. L'adduit synthétique benzyle du Bz-cFapyG est représenté en bleu.

**II.1.1.1. Publication “L’enzyme bactérienne Fpg du système BER reconnaît le FapydG substitué en N7 par un gros adduit *via* un mode de fixation non-productif” (Coste *et al.*, 2008)**

**Résumé de l’article :**

Fpg est une enzyme bactérienne du système de réparation par excision de base qui enlève les purines oxydées de l’ADN. Ce travail montre que Fpg et son homologue eucaryote Ogg1 reconnaissent avec une haute affinité le FapyG et l’encombrant N<sup>7</sup>-benzyle-FapyG (Bz-FapyG). L’analyse comparative des structures cristallines de complexes stables entre Fpg et des ADNs contenant des nucléosides carbocycliques cFapyG et Bz-cFapyG fournit les bases moléculaires permettant d’expliquer que Fpg se fixe sur les deux lésions avec la même affinité sans pour autant être capable de les métaboliser avec la même efficacité. Pour accommoder l’encombrement stérique du groupement benzyle, Fpg sélectionne le rotamère adéquat du groupement formamido du Bz-cFapyG extrahélical, forçant le groupement encombrant à sortir de la poche de fixation. Contrairement au mode de fixation du cFapyG, la reconnaissance particulière du Bz-cFapyG amène Fpg à former un complexe non-productif qui masquerai la lésion aux autres systèmes de réparation tels que le NER, et ralentiraient ainsi son élimination de l’ADN.

# Bacterial Base Excision Repair Enzyme Fpg Recognizes Bulky $N^7$ -Substituted-FapydG Lesion via Unproductive Binding Mode

Franck Coste,<sup>1</sup> Matthias Ober,<sup>2</sup> Yann-Vaï Le Bihan,<sup>1</sup> Maria Angeles Izquierdo,<sup>2</sup> Nadège Hervouet,<sup>1</sup> Heiko Mueller,<sup>2</sup> Thomas Carell,<sup>2,\*</sup> and Bertrand Castaing<sup>1,\*</sup>

<sup>1</sup>Centre de Biophysique Moléculaire, UPR4301, CNRS, rue Charles Sadron, 45071 Orléans cedex 02, France

<sup>2</sup>Center for Integrated Protein Science CIPS<sup>M</sup>, Department of Chemistry and Biochemistry, Ludwig-Maximilians Universität München, Feodor-Lynen Straße 25, D-81377 Munich, Germany

\*Correspondence: castaing@cnrs-orleans.fr (B.C.), thomas.carell@cup.uni-muenchen.de (T.C.)

DOI 10.1016/j.chembiol.2008.05.014

## SUMMARY

Fpg is a bacterial base excision repair enzyme that removes oxidized purines from DNA. This work shows that Fpg and its eukaryote homolog Ogg1 recognize with high affinity FapydG and bulky  $N^7$ -benzyl-FapydG (Bz-FapydG). The comparative crystal structure analysis of stable complexes between Fpg and carbocyclic cFapydG or Bz-cFapydG nucleoside-containing DNA provides the molecular basis of the ability of Fpg to bind both lesions with the same affinity and to differently process them. To accommodate the steric hindrance of the benzyl group, Fpg selects the adequate rotamer of the extrahelical Bz-cFapydG formamido group, forcing the bulky group to go outside the binding pocket. Contrary to the binding mode of cFapydG, the particular recognition of Bz-cFapydG leads the BER enzymes to unproductive complexes which would hide the lesion and slow down its repair by the NER machinery.

## INTRODUCTION

Reactive oxygen species (ROS) and alkylating agents are responsible for numerous types of DNA base damage which can interfere with DNA transactions such as replication, transcription, and recombination (Lindahl, 1993). Base damage persistence in cellular DNA plays a role in cell death, mutagenesis, carcinogenesis, aging, and various degenerative diseases (Beckman and Ames, 1997; Sedgwick et al., 2007; Wilson and Bohr, 2007). The most abundant lesions formed in DNA from guanine are 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapydG) and 7,8-dihydro-8-oxoguanine (8-oxodG) (Douki et al., 1997). Usually resulting from ROS action, Fapy-derivatives can also be generated in DNA from intermediate alkylated products of purines at the  $N^7$  position (Figures 1A and 1B). To counteract the deleterious effects of these damages, cells have evolved many DNA repair strategies which have been conserved from bacteria to higher eukaryotes (Wood et al., 2005). The main way to repair base damages is the base excision repair (BER) pathway.

The formamidopyrimidine-DNA glycosylase (Fpg or MutM) was initially identified as a BER enzyme able to specifically

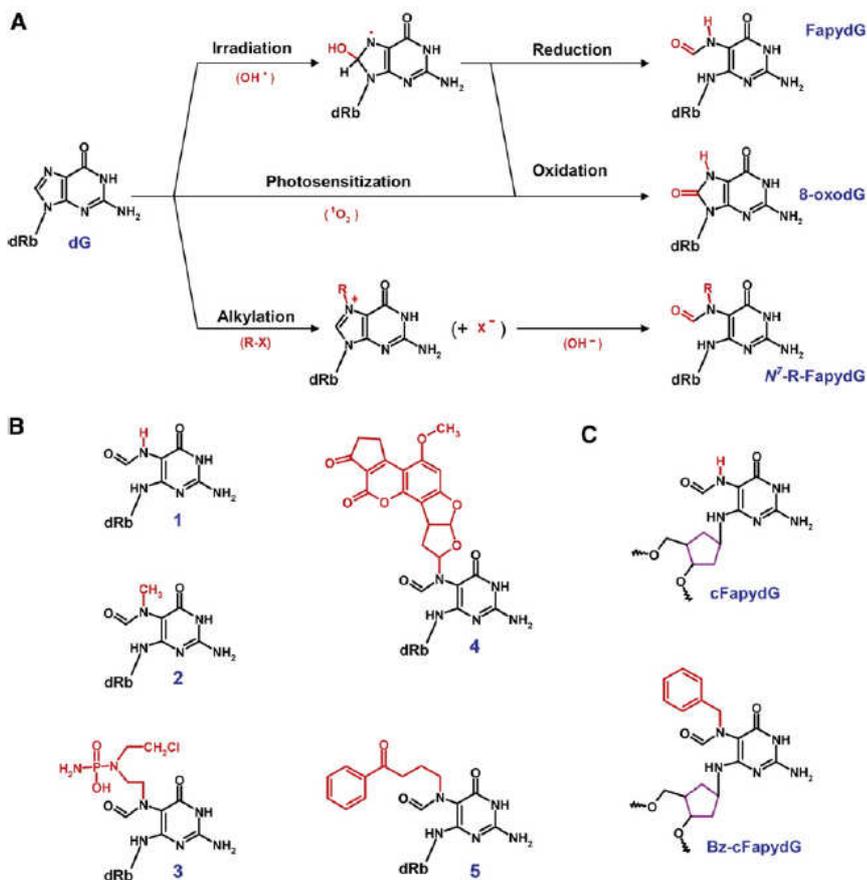
excise FapydG (Chetsanga and Lindahl, 1979). More recently, it has also been shown that Fpg efficiently removes the major oxidation product of purines, 8-oxodG (Tchou et al., 1994; Castaing et al., 1993). In addition to its DNA glycosylase activity, Fpg also displays an efficient AP lyase activity allowing the successive cleavage of the phosphodiester bonds at the 3' and the 5' sides of the abasic (AP) site, respectively. As a consequence, a one nucleoside gap is formed in DNA (O'Connor and Laval, 1989).

The recent crystal structures of Fpg bound to damaged DNA (AP site, 8-oxodG or FapydG) reveal that through binding, Fpg flips the damaged nucleoside out of the DNA helix in an extrahelical conformation inside an enzyme binding pocket and exposes it to the nucleophilic attack of the N-terminal proline, P1 (Serre et al., 2002; Fromme and Verdine, 2003; Coste et al., 2004). However, these comparative data suggest that Fpg uses several different binding modes to recognize the damage. In order to better understand the Fpg binding modes of Fapy-derivatives, we have solved the crystal structures of the wild-type *L. lactis* Fpg enzyme bound to 14-mer DNA duplexes containing either a cFapydG or a  $N^7$ -benzyl-cFapydG (Bz-cFapydG) residue. The latter is a model for bulky Fapy-DNA adducts (Figure 1C). The present work provides new structural and functional insights into the ability of this enzyme class to accommodate bulky DNA adducts in its substrate binding pocket and to process it.

## RESULTS AND DISCUSSION

### Bulky Bz-cFapydG-Containing DNA Is a High Affinity Ligand for the Fpg and Ogg1 Proteins

The chemical replacement of the heterocyclic oxygen of the deoxyribose by a  $-CH_2-$  group in FapydG leads to a Fpg substrate analog (cFapydG, Figure 1C) which is recognized but cannot be processed by the enzyme (Coste et al., 2004; Ober et al., 2003, 2005). Here, we investigate the ability of Fpg to bind to a different analog which mimics a bulky adduct FapydG lesion (Bz-cFapydG, Figure 1C and Supplemental Experimental Procedures, available online). To this end, the binding to Bz-cFapydG-containing DNA was quantified using electrophoretic mobility shift assays (EMSA) and the data were compared with the binding of Fpg to cFapydG-DNA (Figure 2). Both oligonucleotides were recognized by *L. lactis* Fpg (LIFpg) with about the same high affinity (Figure 2B). Previous works have shown that  $N^7$ -alkyl-Fapy derivatives, such as AFB1-FapydG, PM-FapydG

**Figure 1. Guanine Oxidized**

(A) Different pathways to produce the major oxidized products of guanine in DNA.

(B) Fapy derivatives. As shown in (A), unsubstituted FapydG (1) is radiation-induced in DNA and *N*<sup>7</sup>-substituted FapydG residues result from the destabilization of the imidazole ring followed by alkylation at the *N*<sup>7</sup> position of guanine: alkylation by (2) the dimethylsulfate (DMS), (3) the phosphoramidate mustard (PM), (4) the aflatoxin B1 (AFB1), and (5) the 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) (Tudek, 2003).

(C) Fapy analogs used in this study. These analogs were stabilized by the chemical replacement of the deoxyribose of the damage nucleoside by a cyclopentane (Supplemental Experimental Procedures).

(Figure 1B, compounds 4 and 3, respectively), and AF-irodG (AF = aminofluoren, *iro* = imidazole ring-opened), are very poor substrates for Fpg (Chetsanga et al., 1982; Tchou et al., 1994; Alekseyev et al., 2004). The Fpg glycosylase activity is strongly affected by the steric bulk at the *N*<sup>7</sup> position of FapydG (Tudek et al., 1998). Our data therefore suggest that the difficulty in excising bulky-Fapy adducts is not inevitably related to the Fpg ability to bind damaged DNA. It is however possible that the alkyl group *N*<sup>7</sup>-substitution leads the enzyme to form a quasi-abortive complex. A similar explanation can be proposed for the eukaryote Fpg homolog, the Ogg1 protein. Like Fpg, yOgg1 forms a stable complex with cFapydG- and Bz-cFapydG-DNA (Figure 2). Based on the EMSA studies, we therefore propose that Fpg (Ogg1) uses a tight but unproductive binding mode for the recognition of *N*<sup>7</sup>-bulky adducted FapydG lesions quite different from those which were so far structurally characterized (Fromme and Verdine, 2003; Coste et al., 2004). To characterize the molecular basis of this unproductive complex, we initiated a structural study. Because we cannot exclude that the mutation present in the previously crystallized Fpg versions ( $\Delta\text{P1-LIFpg}$ , E2Q-BstFpg) might influence the recognition of the damage, we performed the structural study of the wild-type enzyme bound to cFapydG-DNA.

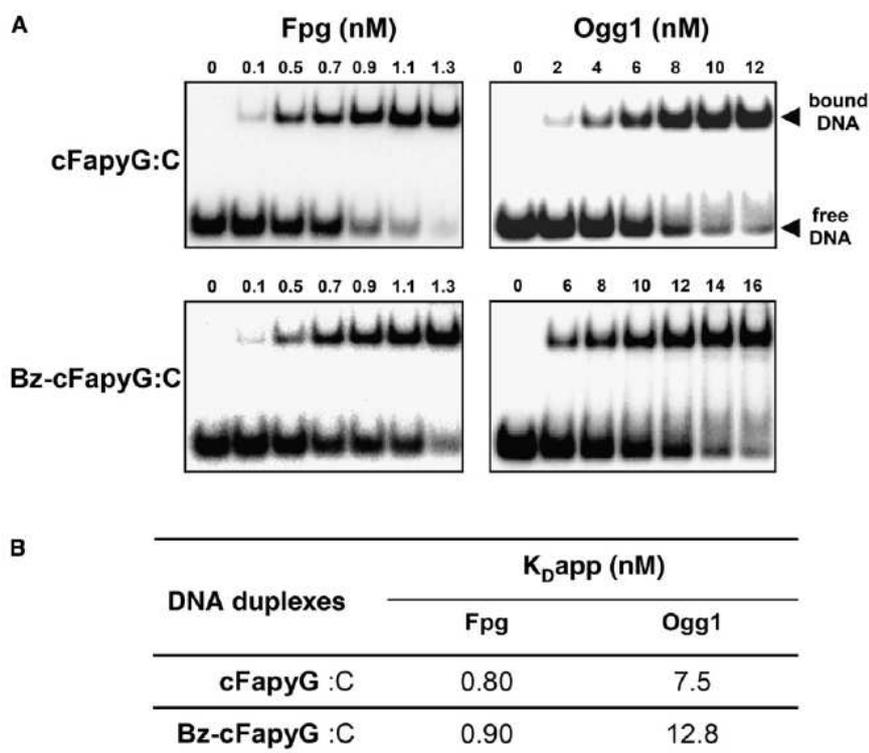
#### Accommodation of the Bz-cFapydG Steric Hindrance by Fpg

The crystal structures of the wild-type LIFpg bound to a 14-mer DNA duplex containing either cFapydG or Bz-cFapydG were

solved (Table 1). Between both models, the global fold of Fpg and DNA is unchanged as compared with previous structures (RMSD of 0.37 Å is observed between the protein  $\text{C}\alpha$ -backbones of both models). Few structural rearrangements of Fpg are limited to the  $\alpha\text{F}$ - $\beta\text{9}$  loop. Both lesions are stabilized in an extrahelical *anti* conformation inside the Fpg binding pocket. No difference was observed between the FapydG bound by the wild-type LIFpg versus the so-far used defective mutant  $\Delta\text{P1-LIFpg}$  (Coste et al., 2004). However, the Fpg binding modes of cFapydG and Bz-cFapydG

are significantly different and damages are not superimposed inside the Fpg binding pocket (Figure S1). The pyrimidine moieties of FapydG derivatives are not coplanar and the Bz-cFapydG appears slightly less extrahelical than cFapydG. Only the sugar moieties of both damages (here the cyclopentane) are quasi-superimposed (Figure S1).

The enzyme does not select the same rotamer of the formamido group of FapydG derivatives (Figure 3). The steric hindrance of the benzyl group of Bz-cFapydG is reduced by changing the conformation of the formamido group (Figure 3A). We thus observe a difference of almost 180° of the dihedral angle of the formamido group between cFapydG and Bz-cFapydG (Figure 3B). In DNA, the exact conformation of the different Fapy lesions is not well understood and, possibly, a mixture of multiple interconverting isomeric forms is present. This structural complexity is increased by the possibility of anomerization (Haraguchi et al., 2001; Greenberg et al., 2001; Brown et al., 2006). Although four *N*-formyl geometrical isomers are possible for the Fapy-derivatives (Figure 3C), only a subgroup was experimentally observed for the nucleobases, nucleosides, and nucleoside-containing DNA. Using NMR spectroscopy and HPLC analysis, it was shown that the *N*<sup>7</sup>-Me-FapyG base, *N*<sup>7</sup>-Me-FapydG, and *N*<sup>7</sup>-S-Et-GSH-FapydG nucleosides exist in equilibrium between the geometrical isomers I and II of Figure 3C which are related to each other by a 180° rotation about the C<sup>5</sup>-*N*<sup>7</sup> and *N*<sup>7</sup>-C<sup>8</sup> bonds (Boiteux et al., 1984; Humphreys and Guengerich, 1991). Tomasz and coworkers have identified the geometrical



**Figure 2. Fpg and Ogg1 Binding to cFapydG- and Bz-cFapydG-Containing DNA**

(A) EMSA. Binding experiments were carried out as described elsewhere (Castaing et al., 1999). As indicated, cFapydG- and Bz-cFapydG-14-mer DNA duplexes were incubated with increased protein concentrations and the resulting mixtures were analyzed by EMSA.

(B) Apparent dissociation constants ( $K_{Dapp}$ ).

isomers II and III as the minor and major rotamers of the  $N^7$ -Et-FapydG nucleoside, respectively, whereas the rotamer I is not significantly populated contrary to  $N^7$ -Me-FapydG (Tomasz et al., 1987). It was also shown that the major separated HPLC fraction of AFB1-FapydG base and the corresponding nucleoside (Figure 1B) can be attributed to two geometrical isomers, the major rotamer III and the minor rotamer II (Lin et al., 1977; Hertzog et al., 1982; Brown et al., 2006). Taken together, these results suggest that the nature of the substituent at the  $N^7$  position modulates the conformational equilibrium. The solution structure of AFB1-FapydG-containing DNA duplex has revealed that the formamide group adopts the rotamer IV (Mao et al., 1998). It was proposed that the *E. coli* Fpg protein preferentially recognizes in DNA the geometrical isomer II of the  $N^7$ -Me-FapydG, which is a very efficient Fpg substrate (Boiteux et al., 1984). Based on the crystal structure of wtLIFpg/cFapydG-DNA, it is easy to replace the  $N^7$ -proton of the FapydG base in its rotamer II by a methyl group without creating a steric clash in the binding pocket (data not shown). This suggests that the  $N^7$ -Me-FapydG can be indeed bound as its geometrical isomer II and processed by the enzyme like the unsubstituted FapydG. Such a mechanism is not possible for Fapy-derivatives  $N^7$ -substituted by bulky adducts. As illustrated by the present work, the enzyme can accommodate the  $N^7$ -bulky lesion as the rotamer IV which enables the bulky group to stay outside the binding pocket (Figure 3A).

#### Fine Structural Determinants for the Recognition of the Flipped-Out Fapy-Nucleobases inside the Fpg Binding Pocket

The ability of the enzyme to bind either the rotamer II (FapydG) or the rotamer IV (Bz-FapydG) of FapydG-derivatives is accompa-

nied by drastic changes of the molecular interactions between Fpg and the extrahelical damaged base (Figure 4). All the hydrogen donors and acceptors of FapydG are contacted by the enzyme (Figure 4C). In particular, all the specific determinants of the damaged guanine (as compared to a normal guanine) are recognized: the  $N^7$ -proton by the side chain of T221, the  $C^8$ -keto by the hydroxyl of Y238 through a w325-mediated interaction, and the  $N^9$ -proton by the side chain of E76 and the main chain of M75 through w215-mediated interactions (Figures 4A and 4C). Between FapydG and Bz-FapydG, only the three interactions with the  $N^1$ ,  $N^2$ , and  $O^6$  groups of the damaged guanine are conserved (highlighted by a gray background in Figure 4C), while  $N^3$  and  $N^7$  of Bz-FapydG are not directly contacted. The spatial position of the  $N^7$ -benzyl group of Bz-FapydG is constrained by the structural water molecules w323 and w182 which play a similar role to that of T221 in the case of FapydG. Whereas 11 amino acid residues of the enzyme are mobilized for the recognition of FapydG, only seven are involved in the interaction with Bz-FapydG. However, most of these interactions are established by residues conserved in Fpg primary structures indicating a high level of specificity in the recognition of both lesions. The flipping out of the lesion by the enzyme exposes C1' of the damaged nucleoside to the nucleophilic attack of the active site P1 residue. Indeed, P1 forms an imino enzyme-DNA intermediate during catalysis which can be easily trapped by irreversible reduction (Zharkov et al., 1997). In both structures, the C1' of the damage is located at 3.6–3.7 Å away from the P1 amino group. Interestingly, because of the rotamer IV selection, the catalytic proline P1 is used by the enzyme for the recognition of the formyl group of the Bz-FapydG whereas no such interaction is established in the case of FapydG binding. Thus, the amino-proton and the  $C\beta$  groups of P1 are contacted directly, and via w185, by the  $C^8$ -keto group of the Bz-FapydG formyl group (Figures 4B and 4C). Considering the essential catalytic role of P1, the ( $C^8$ -P1) interactions observed in the structure of LIFpg bound to Bz-cFapydG-DNA appear to be functionally unfavorable for the subsequent processing of the bulky adduct.

In addition, when we compare the binding mode of FapydG by the wt-LIFpg protein (this work) and by the defective mutant  $\Delta$ P1-LIFpg (Coste et al., 2004), we observe differences in the flexible part of the loop  $\alpha$ F- $\beta$ 9. This loop adopts a precise fold in the wt-LIFpg/cFapydG-DNA model, whereas in the previous

Table 1. Data Collection and Refinement Statistics

	Fpg/DNA models	
	wt-LIFpg/cFapydG	wt-LIFpg/Bz-cFapydG
Data Collection Statistics		
Radiation source	ESRF BM30A	ESRF ID23-1
Wavelength (Å)	0.92004	0.97935
Spacegroup	P4 <sub>1</sub> 2 <sub>1</sub> 2	P4 <sub>1</sub> 2 <sub>1</sub> 2
Cell dimensions	a = b = 92.513 Å, c = 142.867 Å	a = b = 91.356 Å, c = 140.943 Å
Total observations	212432	331651
Unique reflections	41650	47718
Completeness (%)	91.4 (91.4) <sup>a</sup>	100.0 (100.0) <sup>b</sup>
Redundancy	5.1 (4.8) <sup>a</sup>	7.0 (7.1) <sup>b</sup>
R <sub>sym</sub> <sup>c</sup>	6.9 (28.4) <sup>a</sup>	9.9 (45.1) <sup>b</sup>
<I/σ>	17.1 (3.6) <sup>a</sup>	17.0 (3.3) <sup>b</sup>
Refinement and Model Statistics		
Resolution (Å)	17.25-1.95	45.69-1.90
Number of reflections used	39561	45259
R <sub>work</sub> (%) <sup>d</sup>	17.6	17.2
R <sub>free</sub> (%) <sup>d</sup>	20.9	20.0
Average B values		
All atoms (Å <sup>2</sup> )	36.08	23.37
Protein atoms (Å <sup>2</sup> )	32.03	19.88
DNA atoms (Å <sup>2</sup> )	44.71	32.87
Water atoms (Å <sup>2</sup> )	45.34	32.62
Root mean square deviation from ideality		
Bond lengths (Å)	0.020	0.018
Bond angles (°)	1.776	1.745
Torsion angles (°)	5.971	5.825
No. of atoms		
Protein	2128	2147
DNA	569	576
Glycerol	6	6
Water	397	335

<sup>a</sup>Values in parentheses refer to data in the highest resolution shell (2.06–1.95 Å).

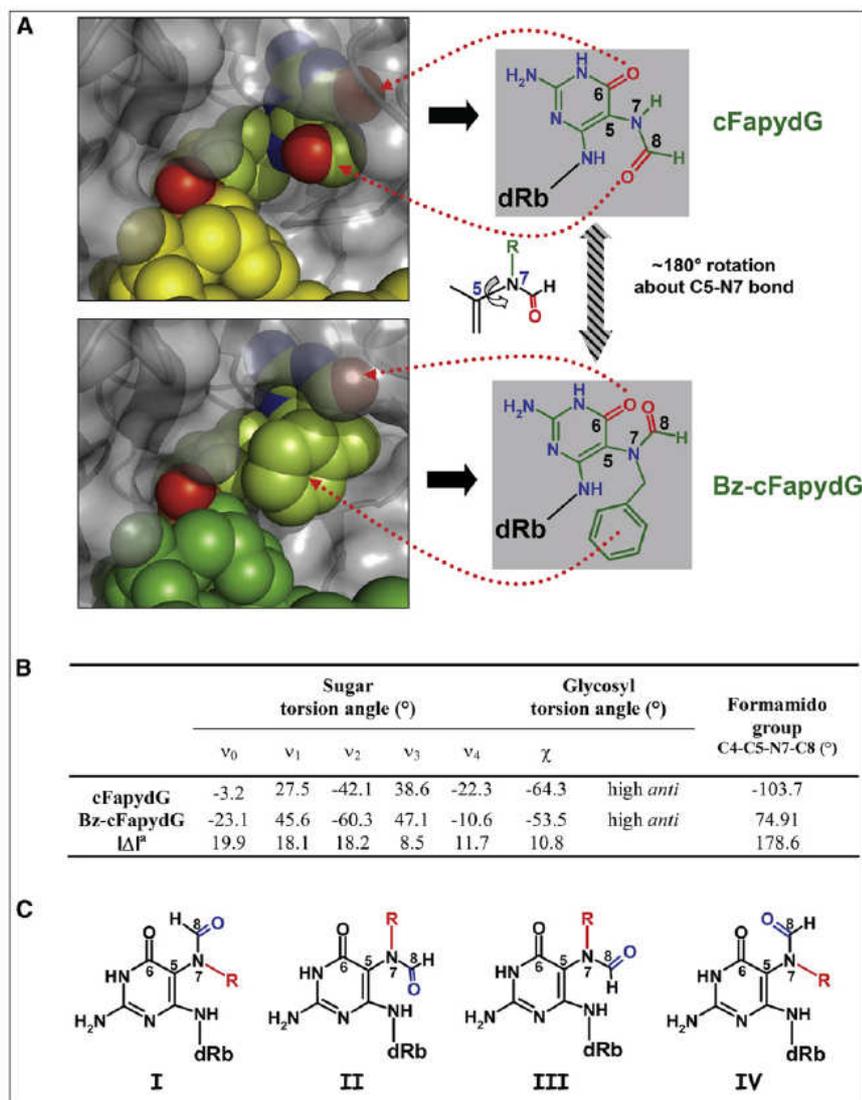
<sup>b</sup>Values in parentheses refer to data in the highest resolution shell (2.00–1.90 Å).

<sup>c</sup> $R_{sym} = \sum |I - \langle I \rangle| / \sum I$ , where  $I$  is the observed intensity and  $\langle I \rangle$  is the average intensity from multiple observations of the symmetry-related reflections.

<sup>d</sup> $R = \sum ||F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$ .  $R_{free}$  is the R-value for a subset of 5% of the reflection data, which were not included in the crystallographic refinement.

$\Delta$ P1-LIFpg/cFapydG crystal structure (Coste et al., 2004), the loop segment R<sub>222</sub>-A<sub>224</sub> was not visible and the loop segment L<sub>225</sub>-T<sub>228</sub> adopted a quite different fold (Figure 5A). Interestingly, the global fold of the visible part of this loop in  $\Delta$ P1-LIFpg/cFapydG was also found in previous crystal structures of LIFpg bound to abasic site analogs (Pereira de Jesus et al., 2005; Figure 5A). Taking into account all these structural studies, we propose that: (1) the very flexible part of the loop  $\alpha$ F- $\beta$ 9 can adopt two global folds which we term "open" and "closed" conformation and (2) P1 is absolutely required for observing the "closed" conformation of the loop when the enzyme recognizes the damaged guanine. The transition from the "open" to the "closed" conformation depends on three strictly conserved hinge residues G<sub>216</sub>, G<sub>226</sub>, and G<sub>229</sub>, separating the two seg-

ments of the flexible part of the  $\alpha$ F- $\beta$ 9 loop (Figure 5B). One of these "open" conformations was observed in the crystal structure of wt-LIFpg bound to tetrahydrofuran-containing DNA (in brown, Figure 5A; Pereira de Jesus et al., 2005). The "closed" conformation of the loop has already been observed in the crystal structures of Fpg from *B. stearothermophilus* (BstFpg) bound to 8-oxodG-containing DNA duplex and in the model of the free Fpg enzyme from *T. thermophilus* (TtFpg) (Fromme and Verdine, 2003; Sugahara et al., 2000). In spite of length and sequence differences (Figure 5B), the "closed" conformation of the loop in the Fpg proteins are perfectly superimposed (except for amino acid insertions; Figure 5C). Consequently, the "closed" conformation is directly associated with the recognition of the damaged guanine. Although the damage can be stabilized in an



**Figure 3. Fpg Rotameric Selection of the Formamido Group of Fapy Derivatives**

(A) Focus views of the extrahelical damaged nucleoside inside the Fpg binding pocket. Left panels show a zoom-in of the substrate binding pocket in LIFpg/cFapydG-DNA (top) and LIFpg/Bz-cFapydG-DNA (bottom). The solvent accessible surface and the ribbon drawing of the protein are indicated in gray. DNA is represented by its space filling model. Carbons of the DNA are colored in vivid green for both lesions and the sugar-phosphate backbone in yellow and green for cFapydG and Bz-cFapydG, respectively. Schematic representations of the structure of the damaged nucleoside are shown in the right panels.

(B) Base lesion geometry (footnote a = differences between cFapydG and Bz-cFapydG).

(C) Four geometrical isomers of Fapy derivatives. One rotamer can be obtained from another by 180° rotation about the C5-N7 and/or N7-C8 bonds.

*anti* or *syn* conformation (for FapydG and 8-oxodG, respectively), the stabilization of the loop in the “closed” conformation depends on the recognition of the O<sup>6</sup> group of the damaged guanine (Figure 5D). The “closed” conformation allows the stabilization of the extrahelical damage and its confinement in the substrate binding pocket. The loop segment S<sub>217</sub>-Y<sub>222</sub> plays the role of a mechanical valve which wraps the damaged base and closes the enzyme substrate binding pocket (Figure 5D). Like a “cap,” the conserved aromatic residue Y<sub>222</sub> fills the area defined by the loop in the “closed” conformation. The dynamic behavior of the loop enables the size and the volume of the binding pocket to be modulated for the various Fpg substrates. We further suggest that the “closed” conformation is essential for the catalytic process mediated by Fpg. This proposition is also supported by molecular dynamics simulations and structural studies in solution (Amara et al., 2004; Perlow-Poehnelt et al., 2004; Buchko et al., 2005; Amara and Serre, 2006). Amara and coworkers in particular have proposed that when Bst-Fpg is bound to normal guanine in DNA, the flexible part of the loop rap-

idly adopts a conformation significantly different from that of the enzyme bound to 8-oxodG-DNA. NMR studies of *E. coli* Fpg bound to 1,3-propanediol-DNA provided similar results (Buchko et al., 2005). We therefore propose that the flexible part of the  $\alpha$ F- $\beta$ 9 loop is necessary to discriminate a substrate from a nonsubstrate inside the binding pocket of the enzyme. Indisputably, an energetic and/or physical coupling exists between the catalytic P1 and the loop since the catalytic defective mutant  $\Delta$ P1-LIFpg bound to cFapydG-DNA displays the “open” conformation (Coste et al., 2004). On the contrary, the “closed” conformation is apparently not required for the AP lyase

activity of the enzyme since all crystal structures of Fpg bound to abasic site analogs show the  $\alpha$ F- $\beta$ 9 loop in the “open” conformation. Interestingly, the loop adopts its “open” conformation when the wt-LIFpg binds to Bz-cFapydG-DNA duplex (peptide chain in blue, Figure 5A). Due to the fact that Bz-FapydG is stabilized in its rotamer IV, the benzyl group of the damage is expelled from the binding pocket on the side of the flexible loop and thus prevents the enzyme from completely imprisoning the base damage inside the binding pocket (Figure 3). The C <sup>$\beta$</sup> -P1 interactions (see below) and the “open” conformation of the observed  $\alpha$ F- $\beta$ 9 loop together provide a structural explanation of the difficulty of Fpg in processing bulky Fapy-adducts as compared with the canonical substrates FapydG and 8-oxodG.

#### Biological Implication of the Recognition by DNA Glycosylases of FapydG Derivatives *N*<sup>7</sup>-Substituted by Bulky Adducts

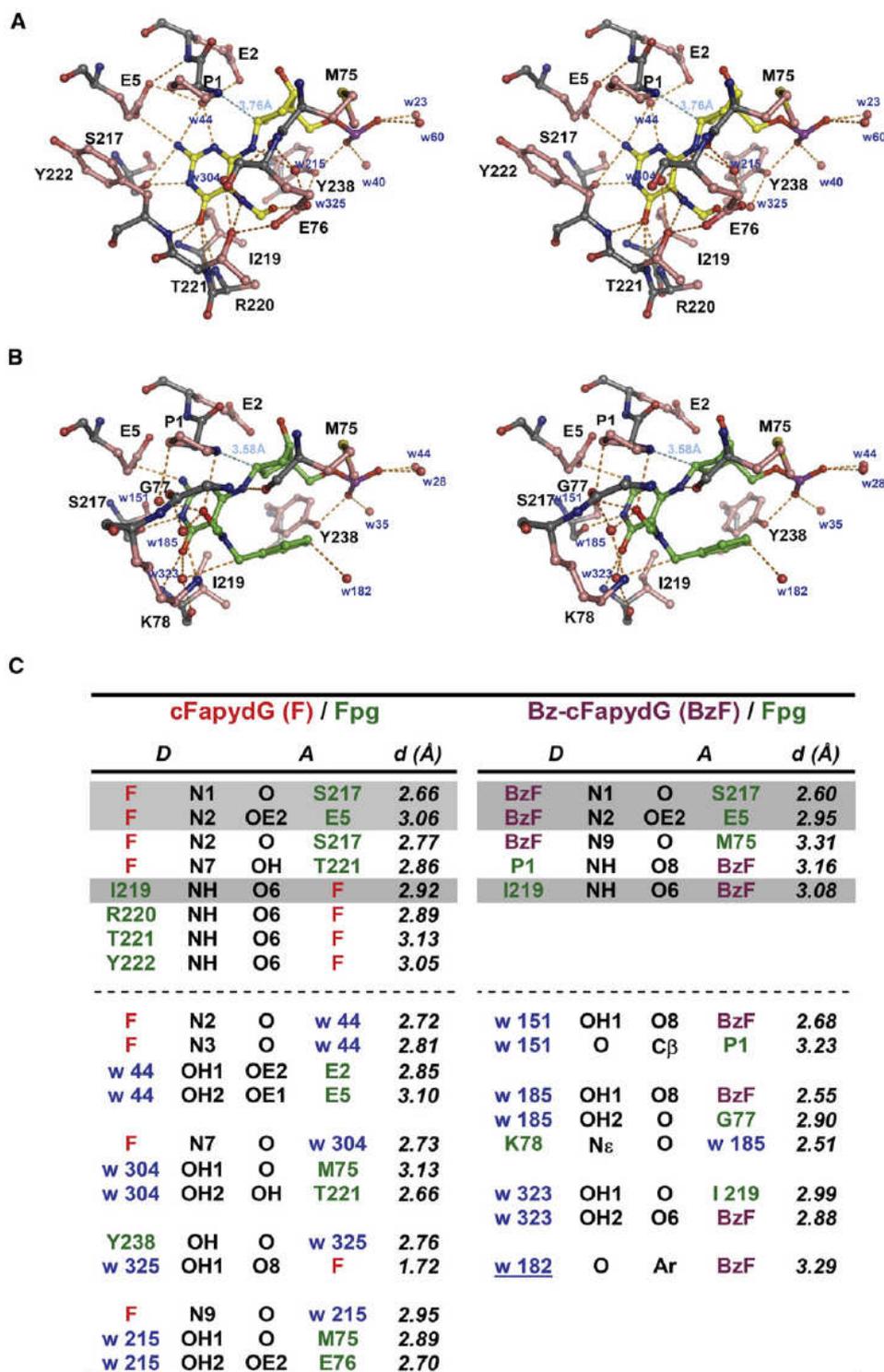
Fapy lesions and especially bulky-Fapy lesions in DNA represent challenges for DNA replication and repair enzymes (Tudek,

2003). Numerous *in vitro* and *in vivo* studies have shown that Fapy lesions constitute fairly strong but not absolute blocks to DNA synthesis (O'Connor et al., 1988; Aguilar et al., 1993; Asagoshi et al., 2002; Smela et al., 2001, 2002; Kalam et al., 2006; Patro et al., 2007). Bypass of the lesions can occur at low frequency in bacteria and eukaryotic cells which generally results in G to T transversions. Very interestingly, previous works have obtained evidence that replication blocking or bypass of bulky Fapy lesions is conditioned by the precise structure of the damage, i.e., its rotameric conformation and/or anomeric configuration (Smela et al., 2002; Brown et al., 2006). To prevent the interference of Fapy-lesions with the replication machinery, cells have evolved DNA repair strategies. Even though the major way to repair Fapy lesions is base excision repair (BER), Essigmann and coworkers have unambiguously demonstrated that AFB1-FapydG is not a substrate for Fpg and Ogg1 as compared to 8-oxodG and  $N^7$ -Me-FapydG (Tudek et al., 1998; Alekseyev et al., 2004; Hu et al., 2005). Until now, the explanation for this observation was the steric bulk of such lesions, preventing the enzyme from stabilizing the damage properly in its binding pocket. Here, we show that Fpg recognizes both the Bz-FapydG lesion and the unsubstituted FapydG lesion with the same affinity (Figure 2). This result correlates with the ability of the enzyme to stabilize both lesions in an extrahelical conformation inside its binding pocket (Figures 3 and 4). To accommodate the benzyl group of Bz-FapydG, the rotamer IV of the formamido group is stabilized by the enzyme, which leads to the expelling of the bulky group outside the binding pocket, whereas the FapydG base moiety of the damage is maintained inside. Based on the binding mode of Bz-FapydG, it is now possible to propose a model for the binding of the extrahelical AFB1-FapydG by Fpg by replacing the benzyl group of Bz-FapydG by the AFB1 adduct (Figure 6). This model supposes that Fpg is first able to melt the DNA duplex at the intercalated damaged site. We therefore propose that the stabilization of the rotamer IV is one of the lesion binding modes selected by the enzyme to accommodate bulky Fapy adducts in its substrate binding pocket. However, recognition of larger bulky Fapy adducts with intercalative properties, such as AFB1, needs further investigation (Figure 1B).

The present work confronted with previous structural and functional studies provides a detailed understanding of the ability of Fpg to recognize and process bulky DNA adducts. Based on the binding mode of FapydG (selection of the rotamer II), we can easily substitute the  $N^7$ -proton of the damaged base by a methyl group without disturbing the local structure observed. The ( $\delta^+$ )  $N^7$ -methyl group can displace the water molecule w304 and establish a direct interaction with the ( $\delta^-$ ) hydroxyl group of T221 (Figure 4). Based on these observations, the rotamer II of the Fapy lesion probably corresponds to the conformation present during catalysis because the  $N^7$ -Me-FapydG is excised in DNA very efficiently by the enzyme. A similar replacement by an ethyl group still seems to be possible but might already slightly disturb the local structure.  $N^7$ -Et-FapydG and AF-iroG are described as substrates 7-fold and 30-fold less efficiently excised by Fpg than  $N^7$ -Me-FapydG, which is in good agreement with our present crystal structures (Boiteux et al., 1989; Tudek et al., 1998). For  $N^7$ -substituents of FapydG that are bulkier than an ethyl group, the rotamer II cannot be

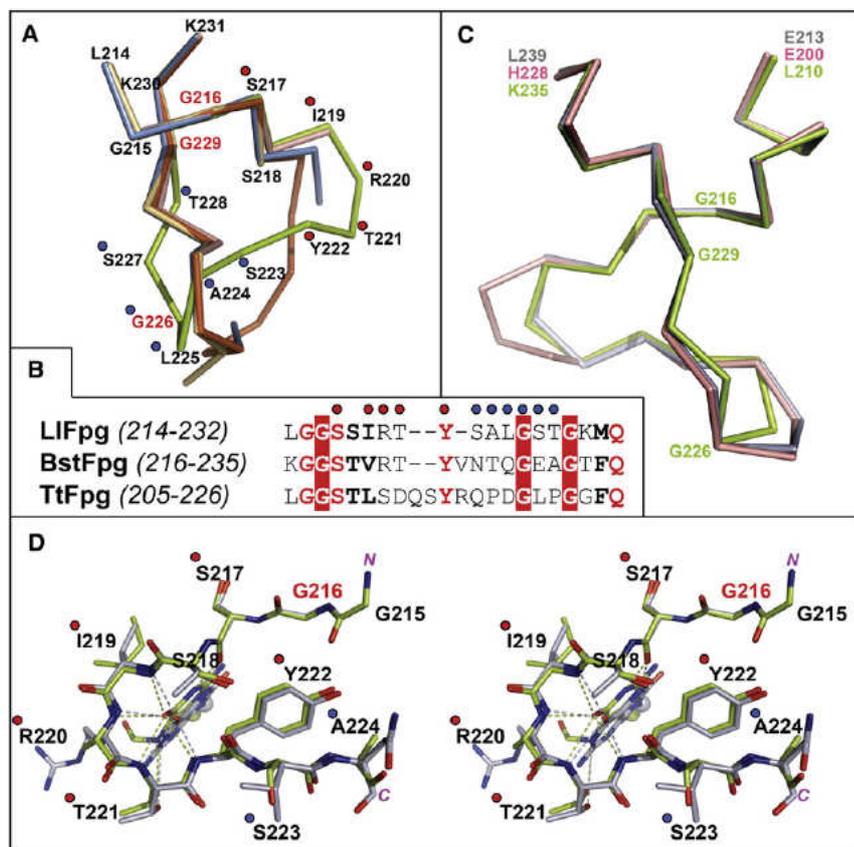
bound by the enzyme without creating a serious clash with the walls of the binding pocket. On the contrary, such adducts can be bound in their conformation IV which results in a reorganization of the interactions and of the water molecules inside the binding pocket. This structural induced-adjustment leads the formyl group of the damage to contact the amino group of the catalytic P1 (Figure 7A). Given the very low electrophilic character of the formyl group, this interaction is strong and stabilizes the protonated state of P1. This interaction therefore increases the real pKa value of P1 and decreases its nucleophilicity. In such a situation, the nucleophilic attack at C1' by P1 is significantly compromised. On the contrary, the catalytic activation of P1 when Fpg binds FapydG in the rotamer II can be related to a key water molecule, w44 (Figure 4). Indeed, the strictly conserved E2 and E5 residues of Fpg collaborate with the  $N^2$  and  $N^3$  of the flipped-out nucleoside to exquisitely position w44 (Figure 7B). Thus, E2 can activate w44 which plays the role of general acid-base catalyst to deprotonate P1. E5 and the damaged nucleoside itself can contribute to drive the activated w44 to its target, the ammonium proton of P1 (Figure 7B). In such a model, the flipped-out damaged nucleoside acts as a cocatalyst as it has been already proposed for the human Ogg1 (Fromme et al., 2003). Such a scenario is supported by the fact that the E2Q mutant of Fpg is completely defective in the DNA glycosylase activity and that the E5Q mutant is significantly less efficient than the wild-type Fpg (Lavrukhin and Lloyd, 2000). In addition, the crystal structure of the abortive complex E2Q-BstFpg/8-oxodG-DNA indicates that the replacement of E2 by Q results in the absence of w44 in the binding site of the enzyme (Figure 7A). The interaction E2-w44 is also observed inside the binding pocket of the wild-type LIFpg bound to abasic site-containing DNA which is in agreement with the fact that the AP lyase activity of the enzyme also depends on the ability of P1 to attack the C1' of the AP site. The absence of w44 inside the binding pocket of E2Q-BstFpg bound to 8-oxodG-DNA suggests that the same observation in the crystal structure of wt-LIFpg bound to Bz-FapydG-DNA is also associated with the inability of the enzyme to process bulky Fapy-DNA adducts as demonstrated for AFB1-FapydG (Alekseyev et al., 2004).

Consequently, bulky Fapy-DNA adducts such as the Bz-FapydG may be *in vivo* molecular traps leading Fpg into unproductive complexes (Figure 7A). By the same argument, the Fpg eukaryote homolog, Ogg1, may also be trapped by bulky Fapy-derivatives because we show that this enzyme also recognizes FapydG and Bz-FapydG with similar efficiency (Figure 2). The formation of such unproductive complexes would hide lesions, which strongly impairs the ability of other DNA repair systems to approach the damaged site. This hypothesis would be supported by the observation that in contrast to its cationic precursor AFB1-dG, the AFB1-FapydG-DNA adduct is highly persistent in bacteria and cells treated with aflatoxin B1 (AFB1), a toxin produced by the common soil fungus *Aspergillus flavus* (reviewed in Wang and Groopman, 1999). Exposure to this toxin is significant in regions of the world where certain foods are improperly stored and it is considered to be one of the most potent known liver carcinogens (Smela et al., 2001). It has been clearly shown that AFB1-dG and AFB1-FapydG lesions are preferentially repaired by the nucleotide excision repair pathway (NER) in bacteria and mammalian cells and,



**Figure 4. Interactions between the Extrahelical Fapy-Derivatives and Fpg inside the Active Site Binding Pocket**

(A and B) Stereoviews of the elaborated interaction network between Fpg and cFapydG (in yellow) and Bz-cFapydG (in green). The C<sup>α</sup> backbone of Fpg is in gray and residue side chains in pink. Covalent links in protein and DNA are drawn by ball-and-stick representation. N, O, S, and P atoms are represented by blue, red, yellow, and magenta balls, respectively. Inferred hydrogen bonds are shown as red dashed lines and water molecules mediating interactions as small red spheres. For both structures, the distance between the ammonium group of P1 and the C1' of the target nucleoside is indicated and depicted by a light-blue dashed line.



**Figure 5. "Open" and "Closed" Conformation of the Very Flexible Part of the Fpg  $\alpha$ F- $\beta$ 9 Loop Involved in the Damaged Base Recognition**

(A) Superposition of the backbone of the  $\alpha$ F- $\beta$ 9 loop (from L214-K231) of the wild-type LIFpg bound to DNA duplex containing either an abasic site analog (1,3-propanediol and tetrahydrofuran in light-brown and brown, respectively; Pereira de Jesús et al., 2005), cFapydG (in green; this work) Bz-cFapydG (in blue; this work), and of the defective mutant  $\Delta$ P1-LIFpg bound to cFapydG (in pink; Coste et al., 2004). The strictly conserved glycine residues (G216, G226, and G229) which act as a hinge in the "closed" and "open" configurations of the loop are indicated in red. Residues labeled by a red sphere are involved by their main chains in the recognition of the  $O^6$  of the damaged guanine and are often missing in the electron density map of the crystal model. Residues labeled by a blue sphere belong to a part of the loop which presents different folds in the "open" and "closed" conformations of the loop.

(B) Sequence alignment of the flexible part of the  $\alpha$ F- $\beta$ 9 loop from *L. lactis* Fpg (LIFpg), *B. stearothermophilus* (BstFpg), and *T. thermophilus* (TtFpg). Identical residues are colored in bold red, similar residues in bold black and nonconserved residues in black. The strictly conserved glycine residues are shown in white on a red background. Labeling by red and blue spheres is defined in (A).

(C) "Closed" conformation of the  $\alpha$ F- $\beta$ 9 loop. Here is represented a superposition of the E200-H228 segment of the loop of free TtFpg (in pink; Sugahara et al., 2000), the E213-L239 segment of the

defective mutant E2Q-BstFpg bound to 8-oxodG-DNA duplex (in gray; Fromme and Verdine, 2003), and the L210-K235 segment of the wild-type LIFpg protein bound to cFapydG-DNA duplex (in green; this work).

(D) Stereoviews of the superposition of the loops of E2Q-BstFpg bound to 8-oxodG-DNA (gray) and wild-type LIFpg bound to cFapydG-DNA (green) showing the interactions of the loop residues with oxidized guanines (see also Figure 4). Residues indicated are those of the wild-type LIFpg (this work). Labeling by red and blue spheres is defined in (A).

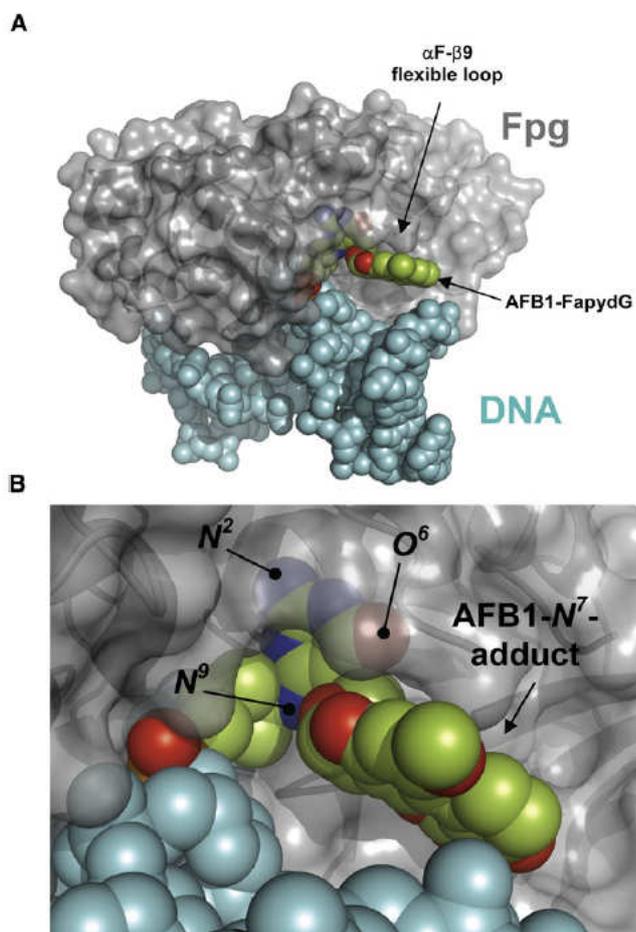
that the contribution of BER enzymes for repair of such lesions in bacteria is negligible (Alekseyev et al., 2004). Based on NMR studies, it was postulated that the enhanced stabilization of DNA duplexes harboring AFB1-FapydG adduct as compared to that of AFB1-dG could be one factor in its ability to escape repair (Mao et al., 1998). However, this explanation is questionable because repair of both lesions by UvrABC endonuclease (bacterial NER) occurs with similar efficiency in vitro (Oleykowski et al., 1993). This suggests that in vivo, some cellular factors can negatively interfere with the repair efficiency of AFB1-FapydG by NER. Based on our present data, we propose that BER enzymes can prevent the efficient repair of bulky FapydG-DNA adducts by NER. We surmise that this repair inhibition would be connected to the ability of BER enzymes such as Fpg and Ogg1 to form stable unproductive complexes with these bulky lesions. Such BER interference is not observed for

$N^7$ -Alkyl-dG, such as the cationic  $N^7$ -Me-dG, because these lesions are not recognized and processed by Fpg (or Ogg1) (Laval et al., 1991).

## SIGNIFICANCE

DNA glycosylases are key enzymes in the cellular defense against the genotoxic (mutagen or lethal) effects of DNA base damage. They initiate the base excision repair pathway by removing a damaged base from the genome. Considerable interest surrounds the molecular mechanism by which these enzymes search, recognize, and excise one base lesion among a vast excess of normal bases. Numerous structural and functional insights about the recognition of oxidized purines (8-oxoguanine and FapyG) by Fpg and hOgg1 have been provided by coupling the synthesis

(C) Direct interactions and water molecule-mediated interactions. Direct interactions between hydrogen bond donors (D) and acceptors (A) established between enzyme residues (in green) and functional groups of the FapydG (F, in red) and Bz-FapydG (BzF, in magenta) are indicated above the dashed line of the table. Water molecules which mediate interactions between the enzyme and the damaged base are indicated in blue below the dashed line of the table. Interactions which are common between the structures of Fpg bound to cFapydG- and Bz-cFapydG are highlighted in gray. The underlined water molecule interacts only with the base moiety of the damage. See also Figure 5 to visualize each interaction in space.



**Figure 6. Proposed Model for the Recognition of the AFB1-FapydG DNA Adduct by Fpg**

(A) Overview of the model and (B) enlargement at the level of the extrahelical damaged nucleoside. The graphic representation and the orientation in (B) of the model are similar to those used in Figure 3A. The benzyl group of Bz-cFapydG in the crystal structure of wt-LIFpg bound to Bz-cFapydG-DNA is replaced by the AFB1 group. Rotation around the C5-N7-C8A-O7 dihedral angle leads to a position for which no steric clashes were observed between the AFB1 adduct and the protein. In such a position, the AFB1 moiety points toward the solvent outside the Fpg binding pocket. The structure of the AFB1 adduct we used to construct the model is that observed in the solution structure of the AFB1-FapydG-DNA (1hm1; Mao et al., 1998).

of damaged nucleobase-containing oligonucleotides with high-resolution crystallographic structures of stable enzyme/DNA complexes. Structural information on the recognition of bulky-base lesion by these enzymes has, however, been lacking to date.

Here, we report the use of synthetic carbocyclic Fapy-derivatives-containing oligonucleotides to solve the X-ray structures of Fpg bound to DNA containing either FapydG or bulky  $N^7$ -benzyl-FapydG (Bz-FapydG). These structures lend key support to the ability of Fpg (or Ogg1) to accommodate bulky Fapy-DNA adducts inside its binding pocket. As already shown for DNA synthesis, the present study highlights the importance of the ability of Fapy lesions to adopt different stable conformations. However and differently

from FapydG, the rotamer IV selection of Bz-FapydG by Fpg observed in the structure forces the enzyme to form a stable but unproductive complex and provides new insights about reactions catalyzed by Fpg.

If we consider Bz-FapydG as a model for other bulky-Fapy-DNA adducts, we propose that in vivo, such unproductive BER enzyme-DNA complexes could negatively interfere with the NER machinery for the efficient repair of such damage. This interference could be responsible for the long persistence of these lesions in organisms, which is connected with a real cancer risk. Further investigations are however needed to extrapolate our findings to larger bulky Fapy adducts especially those which intercalate in DNA.

#### EXPERIMENTAL PROCEDURES

##### Proteins and DNA Duplexes

The open reading frame of the wild-type LIFpg is isolated by PCR using the plasmid pVE1064 as a template (Duwat et al., 1995) and the oligonucleotide primers GGGACATATGCCTGAATTACCAGAAGT and TTTTCCCGGTTA TTTTGGCT-GACAAACTGGGCAAAAATGAG. After NdeI/XmaI digestion of the PCR DNA product, the resulting DNA fragment was inserted by ligation into the linear NdeI/XmaI pFlag-CTC expression vector (Sigma-Aldrich). After sequence checking, the recombinant plasmid (pFlag-WT-LIFpg) was used to transform the Fpg defective *E. coli* strain BH540 (Castaing et al., 1993). The WT LIFpg protein was then overproduced in the recombinant strain BH540/pFlag-WT-LIFpg. The homogeneous WT LIFpg protein was purified as previously described with the following modifications (Pereira de Jesus et al., 2002). After lyses, QMA (Waters) and SP Sepharose FF (Pharmacia) purification steps, the SP active pool was fractionated by AcA54 gel filtration chromatography. The AcA54 pool was then dialyzed and loaded onto an HS POROS cation exchanger (PerSeptive Biosystem) and eluted by a linear NaCl gradient. The homogeneous WT LIFpg protein eluted between 250 and 350 mM NaCl was concentrated and stored at  $-80^{\circ}\text{C}$  after mass spectrometry and N-terminal micro-sequencing checking. The yeast 8-oxoguanine-DNA glycosylase (yOgg1) was donated by Serge Boiteux.

The unmodified single-strand oligonucleotides were purchased from Eurogentec and purified by MonoQ anion exchanger chromatography as previously described (Pereira de Jesus et al., 2002). The 14-mer oligonucleotide containing the cFapydG residue CTCTTT(cFapydG)TTTCTCG (SUP1) was synthesized and purified as previously described (Ober et al., 2003). For synthesis data of  $N^7$ -Benzyl-cFapydG phosphoramidite and oligonucleotide (SUP2), see the Supplemental Experimental Procedures.

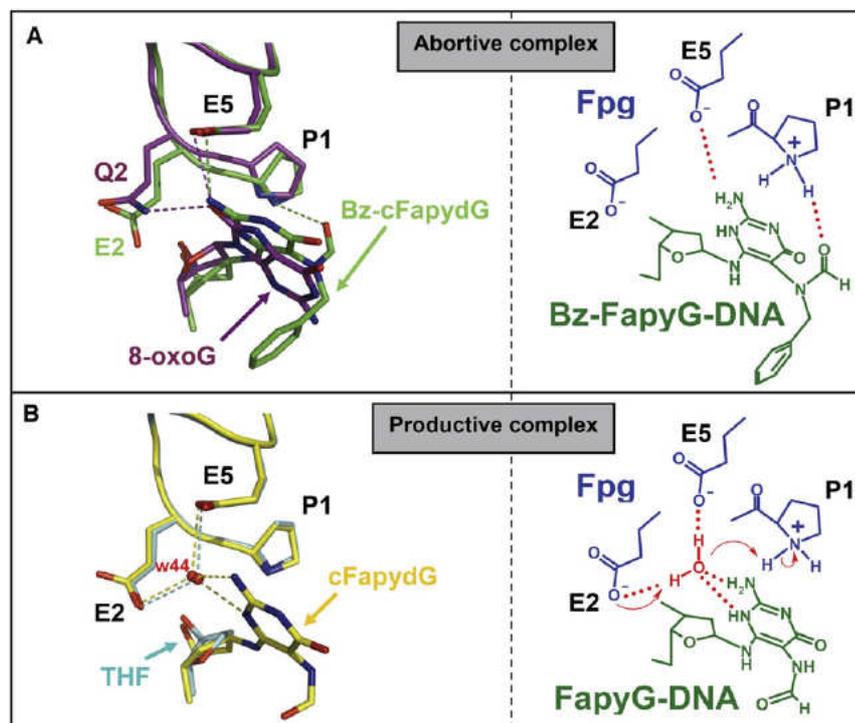
The purified strands SUP1 and SUP2 were then annealed with the complementary strand INF to generate 14-mer DNA duplexes containing either cFapydG or  $N^7$ -Benzyl-cFapydG opposite cytosine. For binding experiments, modified strands SUP1 and SUP2 were  $5'$ - $^{32}\text{P}$ -labeled before annealing as previously described (Serre et al., 2002).

##### Apparent Dissociation Constants of LIFpg and yOgg1 Bound to FapydG Derivative-Containing DNA

Apparent dissociation constants ( $K_{\text{Dapp}}$ ) of both DNA glycosylases (Fpg and Ogg1) for 14-mer DNA duplexes containing cFapydG-derivatives were determined by electrophoresis mobility shift assays. The same experimental conditions were used for both enzymes. Briefly,  $5'$ -radiolabeled DNA duplex (200 pM) was incubated for 20 min at  $4^{\circ}\text{C}$  with or without enzyme in 10  $\mu\text{l}$  of 30 mM HEPES/NaOH (pH 7.6), 140 mM NaCl, 2.4 mM  $\text{Na}_2\text{EDTA}$ , 2.4 mM  $\beta$ -mercaptoethanol, 10% glycerol, and 3% BSA. Assays were then analyzed by EMSA and quantified as previously described (Castaing et al., 1999).

##### Crystallization, X-Ray Data Collection, and Structure Determination

Crystallization of the two complexes was performed as described elsewhere (Coste et al., 2004). X-ray diffraction data were collected at 100K at the European Synchrotron Radiation Facility. The diffraction images were processed with MOSFLM (Leslie et al., 1996) and scaled with SCALA from the CCP4 package. The two structures were solved by molecular replacement with the



**Figure 7. The Rotamer IV of Fapy-Derivatives: A Molecular Trap for Fpg**

(A) Fpg/DNA abortive complex. Left panel: superimposition of E2Q-BstFpg bound to 8-oxodG-DNA (purple) and wt-LIFpg bound to Bz-cFapydG-DNA (green). Right panel: schematic representation of wt-LIFpg bound to Bz-FapydG-DNA showing the stabilization of the P1 ammonium by the formyl group of the damaged nucleoside.

(B) Fpg/DNA productive complex. (Left) Superimposition of crystal models of wt-LIFpg bound to either tetrahydrofuran (THF)-containing DNA (blue) or cFapydG-DNA (yellow). THF is an abasic site analog. (Right) Schematic representation of wt-LIFpg bound to FapydG-DNA showing the activation of the catalytic P1 by the E2 residue through the mediation of the water molecule w44 which plays the role of a general acid-base catalyst. In such a mechanism, the E5 residue and the extrahelical damage collaborate to guide the activated w44 toward the P1 ammonium target (right).

AMoRe program (Navaza, 1994) using the Fpg complex with the Protein Data Bank accession code 1TDZ minus the solvent molecules as a search model. The Dundee PRODRG2 server (van Aalten et al., 1996) was used to generate molecular topology files of the Bz-cFapydG nucleobase. Several cycles of manual model building performed in TURBO-FRODO (Roussel and Cambillau, 1991) and maximum-likelihood refinement performed in Refmac5 (Murshudov et al., 1997) yielded the final models. During the last refinement steps, anisotropic B factor parameters were introduced for the Zn atom. The final model of wt-LIFpg/cFapydG-DNA contains all of the 271 protein residues, 28 nucleotide residues, one glycerol molecule, one zinc ion, and a total of 397 water molecules. Concerning the wt-LIFpg/Bz-cFapydG model, the residues 221 and 222 of the  $\alpha$ F- $\beta$ 9 loop were omitted because of insufficient density in the corresponding electron density maps. The model also contains 28 nucleotide residues, one glycerol molecule, one zinc ion, and a total of 335 water molecules.

Data collection and model refinement statistics are summarized in Figure S1. The crystallographic figures were prepared with the program PyMOL (DeLano, 2002).

#### ACCESSION NUMBERS

The atomic coordinates of the wt-LIFpg/cFapydG-DNA and wt-LIFpg/ $N^7$ -Benzyl-cFapydG-DNA complexes have been deposited in the Protein Data Bank under the accession codes 1XC8 and 3C58, respectively.

#### SUPPLEMENTAL DATA

Supplemental Data include Supplemental Experimental Procedures and one figure and can be found with this article online at <http://www.chembiol.com/cgi/content/full/15/7/706/DC1/>.

#### ACKNOWLEDGMENTS

This work was supported by the Association pour la recherche contre le cancer (ARC, France), the Commission Permanente du Conseil régional du Centre (Biotechnocentre, France), La ligue contre le cancer (Comity of Indre-et-Loire, France), and by Electricité de France (EDF). The work in Germany was sup-

ported by the Deutsche Forschungsgemeinschaft (SFB 646 and the Leibniz Program), as well as by Novartis. M.O. thanks the Boehringer Ingelheim Foundation for a predoctoral fellowship. M.A.I. thanks the Alexander von Humboldt Foundation for a postdoctoral fellowship. We are greatly indebted to the ESRF staff (Grenoble, France) for their help in diffraction data collection, and to F. Culard, S. Boiteux, V. Aucagne, and M. Cadène for critical discussion.

Received: February 16, 2008

Revised: May 14, 2008

Accepted: May 19, 2008

Published: July 18, 2008

#### REFERENCES

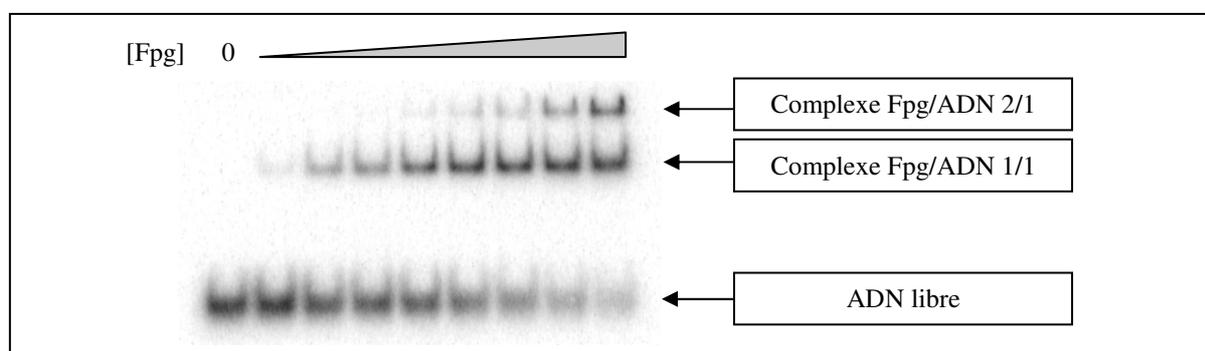
- Aguilar, F., Hussain, S.P., and Cerutti, P. (1993). Aflatoxin B1 induces the transversion of G  $\rightarrow$  T in codon 249 of the p53 tumor suppressor gene in human hepatocytes. *Proc. Natl. Acad. Sci. USA* 90, 8586–8590.
- Alekseyev, Y.O., Hamm, M.L., and Essigmann, J.M. (2004). Aflatoxin B1 formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway in vivo. *Carcinogenesis* 25, 1045–1051.
- Amara, P., and Serre, L. (2006). Functional flexibility of *Bacillus stearothermophilus* formamidopyrimidine DNA-glycosylase. *DNA Repair (Amst.)* 5, 947–958.
- Amara, P., Serre, L., Castaing, B., and Thomas, A. (2004). Insights into the DNA repair process by the formamidopyrimidine-DNA glycosylase investigated by molecular dynamics. *Protein Sci.* 13, 2009–2021.
- Asagoshi, K., Terato, H., Ohyama, Y., and Ide, H. (2002). Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: translesion DNA synthesis, nucleotide insertion, and extension kinetics. *J. Biol. Chem.* 277, 14589–14597.
- Beckman, K.B., and Ames, B.N. (1997). Oxidative decay of DNA. *J. Biol. Chem.* 272, 19633–19636.
- Boiteux, S., Belleney, J., Roques, B.P., and Laval, J. (1984). Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides. *Nucleic Acids Res.* 12, 5429–5439.

- Boiteux, S., Bichara, M., Fuchs, R.P., and Laval, J. (1989). Excision of the imidazole ring-opened form of N-2-aminofluorene-C(8)-guanine adduct in poly(dG-dC) by *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Carcinogenesis* 10, 1905–1909.
- Brown, K.L., Deng, J.Z., Iyer, R.S., Iyer, L.G., Voehler, M.W., Stone, M.P., Harris, C.M., and Harris, T.M. (2006). Unraveling the aflatoxin-FAPY conundrum: structural basis for differential replicative processing of isomeric forms of the formamidopyrimidine-type DNA adduct of aflatoxin B1. *J. Am. Chem. Soc.* 128, 15188–15199.
- Buchko, G.W., McAteer, K., Wallace, S.S., and Kennedy, M.A. (2005). Solution-state NMR investigation of DNA binding interactions in *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg): a dynamic description of the DNA/protein interface. *DNA Repair (Amst.)* 4, 327–339.
- Castaing, B., Geiger, A., Seliger, H., Nehls, P., Laval, J., Zelwer, C., and Boiteux, S. (1993). Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant FPG proteins. *Nucleic Acids Res.* 21, 2899–2905.
- Castaing, B., Fourrey, J.L., Hervouet, N., Thomas, M., Boiteux, S., and Zelwer, C. (1999). AP site structural determinants for Fpg specific recognition. *Nucleic Acids Res.* 27, 608–615.
- Chetsanga, C.J., and Lindahl, T. (1979). Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. *Nucleic Acids Res.* 6, 3673–3683.
- Chetsanga, C.J., Polidori, G., and Mainwaring, M. (1982). Analysis and excision of ring-opened phosphoramidate mustard-deoxyguanine adducts in DNA. *Cancer Res.* 42, 2616–2621.
- Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C., and Castaing, B. (2004). Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. *J. Biol. Chem.* 279, 44074–44083.
- DeLano, W.L. (2002). *The PyMOL User's Manual* (San Carlos, CA: DeLano Scientific).
- Douki, T., Martini, R., Ravanat, J.-L., Turesky, R.J., and Cadet, J. (1997). Measurement of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 8-oxo-7,8-dihydroguanine in isolated DNA exposed to gamma radiation in aqueous solution. *Carcinogenesis* 18, 2385–2391.
- Duwat, P., de Oliveira, R., Ehrlich, S.D., and Boiteux, S. (1995). Repair of oxidative DNA damage in gram-positive bacteria: the *Lactococcus lactis* Fpg protein. *Microbiology* 141, 411–417.
- Fromme, J.C., and Verdine, G.L. (2003). DNA lesion recognition by the bacterial repair enzyme MutM. *J. Biol. Chem.* 278, 51543–51548.
- Fromme, J.C., Bruner, S.D., Yang, W., Karplus, M., and Verdine, G.L. (2003). Product-assisted catalysis in base-excision DNA repair. *Nat. Struct. Biol.* 10, 204–211.
- Greenberg, M.M., Hantosi, Z., Wiederholt, C.J., and Rithner, C.D. (2001). Studies on N4-(2-deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy.dA) and N6-(2-deoxy-D-pentofuranosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy.dG). *Biochemistry* 40, 15856–15861.
- Haraguchi, K., Delaney, M.O., Wiederholt, C.J., Sambandam, A., Hantosi, Z., and Greenberg, M.M. (2001). Synthesis and characterization of oligonucleotides containing formamidopyrimidine lesions (Fapy.dA, Fapy.dG) at defined sites. *Nucleic Acids Res. (Suppl.)* 1, 129–130.
- Hertzog, P.J., Smith, J.R., and Garner, R.C. (1982). Production of monoclonal antibodies to guanine imidazole ring-opened aflatoxin B1DNA, the persistent DNA adduct in vivo. *Carcinogenesis* 3, 825–828.
- Hu, J., de Souza-Pinto, N.C., Haraguchi, K., Hogue, B.A., Jaruga, P., Greenberg, M.M., Dizdaroglu, M., and Bohr, V.A. (2005). Repair of formamidopyrimidines in DNA involves different glycosylases: role of the OGG1, NTH1, and NEIL1 enzymes. *J. Biol. Chem.* 280, 40544–40551.
- Humphreys, W.G., and Guengerich, F.P. (1991). Structure of formamidopyrimidine adducts as determined by NMR using specifically 15N-labeled guanine. *Chem. Res. Toxicol.* 4, 632–636.
- Kalam, M.A., Haraguchi, K., Chandani, S., Loechler, E.L., Moriya, M., Greenberg, M.M., and Basu, A.K. (2006). Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res.* 34, 2305–2315.
- Laval, J., Lopès, F., Madelmont, J.C., Godenèche, D., Meyniel, G., Habraken, Y., O'Connor, T.R., and Boiteux, S. (1991). Excision of imidazole ring-opened N7-hydroxyethylguanine from chloroethylnitrosourea-treated DNA by *Escherichia coli* formamidopyrimidine-DNA glycosylase. *IARC Sci. Publ.* 105, 412–416.
- Lavrukhin, O.V., and Lloyd, R.S. (2000). Involvement of phylogenetically conserved acidic amino acid residues in catalysis by an oxidative DNA damage enzyme formamidopyrimidine glycosylase. *Biochemistry* 39, 15266–15271.
- Leslie, A.G., Brick, P., and Wonacott, A.J. (1996). CCP4 Newslett. 18, 33–39.
- Lin, J.K., Miller, J.A., and Miller, E.C. (1977). 2,3-Dihydro-2-(guan-7-yl)-3-hydroxy-aflatoxin B1, a major acid hydrolysis product of aflatoxin B1-DNA or -ribosomal RNA adducts formed in hepatic microsome-mediated reactions and in rat liver in vivo. *Cancer Res.* 37, 4430–4438.
- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* 362, 709–715.
- Mao, H., Deng, Z., Wang, F., Harris, T.M., and Stone, M.P. (1998). An intercalated and thermally stable FAPY adduct of aflatoxin B1 in a DNA duplex: structural refinement from 1H NMR. *Biochemistry* 37, 4374–4387.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* 53D, 240–255.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallogr.* 50A, 157–163.
- Ober, M., Linne, U., Gierlich, J., and Carell, T. (2003). The two main DNA lesions 8-Oxo-7,8-dihydroguanine and 2,6-diamino-5-formamido-4-hydroxypyrimidine exhibit strongly different pairing properties. *Angew. Chem. Int. Ed.* 42, 4947–4951.
- Ober, M., Müller, H., Pieck, C., Gierlich, J., and Carell, T. (2005). Base pairing and replicative processing of the formamidopyrimidine-dG DNA lesion. *J. Am. Chem. Soc.* 127, 18143–18149.
- O'Connor, T.R., and Laval, J. (1989). Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites. *Proc. Natl. Acad. Sci. USA* 86, 5222–5226.
- O'Connor, T.R., Boiteux, S., and Laval, J. (1988). Ring-opened 7-methylguanine residues in DNA are a block to in vitro DNA synthesis. *Nucleic Acids Res.* 16, 5879–5894.
- Oleykowski, C.A., Mayernik, J.A., Lim, S.E., Groopman, J.D., Grossman, L., Wogan, G.N., and Yeung, A.T. (1993). Repair of aflatoxin B1 DNA adducts by the UvrABC endonuclease of *Escherichia coli*. *J. Biol. Chem.* 268, 7990–8002.
- Patro, J.N., Wiederholt, C.J., Jiang, Y.L., Delaney, J.C., Essigmann, J.M., and Greenberg, M.M. (2007). Studies on the replication of the ring opened formamidopyrimidine, Fapy.dG in *Escherichia coli*. *Biochemistry* 46, 10202–10212.
- Pereira de Jesus, K., Serre, L., Hervouet, N., Bouckson-Castaing, V., Zelwer, C., and Castaing, B. (2002). Crystallization and preliminary X-ray crystallographic studies of a complex between the *Lactococcus lactis* Fpg DNA-repair enzyme and an abasic site containing DNA. *Acta Crystallogr.* 58D, 679–682.
- Pereira de Jesus, K., Serre, L., Zelwer, C., and Castaing, B. (2005). Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. *Nucleic Acids Res.* 33, 5936–5944.
- Perlow-Poehnell, R.A., Zharkov, D.O., Grollman, A.P., and Broyde, S. (2004). Substrate discrimination by formamidopyrimidine-DNA glycosylase: distinguishing interactions within the active site. *Biochemistry* 43, 16092–16105.
- Roussel, A., and Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory 86* (Mountain View, CA: Silicon Graphics Corp.).
- Sedgwick, B., Bates, P.A., Paik, J., Jacobs, S.C., and Lindahl, T. (2007). Repair of alkylated DNA: recent advances. *DNA Repair (Amst.)* 6, 429–442.
- Serre, L., Pereira de Jesus, K., Boiteux, S., Zelwer, C., and Castaing, B. (2002). Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA

- glycosylase bound to an abasic site analogue-containing DNA. *EMBO J.* 21, 2854–2865.
- Smela, M.E., Currier, S.S., Bailey, E.A., and Essigmann, J.M. (2001). The chemistry and biology of aflatoxin B(1): from mutational spectrometry to carcinogenesis. *Carcinogenesis* 22, 535–545.
- Smela, M.E., Hamm, M.L., Henderson, P.T., Harris, C.M., Harris, T.M., and Essigmann, J.M. (2002). The aflatoxin B(1) formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA* 99, 6655–6660.
- Sugahara, M., Mikawa, T., Kuramitsu, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000). Crystal structure of a repair enzyme of oxidatively damaged DNA, MutM (Fpg), from an extreme thermophile, *Thermus thermophilus* HB8. *EMBO J.* 19, 3857–3869.
- Tchou, J., Bodepudi, V., Shibutani, S., Antoshchkin, I., Miller, J., Grollman, A.P., and Johnson, F. (1994). Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem.* 269, 15318–15324.
- Tomasz, M., Lipman, R., Lee, M.S., Verdine, G.L., and Nakanishi, K. (1987). Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. *Biochemistry* 26, 2010–2027.
- Tudek, B. (2003). Imidazole ring-opened DNA purines and their biological significance. *J. Biochem. Mol. Biol.* 36, 12–19.
- Tudek, B., van Zeeland, A.A., Kusmierek, J.T., and Laval, J. (1998). Activity of *Escherichia coli* DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. *Mutat. Res.* 407, 169–176.
- van Aalten, D.M.F., Bywater, R., Findlay, J.B.C., Hendlich, M., Hooft, R.W.W., and Vriend, G. (1996). PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided Mol. Des.* 10, 255–262.
- Wang, J.S., and Groopman, J.D. (1999). DNA damage by mycotoxins. *Mutat. Res.* 424, 167–181.
- Wilson, D.M., 3rd, and Bohr, V.A. (2007). The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair (Amst.)* 6, 544–559.
- Wood, R.D., Mitchell, M., and Lindahl, T. (2005). Human DNA repair genes. *Mutat. Res.* 577, 275–283.
- Zharkov, D.O., Rieger, R.A., Iden, C.R., and Grollman, A.P. (1997). NH<sub>2</sub>-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. *J. Biol. Chem.* 272, 5335–5341.

### II.1.1.2. Effet de la base opposée aux FapyGs sur la reconnaissance de ces lésions par LIFpg

La métabolisation des lésions par Fpg est connue pour être influencée par le contexte d'appariement de celles-ci (Tableau 7 de l'introduction). De plus, la fixation de Fpg sur des lésions est aussi influencée par leur contexte d'appariement (Castaing *et al.*, 1993; Tchou *et al.*, 1994; Wiederholt *et al.*, 2003; Zaika *et al.*, 2004). Ainsi, en plus des tests de fixation de Fpg sur les duplexes cFapyG/C et Bz-cFapyG/C, nous avons aussi réalisé des gel-retards sur le cFapyG et le Bz-cFapyG appariés à une adénine, une guanine ou une thymine (conditions standard, voir le « Matériel et méthodes »). La Figure 35 illustre ce type d'expérience.



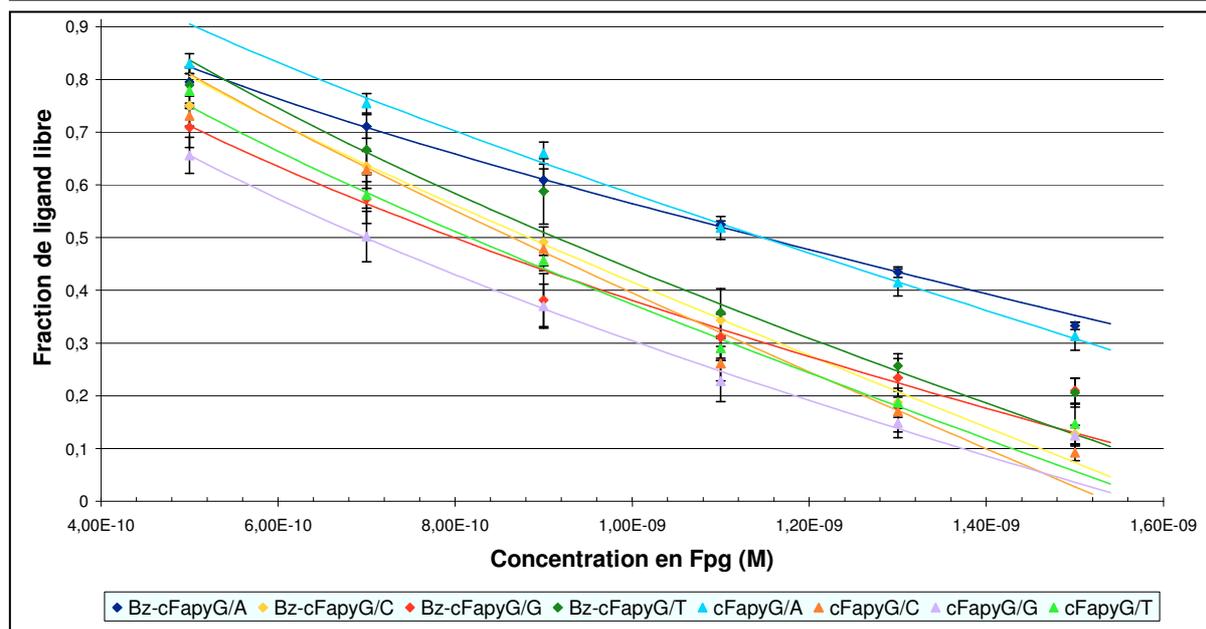
**Figure 35 :** Gel en conditions natives (gel-retard) correspondant au titrage de la fixation de Fpg sur un duplexe cFapyG/C.

Les concentrations en Fpg pour les pistes 2 à 9 sont de 0,1 ; 0,5 ; 0,7 ; 0,9 ; 1,1 ; 1,3 ; 1,5 et 3,2 nM, respectivement (voir détails dans le « Matériel et méthodes »).

Les bandes retardées par rapport à la migration de l'ADN libre sont attribuées à des complexes ADN/protéine (Figure 35). La quantité des complexes formés augmente bien avec la concentration de Fpg. En outre, plusieurs bandes correspondant à des complexes sont visibles, et peuvent être interprétées comme suit : l'ADN libre est d'abord titré dans un complexe de stoechiométrie 1/1 avec la protéine Fpg (bande la moins retardée, fixation spécifique d'une molécule de protéine sur un duplexe d'ADN lésé), puis dans un complexe de stoechiométrie 1/2 (bande la plus retardée, fixation non-spécifique d'une seconde molécule de protéine sur le complexe 1/1, probablement sur les extrémités non-lésées de l'ADN qui ne sont pas impliquées dans le premier complexe).

La quantification de l'intensité des bandes sur différentes expériences de gel retard nous donne les courbes de la Figure 36. A partir de ces données, la concentration en protéine Fpg nécessaire pour observer 50% de complexe correspond à une bonne estimation de la constante de dissociation apparente ( $K_{D\ app}$ ) des complexes considérés (voir le « Matériel et méthodes ») (Tableau 8).

**Figure 36 :** Evolution des fractions de ligands libres en fonction de la concentration en protéine Fpg pour les différents contextes d'appariement du cFapyG et du Bz-cFapyG.



**Tableau 8 :** Constantes de dissociations apparentes de Fpg pour les différents appariements possibles du cFapyG et du Bz-cFapyG.

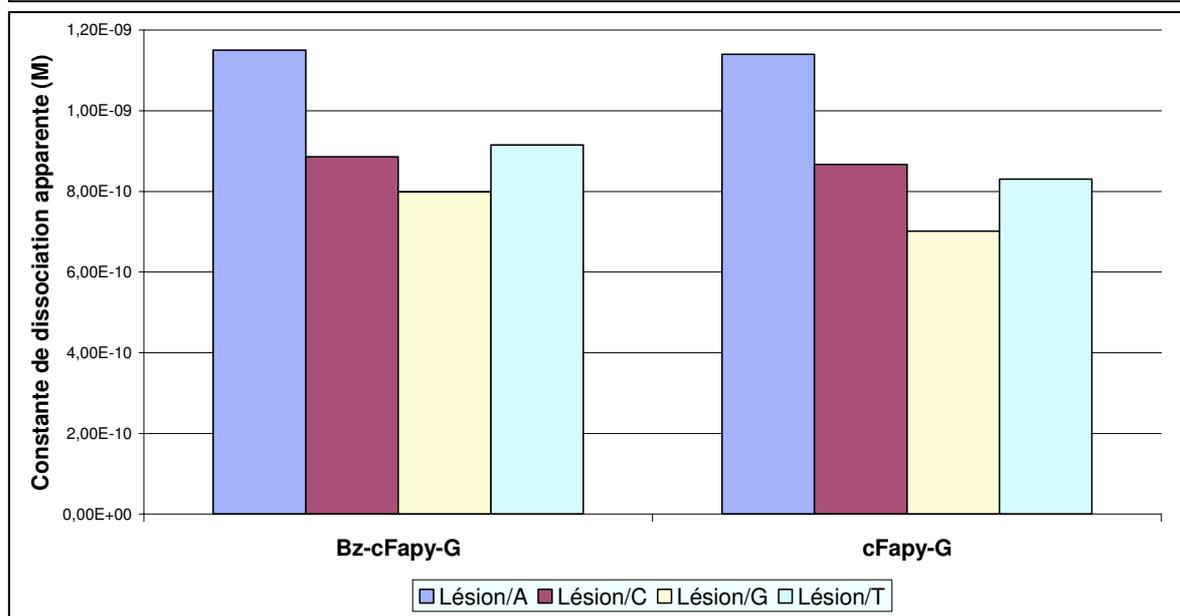
Sonde oligonucléotidique	cFapyG/C	cFapyG/T	cFapyG/G	cFapyG/A
<b>K<sub>D</sub> app (nM)</b>	0,87	0,83	0,70	1,14
Sonde oligonucléotidique	Bz-cFapyG/C	Bz-cFapyG/T	Bz-cFapyG/G	Bz-cFapyG/A
<b>K<sub>D</sub> app (nM)</b>	0,89	0,92	0,80	1,15

Quelle que soit la sonde d'ADN utilisée, les constantes de dissociation sont toutes de l'ordre du nanomolaire (Tableau 8). Cela nous conduit donc à la conclusion que le cFapyG et le Bz-cFapyG sont tous les deux des ligands de haute affinité pour Fpg, et cela quel que soit le contexte d'appariement de ces analogues de lésions. Ni la présence d'un groupement benzyle en N7 du FapyG, ni le contexte d'appariement, n'ont d'effet notable sur l'affinité de l'enzyme pour ces sondes d'ADN.

Une représentation sous forme d'histogramme des différentes valeurs de constantes de dissociations apparentes est présentée dans la Figure 37.

**Figure 37 : Comparaison des constantes de dissociation apparentes obtenues avec *LIFpg* sur cFapyG et le Bz-cFapyG.**

Les notations Lésion/A, Lésion/C, Lésion/G et Lésion/T font référence à la base opposée à la lésion, respectivement A, C, G et T.

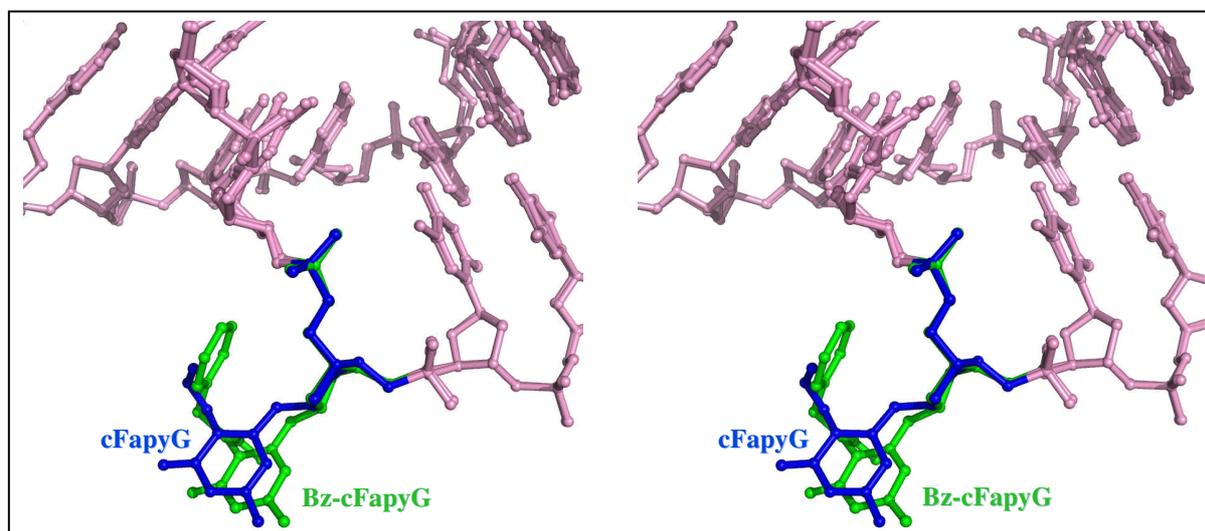


L'amplitude des différences dans les constantes de dissociation dues aux bases opposées aux lésions est légèrement plus importante avec le cFapyG qu'avec le Bz-cFapyG (Figure 37). En effet, la différence entre le meilleur et le plus mauvais ligand est d'environ 1,6 fois avec le cFapyG, et seulement d'environ 1,4 fois avec le Bz-cFapyG. Cet effet est maximal avec une guanine en face de la lésion, la valeur obtenue étant légèrement moins importante pour le cFapyG que pour le Bz-cFapyG. Cependant, cette différence étant très faible, on peut considérer que celle-ci n'est pas significative. De plus, les appariements FapyG/C et FapyG/A sont les plus importants à considérer, car ce sont les seuls qui aient été détectés lors des études de la fidélité des ADN polymérases sur cette lésion (Wiederholt *et al.*, 2002b). Aucune différence significative entre les fixations de Fpg sur le cFapyG et le Bz-cFapyG dans ces deux contextes d'appariement n'est observée, ce qui indique que la présence du groupement benzyle n'influence pas le comportement de fixation de Fpg quant à la spécificité de la base opposée à la lésion, l'adénine opposée à une lésion constituant dans chaque cas le moins bon ligand. Enfin, l'ordre de préférence de la base opposée à la lésion est légèrement différent pour les deux analogues de lésions. En allant du meilleur au plus mauvais ligand, cet ordre est  $G > T \geq C > A$  pour le cFapyG et  $G > C \geq T > A$  pour le Bz-cFapyG. Cependant, ces différences sont très ténues, et ne sont donc pas biologiquement significatives.

### II.1.1.3. Conclusion et discussion

Les valeurs de  $K_D$  de la publication (Coste *et al.*, 2008) indiquent que le cFapyG et le Bz-cFapyG sont tous deux des ligands de haute affinité pour Fpg, et un peu moins pour yOgg1. Sachant que les FapyGs avec des adduits encombrants en N7 ont souvent été décrits comme étant des substrats relativement mauvais pour Fpg (Chetsanga *et al.*, 1982; Tudek *et al.*, 1998; Alekseyev *et al.*, 2004), nous avons proposé que la difficulté à exciser les FapyGs substitués en N7 par des groupements encombrants n'était pas liée à la capacité de Fpg (ou yOgg1) à se fixer sur l'ADN lésé, mais plutôt à la formation d'un complexe catalytiquement non-productif avec ce type de dommage.

Afin de vérifier cette hypothèse, une étude structurale a été réalisée avec la Fpg de type sauvage de *L. lactis* complexée non covalamment à un duplexe cFapyG/C ou un duplexe Bz-cFapyG/C. L'analyse structurale indique que les complexes formés avec les deux analogues de substrats sont globalement très semblables (RMSD de 0,37 Å sur les atomes Ca de Fpg), avec des différences notables localisées au niveau des résidus 218 à 228 de la boucle  $\alpha$ F- $\beta$ 9 de Fpg, et au niveau de la conformation de la partie base nucléique des analogues de lésions. En l'occurrence, les analogues de lésions stabilisés dans la poche du site actif de Fpg ne sont pas coplanaires, et le Bz-cFapyG apparaît un peu moins extrahélicale que le cFapyG (Figure 38).



**Figure 38 :** Vue stéréoscopique d'une superposition des ADN contenant les analogues de lésions cFapyG ou Bz-cFapyG extrahélicaux dans les complexes L/Fpg/ADN.

Les structures sont celles des complexes L/Fpg/cFapyG et L/Fpg/Bz-cFapyG (codes PDB 1XC8 et 3C58, respectivement). L'ADN est coloré en rose, hormis pour les analogues de lésions, qui sont respectivement colorés en bleu et en vert pour le cFapyG et le Bz-cFapyG.

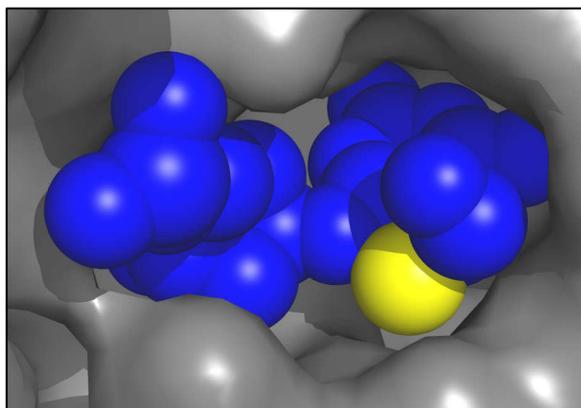
De plus, l'enzyme ne sélectionne pas le même rotamère du groupement formamido pour le cFapyG et le Bz-cFapyG (Figure 3 de Coste *et al.*, 2008), avec une rotation de presque 180° autour de la liaison C5-N7. Ce changement de rotamère permet à l'enzyme d'accommoder l'encombrement stérique du groupement benzyle en le repoussant à l'extérieur de la poche du site actif.

L'analyse structurale fine de ces modèles met en évidence que tous les groupements donneurs et accepteurs de la partie base nucléique du cFapyG sont contactés par des résidus de la protéine *via* des liaisons hydrogène, alors que ce n'est pas le cas pour le Bz-cFapyG (Figure 4 de Coste *et al.*, 2008). Les résidus de Fpg impliqués dans ces réseaux de liaisons hydrogène sont de plus différents dans les deux structures. Cela nous a conduit à émettre l'hypothèse que les différences dans le réseau de liaison hydrogène entre ces structures étaient favorables à la catalyse dans le cas du FapyG, et défavorable dans le cas du Bz-FapyG. Notamment, et uniquement dans le cas du Bz-cFapyG, des liaisons hydrogène sont établies entre l'azote du résidu P1 de l'enzyme et le groupement formamido de l'analogue de lésion. Cette interaction est susceptible de favoriser l'état protoné de P1, ce qui diminue le caractère nucléophile de cet azote, induisant donc vraisemblablement une moins bonne activité ADN glycosylase. Avec le Bz-cFapyG, tout se passe en fait comme si le pKa du résidu P1 de Fpg était augmenté. Dans la structure de Fpg complexée au cFapyG, l'azote de P1 interagit avec une molécule d'eau (w44 dans la structure, Figure 4 de Coste *et al.*, 2008), qui est elle-même en contact avec E2, un résidu catalytiquement important de la protéine. Ainsi, nous avons envisagé que le résidu E2 participe à la déprotonation du résidu P1 *via* cette molécule d'eau (voir le chapitre suivant), rendant donc le résidu P1 nucléophile, et favorisant ainsi l'activité ADN glycosylase de l'enzyme.

D'autre part, les études comparatives de la position de la boucle  $\alpha$ F- $\beta$ 9 dans nos structures et dans celles déjà disponibles pour Fpg nous ont conduit à proposer qu'il existait deux conformations majeures de celle-ci, une conformation « ouverte » et une conformation « fermée » (Figure 5 de Coste *et al.*, 2008). Nous proposons ici que la conformation « fermée » est nécessaire pour la reconnaissance correcte des guanines oxydées et pour l'excision efficace de ces lésions. Par contre, la conformation « fermée » de cette boucle ne semble pas nécessaire pour l'activité AP lyase de l'enzyme étant donné que toutes les structures de Fpg liée à des analogues de sites abasiques ont montré une conformation « ouverte » (Fromme *et al.*, 2002; Serre *et al.*, 2002; Pereira de Jésus *et al.*, 2005). L'observation que cette boucle adopte une conformation « ouverte » avec le Bz-cFapyG pour des raisons d'accommodation du

groupement benzyle constitue encore un bon argument pour dire que le Bz-FapyG est stabilisé dans la poche du site actif de Fpg dans une conformation catalytiquement non-productive.

Considérant que la structure avec le cFapyG représente probablement un état catalytiquement actif, et que celle avec le Bz-cFapyG un état catalytiquement inactif (ou plutôt très faiblement actif), des essais de reconstruction des adduits méthyle ou aflatoxine B<sub>1</sub> (AFB<sub>1</sub>) dans nos structures ont été réalisés. Le N<sup>7</sup>-AFB<sub>1</sub>-FapyG ne peut pas être reconstruit dans la structure de Fpg complexée au cFapyG sans provoquer des problèmes stériques avec les résidus tapissant la poche du site actif de Fpg. Par contre, cela est possible dans la structure de Fpg complexée au Bz-cFapyG (Figure 6 de Coste *et al.*, 2008). Cela constitue donc un bon argument pour dire que les FapyGs substitués en N7 par de gros adduits sont stabilisés dans la poche du site actif de Fpg suivant une conformation du même type que celle que l'on voit dans notre structure avec le Bz-cFapyG, expliquant ainsi que ces lésions soient de mauvais substrats pour Fpg (Chetsanga *et al.*, 1982; Chetsanga *et al.*, 1983; Tudek *et al.*, 1998; Alekseyev *et al.*, 2004). La reconstruction de l'adduit méthyle dans la structure de Fpg complexée au cFapyG n'est par contre pas associée à des problèmes stériques particuliers (Figure 39). On peut donc raisonnablement penser que le N<sup>7</sup>-Me-FapyG est stabilisé dans la poche du site actif de Fpg dans une conformation du même type que celle que l'on voit avec le cFapyG. Ceci explique pourquoi cette lésion est un très bon substrat pour l'enzyme (Asagoshi *et al.*, 2000).



**Figure 39 : Reconstruction de l'adduit méthyle en N7 du FapyG.**

La structure de base est celle de *LIFpg* complexée sur le duplexe cFapyG/C (code PDB 1XC8). La surface de la protéine est colorée en gris. L'analogue de lésion N<sup>7</sup>-Me-cFapyG est représenté par des sphères et coloré en bleu, hormis pour l'adduit méthyle en N7 qui est coloré en jaune. La taille des sphères est dictée par le rayon de covalence des atomes correspondants.

Finalement, la conclusion majeure de la publication est que le mode d'accommodation du FapyG dans la poche du site actif de Fpg est catalytiquement productif, et que celui du Bz-FapyG est catalytiquement non productif, expliquant ainsi l'efficacité réduite de Fpg sur les FapyG substitués en N7 par de gros adduits. En conséquence, la formation de complexes

abortifs stables entre Fpg et les FapyG substitués en N7 par de gros adduits pourrait poser plusieurs problèmes aux cellules. En effet, la reconnaissance de ce type de lésion par Fpg diminuerait sa disponibilité pour la réparation des lésions pour lesquelles elle est très efficace (8-oxo-G, FapyG, N<sup>7</sup>-Me-FapyG, oxazolone...), diminuant donc la capacité de réparation des lésions oxydatives dans ces cellules. D'autre part, la formation de complexes stables mais non productifs (ou peu productifs) entre Fpg et ce type de dommages implique que les enzymes du système NER, qui sont normalement sensées éliminer les lésions très encombrantes qui bloquent la réplication, soient gênées dans leur activité sur ce type de lésions. De plus, cette conclusion peut être élargie à l'homologue fonctionnel eucaryote Ogg1 de Fpg, et pourrait expliquer la persistance de telles lésions dans les organismes soumis à l'aflatoxine B<sub>1</sub> (Wang *et al.*, 1999).

Néanmoins, pour réellement conclure que les modes différents d'accommodation des lésions FapyG et Bz-FapyG dans la poche du site actif de Fpg correspondent à un mode de fixation catalytiquement productif pour le FapyG et improductif pour le Bz-FapyG, il aurait fallu faire des essais d'activité sur les formes excisables de ces lésions (greffées sur un désoxyribose) afin de prouver que le Bz-FapyG est réellement un plus mauvais substrat que le FapyG pour Fpg. En effet, le Bz-FapyG étant un modèle d'étude qui a été synthétisé uniquement sous forme carbocyclique pour nos études structurales, aucune donnée n'est bien évidemment disponible quant à l'efficacité de son excision par Fpg. Dans cette optique, la réalisation de tests d'activités sur le Bz-FapyG aurait donc aussi permis de vérifier que cette molécule était réellement un bon modèle d'étude pour la compréhension de la différence d'activité de Fpg sur les FapyG substitués en N7 par des adduits de taille variable. De plus, dans la publication nos conclusions ont été élargies à Ogg1, or nous n'avons fait que vérifier que la fixation de cette protéine sur nos deux analogues de lésions était du même ordre de grandeur. Pour réellement pouvoir faire les mêmes conclusions sur Ogg1 que sur Fpg, il faudrait aussi résoudre les structures de cette protéine en complexe avec nos analogues de lésions afin de vérifier que celle-ci utilise des modes de stabilisations différents pour le cFapyG et le Bz-cFapyG. En outre, tout comme pour Fpg, il faudrait aussi faire des essais d'activité de Ogg1 sur les formes excisables de ces lésions.

Par ailleurs, la base opposée au dommage affecte peu la fixation de Fpg sur le FapyG ou le Bz-FapyG. En effet, tous les duplexes testés sont des ligands de haute affinité pour Fpg (Tableau 8). Ainsi, ni la présence du groupement benzyle sur le FapyG, ni la nature de la base

opposée au dommage, n'induisent de modifications significatives de l'affinité de Fpg pour ces sondes oligonucléotidiques. On pourrait penser que la faible différence sur les constantes de dissociations obtenues avec le cFapyG et le Bz-cFapyG dans les différents contextes d'appariements pourrait impliquer qu'il n'y ait que peu de différence d'activité entre les différents appariements possibles des lésions correspondantes (greffées sur un désoxyribose). Cependant, pour réellement pouvoir inspecter ces différents aspects fonctionnels, il aurait fallu pouvoir tester l'activité de Fpg sur le FapyG et le Bz-FapyG (greffés sur des désoxyriboses) avec les différents appariements possibles. Par contre, on peut tout de même indiquer que l'adénine opposée aux lésions FapyG et Bz-FapyG constitue dans chaque cas le moins bon ligand pour Fpg (Tableau 8 et Figure 37). Ceci est conforme aux connaissances que l'on a sur l'efficacité de fixation de Fpg sur ses substrats, avec la règle du « tout sauf un A » (Castaing *et al.*, 1993; Tchou *et al.*, 1994; Wiederholt *et al.*, 2003; Zaika *et al.*, 2004). Cependant, l'amplitude des différences de constantes de dissociations entre nos différents duplexes d'ADN est plus réduite que celle observée pour les publications précédentes. La plus grande différence que nous ayons observée entre nos duplexes était de 1,6 fois pour le cFapyG apparié à une adénine ou une cytosine, alors que les études précédentes indiquaient une amplitude maximale des différences sur les  $K_D$  de 27 fois (Castaing *et al.*, 1993), de 38 fois (Tchou *et al.*, 1994), de 20 à 38 fois (Wiederholt *et al.*, 2003), et de 34 fois (Zaika *et al.*, 2004). Cependant, plusieurs différences notables entre notre étude et les précédentes sont à prendre en compte.

Premièrement, la plupart des études indiquées précédemment ont déterminé l'effet de l'appariement de la 8-oxo-G sur les  $K_D$  de Fpg pour ces ADN. Ainsi, on peut se demander si la nature de la lésion n'influerait pas sur les résultats, avec une différence dans les  $K_D$  plus importante avec la 8-oxo-G qu'avec le FapyG ou le Bz-FapyG en fonction de la base opposée au dommage. Ce comportement ne serait d'ailleurs pas si étonnant si on considère que le potentiel mutagène de la 8-oxo-G (lésion très mutagène, Michaels *et al.*, 1992b; Krahn *et al.*, 2003; Freisinger *et al.*, 2004; Hsu *et al.*, 2004) est bien supérieur à celui des FapyGs (lésions majoritairement bloquantes pour la réplication, Boiteux *et al.*, 1983; O'Connor *et al.*, 1988; Tudek *et al.*, 1992; Tudek *et al.*, 1999; Tudek, 2003; Asagoshi *et al.*, 2002; Wiederholt *et al.*, 2002b; Patro *et al.*, 2007). Ainsi, Fpg aurait subi une forte pression de sélection concernant la discrimination des différents contextes d'appariement de la 8-oxo-G pour contrecarrer sa forte mutagénicité, alors que ce ne serait pas le cas avec les FapyGs. Cette explication reste cependant peu probable, étant donné que la fixation de *EcFpg* sur la 8-oxo-G et le FapyG naturels opposés à un A ou un C a déjà été étudiée (Wiederholt *et al.*, 2003). Cette étude a montré des fortes différences entre les  $K_D$  obtenus en fonction du contexte d'appariement que

ce soit pour la 8-oxo-G ou pour le FapyG (différence de 20 fois et 38 fois entre les  $K_D$  obtenus avec un C ou un A en face du FapyG et de la 8-oxo-G, respectivement).

Une deuxième différence peut aussi avoir joué sur les résultats obtenus : les études précédentes (Castaing *et al.*, 1993; Tchou *et al.*, 1994; Wiederholt *et al.*, 2003; Zaika *et al.*, 2004) ont utilisé de la 8-oxo-G ou le FapyG sous formes naturelles, alors que nous avons utilisé des FapyGs greffés sur un cyclopentane. On sait que nos analogues carbocycliques de lésions ne peuvent être discernés des lésions naturelles pour ce qui est de la stabilité des paires de bases qu'elles forment, et ne perturbent pas la structure de l'ADN (Ober *et al.*, 2005). Cependant, on peut se demander si l'utilisation de tels analogues de lésions ne pourrait pas jouer sur l'amplitude des différences entre les constantes de dissociations obtenues avec Fpg en fonction de la base opposée au dommage. En effet, des différences de constantes de dissociations de Fpg pour une lésion ou son analogue non métabolisable ont déjà été observées (Wiederholt *et al.*, 2003). De plus, une étude du comportement de Fpg de *E. coli* sur un analogue du FapyA (où le N9 de la lésion a été remplacé par un  $CH_2$ ) a montré peu de différences entre les  $K_D$  obtenus en fonction de la base opposée au dommage (Wiederholt *et al.*, 2002a). Enfin, une étude de l'affinité de *EcFpg* pour un FapyG naturel opposé à un C ou un A a montré une différence de 20 fois entre ces constantes de fixation (Wiederholt *et al.*, 2003), ce qui correspond à une différence nettement supérieure à celle que l'on a observée dans notre étude. Tout ceci suggère qu'un analogue de lésion n'a pas forcément les mêmes propriétés que la lésion naturelle correspondante.

Une troisième différence notable entre notre étude et les précédentes concerne l'origine de la Fpg utilisée : *L. lactis* pour notre étude, et *E. coli* pour les trois précédentes (Castaing *et al.*, 1993; Tchou *et al.*, 1994; Zaika *et al.*, 2004). On peut ainsi se demander si *EcFpg* ne serait pas plus discriminative que *LFpg* pour la base opposée au dommage. Ainsi, l'effet mutagène des lésions est peut être plus important avec les ADN polymérases de *E. coli* que celles de *L. lactis*, d'où la sélection d'une Fpg très stricte concernant la discrimination du contexte d'appariement des lésions chez *E. coli*, et pas chez *L. lactis*. Il est d'ailleurs important de rappeler ici que des différences de comportement pour une même ADN polymérase provenant de deux espèces bactériennes ont déjà été observées (Lowe *et al.*, 1996; Hsu *et al.*, 2004).

Enfin, d'autres différences entre notre étude et les précédentes peuvent aussi avoir joué sur les résultats. En effet, la séquence dans laquelle la lésion est insérée est un paramètre qui est susceptible de moduler les propriétés de Fpg, comme cela a déjà été vu pour d'autres études (Graves *et al.*, 1992b; Asagoshi *et al.*, 2000). D'autre part, les conditions d'analyse représentent aussi un paramètre susceptible de jouer sur les propriétés de Fpg, donc pouvant

aussi expliquer en partie les différences entre notre étude et les précédentes (tampons utilisés, température...). Ainsi, il a été montré que la force ionique du milieu et sa concentration en ions  $Mg^{2+}$  influent sur la spécificité de Fpg et Ogg1 pour la base opposée à la lésion 8-oxo-G (Sidorenko *et al.*, 2008).

Pour pouvoir réellement conclure sur les points précédents, il faudrait refaire les expériences d'interaction en fonction de la base opposée au dommage, mais en comparant les Fpg de *L. lactis* et de *E. coli* dans les mêmes conditions expérimentales, sur la 8-oxo-G et le FapyG, en série naturelles ou greffées sur un cyclopentane, avec le même contexte de séquence pour toutes les lésions.

D'autre part, les tests de fixation de  $\gamma$ Ogg1 sur les duplexes G/C, cFapyG/C et Bz-cFapyG/C ont montré assez peu de différences. En effet, l'enzyme  $\gamma$ Ogg1 montre des constantes de dissociations de l'ordre d'une centaine de nanomolaires pour le cFapyG et le Bz-cFapyG (Figure 2 de Coste *et al.*, 2008), indiquant qu'ils sont tous deux des ligands relativement affins pour  $\gamma$ Ogg1. Cependant, nous avons aussi déterminé ce paramètre pour un duplexe G/C (sans lésion), et nous avons obtenu une valeur de 224 nM (données non présentées dans la publication de Coste *et al.*, 2008). La constante de dissociation apparente de  $\gamma$ Ogg1 pour un ADN sans lésion est donc seulement deux fois plus importante que pour un ADN lésé, ce qui indique que l'efficacité de fixation n'est pas franchement un point crucial pour l'efficacité de l'activité ADN glycosylase de Ogg1. En effet, on sait que l'excision d'un G normal par cette enzyme est une réaction inefficace voire inexistante, alors que l'excision par Ogg1 du  $N^7$ -Me-FapyG est efficace (Asagoshi *et al.*, 2000). Dans ce cadre, nos données indiquant que la fixation de Ogg1 sur le cFapyG et le Bz-cFapyG sont similaires ne prouvent aucunement que le FapyG et le Bz-FapyG sont aussi bien excisées par Ogg1. Au contraire, le comportement de Ogg1 sur l'ADN lésé ou non lésé aurait plutôt tendance à indiquer qu'une faible différence d'efficacité de fixation peut être associée à une forte différence d'efficacité de réparation. Finalement, en l'absence de données d'activité comparatives sur le FapyG et le Bz-FapyG (greffés sur un désoxyribose), les données de fixation précédentes ne sont que très peu informatives car elles ne donnent aucune information sur l'efficacité de l'activité de Ogg1 sur de telles lésions.

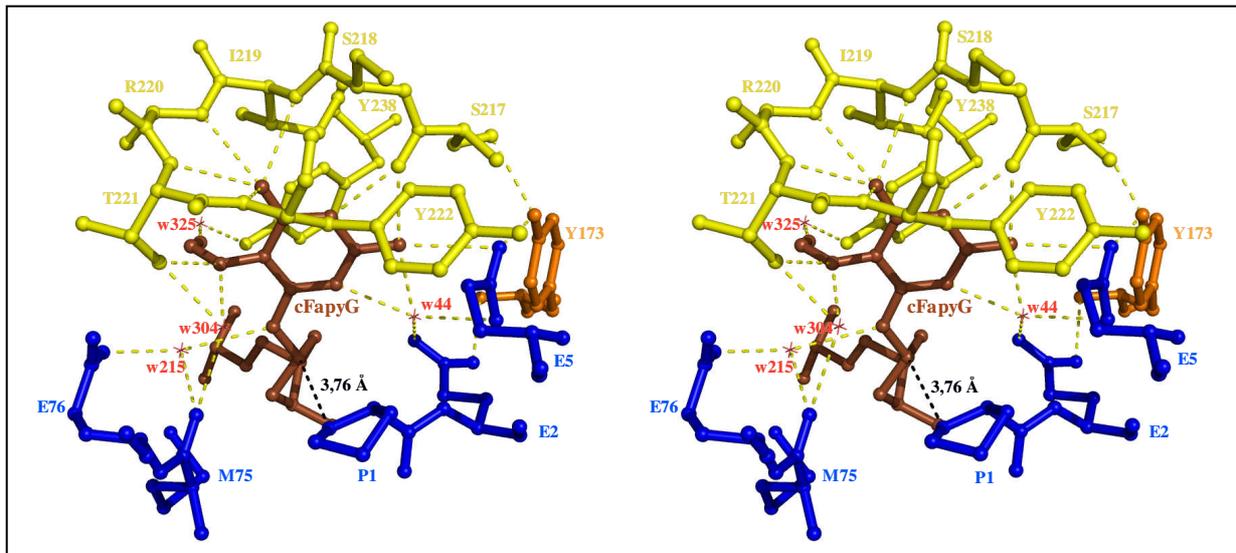
Finalement, comme illustré dans le paragraphe précédent avec Ogg1, il est important de rappeler que les faibles différences d'efficacité de fixation de Fpg et Ogg1 sur toutes nos

sondes oligonucléotidiques n'impliquent pas que les différences d'activité ADN glycosylase sur les lésions naturelles correspondantes soient faibles.

### II.1.2. Eléments fonctionnels relatifs à la reconnaissance et à l'excision de la 8-oxoguanine par la protéine Fpg de *L. lactis*

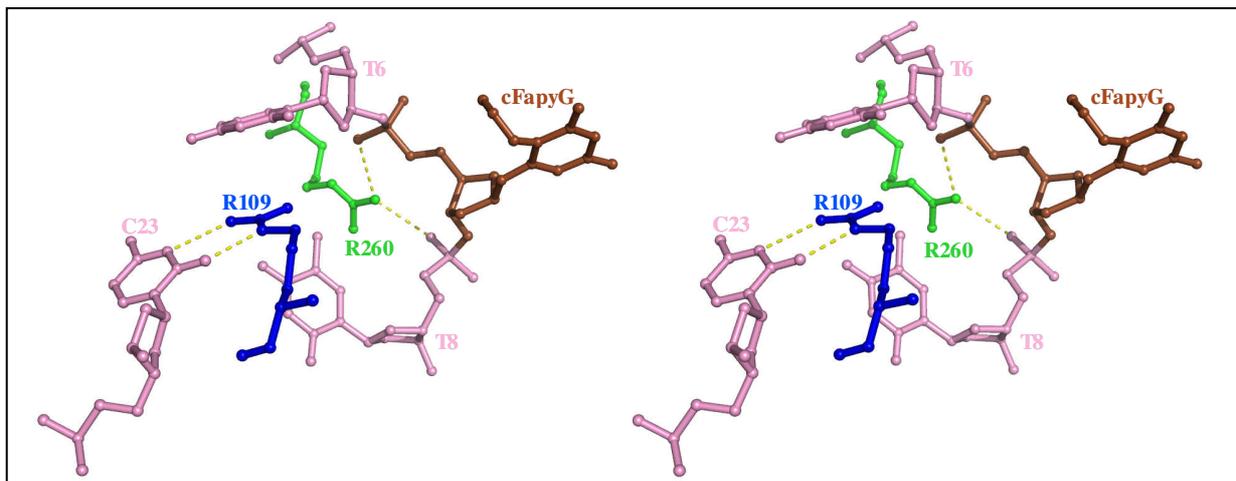
Les structures cristallographiques précédemment obtenues pour la Fpg de *L. lactis* (*LlFpg*) complexée à des ADN contenant un cFapyG ou un Bz-cFapyG ont permis de mettre en évidence les déterminants structuraux et/ou fonctionnels de l'enzyme indispensables à la reconnaissance de ces dommages dans l'ADN. Comme cela avait déjà été montré lors d'études structurales sur les Fpgs de *E. coli* (*EcFpg*) et de *B. stearothermophilus* (*BstFpg*), *LlFpg* expose la base lésée aux résidus catalytiques après l'avoir extrudée de la double hélice d'ADN et stabilisée en position extrahélicale dans la poche de son site actif (chapitre I.7.1 de l'introduction concernant les généralités sur la protéine Fpg). Sur la base de ces données structurales, nous avons décidé de réaliser une mutagenèse dirigée en ciblant les résidus amino-acides qui tapissent la poche du site actif de *LlFpg*. Ces résidus sont en interaction directe avec la base extrudée, avec des résidus qui sont eux-mêmes en interaction directe avec la base extrudée (Figure 40), avec les phosphates du squelette phosphodiester de l'ADN au voisinage de la lésion, ou avec la base opposée à la lésion (Figure 41).

Les mutations ponctuelles introduites dans la séquence du gène *fpg* de *L. lactis* conduisent aux substitutions suivantes sur *LlFpg* : P1G, E2N, E2R, E2T, E2L, E2M, E2D, E2Q, E5Q, M75C, M75L, E76Y, R109I, R109S, Y173F, S217A, S217Q, I219D, T221A, Y222A, R260I et R260L (voir structure primaire de Fpg, Figure 22 de l'introduction). La grande majorité des résidus de *LlFpg* en interaction directe avec le cFapyG ont été ciblés, de même que les résidus dont la position est équivalente chez *LlFpg* à ceux qui sont en interaction directe avec la 8-oxo-G dans la structure de *BstFpg* mutante E2Q complexée à une ADN contenant une 8-oxo-G (code PDB 1R2Y, Fromme et al., 2003a). Afin de simplifier le texte, le mot « mutant » sera employé pour désigner les protéines Fpg ayant subi une substitution d'un résidu amino-acide, et non pas pour désigner les bactéries de type mutant.



**Figure 40 :** Vue stéréoscopique des résidus de *LIFpg* impliqués dans la stabilisation de la partie base nucléique du cFapyG.

Issu de la structure de *LIFpg* complexée à un duplexe cFapyG/C (code PDB 1XC8, Coste *et al.*, 2008). Les résidus de la protéine sont soumis au même code couleur que pour les Figures 23 et 24 de l'introduction du manuscrit, suivant la définition de Sugahara *et al.*, 2000. Le cFapyG et les molécules d'eau sont représentés en marron et en rouge, respectivement. Les liaisons hydrogène sont représentées par des pointillés jaunes. La distance entre l'azote du résidu P1 et le C1' du cFapyG est indiquée en noir, et est localisée par un pointillé noir.

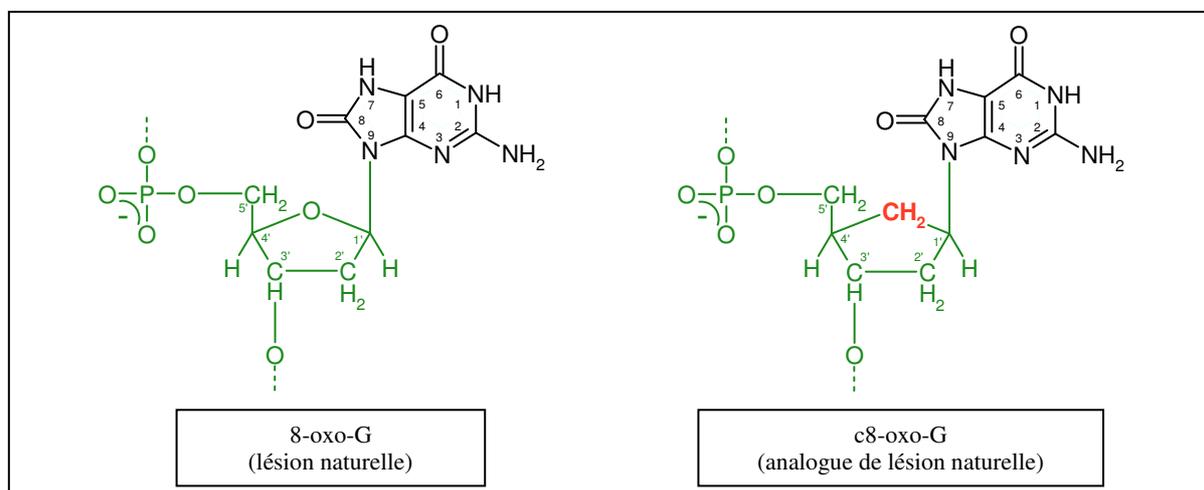


**Figure 41 :** Vue stéréoscopique des résidus impliqués dans la stabilisation des deux phosphates entourant la lésion cFapyG et dans la stabilisation de la cytosine opposée à la lésion.

Issu de la structure de *LIFpg* complexée à un duplexe cFapyG/C (code PDB 1XC8, Coste *et al.*, 2008). Les résidus de la protéine sont soumis au même code couleur que pour les Figures 23 et 24 de l'introduction du manuscrit, suivant la définition de Sugahara *et al.*, 2000. L'ADN est coloré en rose, hormis le cFapyG qui est coloré en marron. Les liaisons hydrogène sont représentées par des pointillés jaunes.

Dans cette étude, nous nous sommes intéressés à la caractérisation de nos différents mutants protéiques du point de vue de leur fixation à un ADN contenant un analogue de carbocyclique de la 8-oxo-G (c8-oxo-G), et du point de vue de leur capacité à exciser une 8-oxo-G naturelle (Figure 42).

Afin d'éviter d'alourdir le texte, j'ai choisi de vous décrire en détail uniquement les résultats obtenus avec les mutants de *LIFpg* en position 2, car nous avons étendu notre analyse jusqu'à la caractérisation des protéines E2Q et E2D purifiées. Les résultats obtenus avec les autres mutants protéiques seront uniquement présentés dans le texte et résumés à la fin du sous-chapitre suivant.



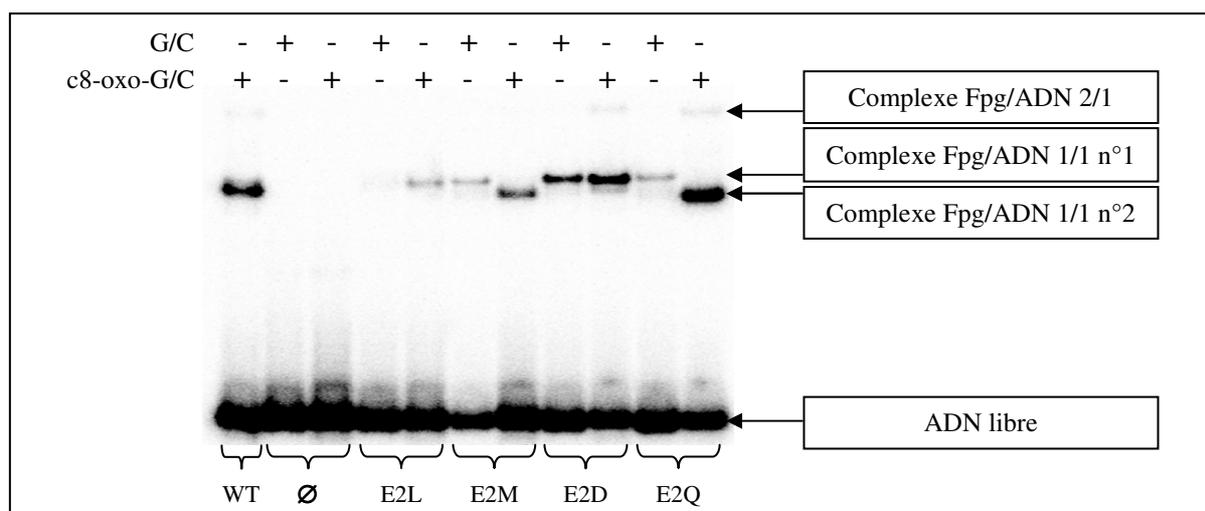
**Figure 42 :** Structure plane de la 8-oxo-G et de son analogue c8-oxo-G. La partie 8-oxo-G des deux molécules est représentée en noir, et la partie squelette phosphate/sucre ou phosphate/cyclopentane de la 8-oxo-G et du c8-oxo-G sont représentées en vert. Le carbone remplaçant l'oxygène du désoxyribose pour le c8-oxo-G est représenté en rouge.

### II.1.2.1. Analyse fonctionnelle des mutants de *LIFpg* dans un contexte d'extraits cellulaires bruts

Après introduction des mutations sur le plasmide codant pour *LIFpg* de type sauvage, les protéines recombinantes correspondantes ont été surproduites dans une souche de *E. coli* pour laquelle le gène codant pour *EcFpg* a été délété (voir le « Matériel et méthodes »). Après avoir lysé les cellules et estimé la concentration en *LIFpg* et en protéines totales dans les extraits cellulaires bruts obtenus, nous avons normalisé la quantité de protéine *LIFpg* dans nos différents extraits contenant les différentes formes mutantes de la protéine (voir le « Matériel et méthodes »). Nous avons ensuite réalisé des tests fonctionnels sur la 8-oxo-G ou son analogue carbocyclique avec nos mutants de *LIFpg* dans un contexte d'extraits cellulaires bruts. La mise au point de tels tests utilisant des extraits cellulaires peut nous permettre de discriminer les mutants intéressants de ceux qui ne le sont pas en gagnant beaucoup de temps par rapport aux mêmes manipulations sur les protéines purifiées (la purification d'un mutant de Fpg dure environ une semaine).

**- Analyse de la fixation des mutants de *Ll*Fpg sur des ADN contenant une guanine ou un c8-oxo-G dans un contexte d'extraits cellulaires bruts**

Sur cette base, nous avons mesuré par électrophorèse en gel retard la capacité de nos différents mutants à se fixer sur des duplexes de 14 paires de bases G/C (sans lésion) et c8-oxo-G/C (avec un analogue de lésion) (Tableau 20 du « Matériel et méthodes ») dans un contexte d'extrait cellulaire brut (conditions alternatives n°1, voir le « Matériel et méthodes »). Ces manipulations ont donc eu pour but de vérifier si nos mutants protéiques possédaient toujours la capacité de se fixer à l'ADN, et si cette fixation était toujours spécifique des duplexes d'ADN contenant l'analogue de lésion c8-oxo-G par rapport à ceux ne contenant pas de lésion. Un exemple de ce type de manipulations sur les extraits exprimant les mutants E2L, E2M, E2D et E2Q est présenté dans la Figure 43. Seuls les résultats obtenus avec les mutants de la position 2 de *Ll*Fpg seront présentés dans ce sous-chapitre, par contre, le Tableau 9 du sous-chapitre suivant présentera un récapitulatif des données obtenues avec tous nos mutants.



**Figure 43 :** Gel en conditions natives correspondant à la fixation de Fpg WT ou mutantes sur un duplexe c8-oxo-G/C ou G/C.

La concentration finale en Fpg pour chaque mutant est d'environ 32 nM. La notation Ø correspond à un extrait brut ne surproduisant pas Fpg. Les mutants utilisés pour ce gel sont indiqués en bas de celui-ci.

Les bandes retardées par rapport à la migration de l'ADN libre sont attribuées à des complexes ADN/protéine (Figure 43). De tels complexes entre Fpg et les duplexes G/C et c8-oxo-G/C sont présents avec tous nos mutants de *Ll*Fpg (les résultats obtenus pour les mutants E2N, E2R et E2T ne sont pas présentés, mais donnent des résultats similaires), hormis pour l'extrait ne surproduisant pas Fpg (noté Ø dans la Figure 43). Cela indique d'une part que la

souche bactérienne utilisée ne surproduit aucune protéine autre que Fpg se fixant à l'ADN (sinon on verrait d'autres bandes de complexes pour l'essai qui ne surproduit pas Fpg). D'autre part, cela indique que tous nos mutants de Fpg en position 2 ont toujours la capacité de se fixer à l'ADN.

Par contre, les complexes sont formés en quantités variables selon le mutant utilisé pour une même sonde d'ADN (ex : pistes 5 et 11, Figure 43). Si on considère que nos estimations des concentrations en protéine Fpg dans nos extraits sont fiables, ce point nous indique que nos différents mutants se fixent à l'ADN avec des affinités différentes, et donc que la mutation du résidu E2 joue sur l'efficacité de fixation de Fpg à l'ADN. A ce niveau, on peut indiquer que ce sont les mutants E2Q et E2D qui semblent être les moins touchés quant à leur efficacité de fixation à l'ADN parmi tous les mutants testés de Fpg en position 2.

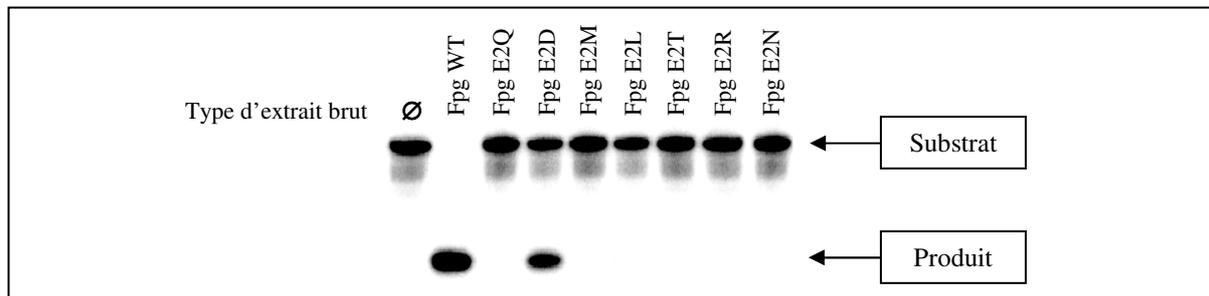
D'autre part, les complexes sont formés en quantités variables selon la sonde oligonucléotidique utilisée pour un même mutant (avec ou sans lésion, ex : pistes 10 et 11, Figure 43), la quantité de complexe étant toujours plus importante avec l'ADN contenant un c8-oxo-G qu'avec celui contenant un G/C. Ceci nous indique qu'aucun de nos mutants n'a perdu sa capacité à se fixer préférentiellement à l'ADN endommagé par rapport à l'ADN normal.

Enfin, les complexes observés présentent différentes mobilités électrophorétiques, que l'on a attribuées à deux types de complexes de stœchiométrie protéine/ADN 1/1 et un complexe de stœchiométrie protéine/ADN 2/1 (Figure 43). Dans la plupart des cas, la bande correspondant au complexe observé avec la sonde G/C migre moins que celle observée avec la sonde spécifique c8-oxo-G/C. C'est en particulier le cas de l'extrait bactérien surexprimant la Fpg de type sauvage (données non présentées). De façon surprenante, ce phénomène n'est pas observé avec les extraits bactériens surproduisant les protéines mutantes E2L et E2D. Ces bandes correspondent vraisemblablement à différentes conformations des complexes. Cependant en l'état actuel de nos connaissances, nous n'avons pas d'explication simple pour interpréter ce phénomène.

#### **- Analyse de l'activité ADN glycosylase des mutants de L/Fpg sur des ADN contenant une 8-oxo-G dans un contexte d'extraits cellulaires bruts**

L'activité ADN glycosylase des différents mutants de L/Fpg dans un contexte d'extraits bruts a été testée sur un duplexe de 24 paires de bases contenant une 8-oxo-G appariée à une cytosine (Tableau 20 du « Matériel et méthodes »). Ces manipulations ont donc eu pour but de

vérifier si nos mutants protéiques possédaient toujours la capacité d'exciser une 8-oxo-G dans l'ADN. Un exemple de ce type de manipulations sur les extraits exprimant la *LIFpg* de type sauvage et ses mutants en position 2 est présenté dans la Figure 44.



**Figure 44 : Gel dénaturant l'ADN correspondant aux tests d'activité des mutants de Fpg dans les extraits bruts sur le duplexe 8-oxo-G/C.**

La concentration en Fpg pour chaque mutant est d'environ 14  $\mu$ M. La notation  $\emptyset$  correspond à un extrait brut ne surproduisant pas Fpg, soit un contrôle négatif. Les mutants utilisés sont indiqués au dessus du gel. La piste avec l'extrait surproduisant la Fpg de type sauvage (WT) correspond à un contrôle positif.

Parmi les mutants de *LIFpg* testés pour leur activité ADN glycosylase sur la 8-oxo-G, on peut distinguer trois classes de mutants :

- (i) Ceux qui sont peu affectés (mutants M75L, S217A, T221A et R260I), pour lesquels une concentration finale en protéine de quelques nanomolaires suffit pour détecter une activité non négligeable.
- (ii) Ceux qui sont affectés de façon significative mais qui conservent une légère activité (mutant E2D, piste 4 de la Figure 44, et mutant P1G, Serre *et al.*, 2002), pour lesquels une concentration finale en protéine de plusieurs micromolaires est nécessaire pour détecter une activité non négligeable.
- (iii) Ceux qui sont totalement dépourvus d'activité (E2N, E2R, E2T, E2L, E2M, E2Q, E5Q, M75C, E76Y, R109I, R109S, Y173F, S217Q, I219D, Y222A et R260L).

Concernant les mutants de *LIFpg* en position 2, seul le mutant E2D est associé à une activité résiduelle. Par contre, ce mutant est nettement plus affecté dans son activité ADN glycosylase que les mutants de *LIFpg* de classe (i) (données non présentées). En effet, dans nos extraits, il faut environ 14  $\mu$ M de Fpg mutante E2D pour avoir environ 50% de coupure, contre 0,8 nM de *LIFpg* de type sauvage.

**- Récapitulatif des analyses fonctionnelles des mutants de *LIFpg* dans un contexte d'extraits cellulaires bruts**

Finalement, les données d'activité ADN glycosylase et de fixation pour nos différents mutants de *LIFpg* dans un contexte d'extraits cellulaires sont récapitulées dans le Tableau 9.

<b>Tableau 9 : Récapitulatif des analyses fonctionnelles des mutants de <i>LIFpg</i>.</b>												
<b><i>LIFpg</i></b>	<b>P1G</b>	<b>E2N</b>	<b>E2R</b>	<b>E2T</b>	<b>E2L</b>	<b>E2M</b>	<b>E2D</b>	<b>E2Q</b>	<b>E5Q</b>	<b>M75C</b>	<b>M75L</b>	<b>E76Y</b>
Activité ADN glycosylase	oui (faible)	non	non	non	non	non	oui (faible)	non	non	non	oui	non
Fixation à l'ADN	oui	non	non	oui	non							
<b><i>LIFpg</i></b>	<b>R109I</b>	<b>R109S</b>	<b>Y173F</b>	<b>S217A</b>	<b>S217Q</b>	<b>I219D</b>	<b>T221A</b>	<b>Y222A</b>	<b>R260I</b>	<b>R260L</b>	<b>WT</b>	
Activité ADN glycosylase	non	non	non	oui	non	non	oui	non	oui	non	oui	
Fixation à l'ADN	non	non	non	oui	oui	non	oui	oui	oui	non	oui	

On peut ainsi définir quatre classes de mutants :

- (i) Les mutants non affectés par la mutation (en rouge dans le Tableau 9).
- (ii) Les mutants P1G et E2D, qui sont affectés uniquement dans leur activité ADN glycosylase et pas dans leur fixation à l'ADN, mais qui conservent tout de même une légère activité (en orange dans le Tableau 9).
- (iii) Les mutants totalement défectueux en activité ADN glycosylase, mais toujours capables de se fixer à l'ADN (en vert dans le Tableau 9).
- (iiii) Les mutants incapables de se fixer à l'ADN, et donc totalement défectueux en activité ADN glycosylase (en bleu dans le Tableau 9).

**II.1.2.2. Validation des résultats obtenus avec les extraits cellulaires bruts via l'analyse fonctionnelle de mutants de *LIFpg* purifiés**

Les résultats que l'on a obtenus dans un contexte d'extrait cellulaire brut avec le mutant P1G de *LIFpg* sont tout à fait cohérents avec les données précédentes de l'équipe sur ce mutant purifié (Serre *et al.*, 2002). En effet, ce mutant purifié est toujours capable de se fixer à un

ADN contenant une 8-oxo-G avec une affinité proche de celle de la protéine de type sauvage, et est fortement déficient en activité ADN glycosylase sur la 8-oxo-G même s'il possède toujours une activité résiduelle à haute concentration protéique.

De même, les résultats que l'on a obtenus dans un contexte d'extrait cellulaire brut avec le mutant E2Q de *L/Fpg* sont tout à fait en accord avec les données de la bibliographie sur ce mutant purifié chez *E. coli* et *B. stearothermophilus* (Lavrukhin *et al.*, 2000; Fromme *et al.*, 2003a). En effet, ce mutant purifié est toujours capable de se fixer à un ADN, mais est totalement déficient en activité 8-oxo-G-ADN glycosylase. Par contre, nous avons utilisé dans notre étude des mutants de la Fpg de *L. lactis*, ce qui fait que les résultats de la bibliographie ne sont pas forcément comparables à ceux que nous avons obtenus.

Pour finir de confirmer que les résultats que nous avons obtenus avec les extraits cellulaires bruts sont valables, nous avons choisi de caractériser fonctionnellement les mutants E2Q et E2D de *L/Fpg* purifiés, en comparant les résultats à ceux que l'on obtient avec l'enzyme de type sauvage purifiée. Les trois formes de *L/Fpg* ont été purifiées jusqu'à homogénéité en utilisant exactement le même protocole (voir le « Matériel et méthodes »).

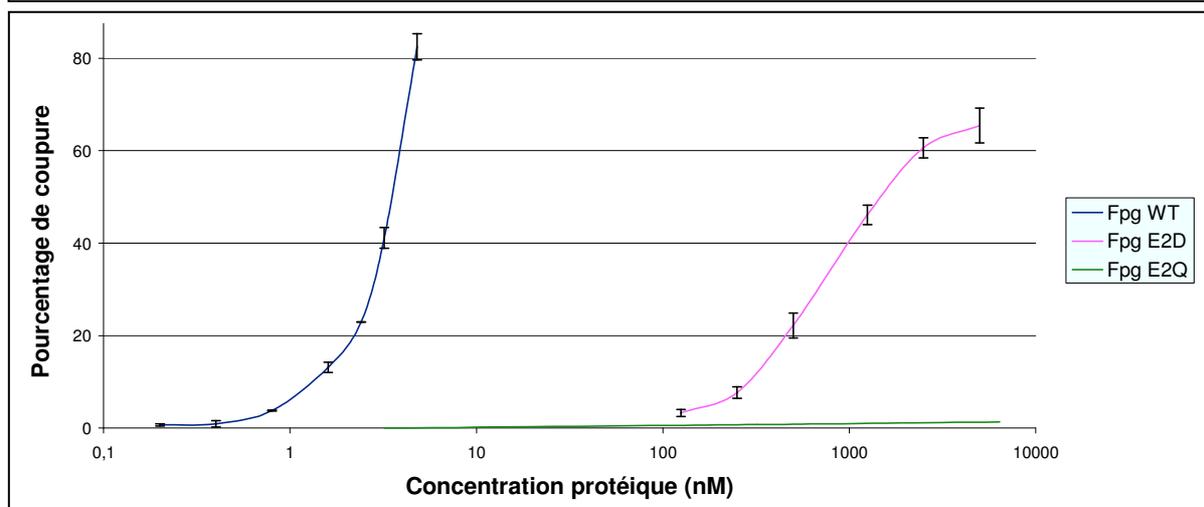
Ainsi, nous avons vérifié que les trois formes de *L/Fpg* testées (WT, E2D et E2Q) possèdent une affinité similaire pour l'ADN contenant un c8-oxo-G apparié à une cytosine ( $K_D$  apparent de 4-5 nM, Tableau 10).

**Tableau 10 :** Constantes de dissociations apparentes des *L/Fpgs* WT ou mutantes E2Q et E2D pour le duplexe c8-oxo-G/C de 14 paires de bases.

Enzymes	Fpg-WT	Fpg-E2Q	Fpg-E2D
$K_D$ app (nM)	5,4	5,2	3,4

De plus, comme cela avait été suggéré par les expériences précédentes, le mutant E2Q de *L/Fpg* est complètement inactif (courbe verte, Figure 45), tandis que le mutant E2D est toujours associé à une activité ADN glycosylase résiduelle (courbe rose, Figure 45).

**Figure 45 : Evolution de la coupure en 20 min et à 37°C par *L/Fpg* WT ou mutantes E2Q et E2D d'un duplexe 8-oxo-G/C de 24 paires de bases.**



En effet, il faut environ 300 fois plus de *L/Fpg* mutante E2D que d'enzyme WT pour couper 40 % du duplexe 8-oxo-G/C. De plus, on a moins de 2% de coupure de ce duplexe en utilisant une concentration en *L/Fpg* mutante E2Q de 6,4  $\mu$ M (soit à peine plus que le bruit de fond lors de nos quantifications de l'intensité des bandes sur gel), alors que 0,8 nM d'enzyme de type sauvage donne déjà 3,8 % de coupure dans les mêmes conditions.

Finalement, l'analyse fonctionnelle que nous avons réalisée sur les *L/Fpg* de type sauvage et mutantes E2D et E2Q purifiées donne des résultats identiques à ceux que nous avons obtenus avec les extraits cellulaires bruts, ces résultats étant conformes aux données de la bibliographie (Lavrukhin *et al.*, 2000; Fromme *et al.*, 2003a), ce qui confirme sans ambiguïté la validité de nos tests fonctionnels sur les extraits.

### **II.1.2.3. Conclusion et discussion**

Les résultats que nous avons obtenus avec les extraits cellulaires bruts ont montré une bonne corrélation avec ceux que l'on a obtenus avec les enzymes de type mutant purifiées (P1G, E2D et E2Q), ainsi qu'avec les données de la bibliographie (voir le sous-chapitre précédent). Nous pourrions donc à l'avenir nous contenter de faire des essais préliminaires sur des extraits bruts afin de discriminer les mutants intéressants à purifier pour une analyse structure/fonction.

L'interprétation des résultats d'expériences de mutagenèse dirigée reste cependant difficile car on ne sait généralement pas si l'effet d'une mutation donnée résulte de l'absence

du résidu amino-acide initialement présent dans la séquence de l'enzyme de type sauvage, ou si cet effet provient de son remplacement par un autre résidu. Afin d'éviter au maximum ce biais, il faudrait chaque fois remplacer un résidu par les 19 autres acides aminés possibles. Nous avons partiellement suivi cette stratégie pour les mutants de *LlFpg* en position 2.

Bien qu'incomplets, nous allons tout de même tenter de faire une analyse fonctionnelle de nos résultats de mutagenèse. Une attention particulière sera donnée aux rôles des résidus P1, E2, Y222 et R260 de *LlFpg*. La Figure 40 présente une vue structurale détaillée du site actif de *LlFpg* de type sauvage lorsque celle-ci est complexée à un ADN contenant un analogue de lésion cFapyG (code PDB 1XC8, Coste *et al.*, 2008). Le choix des acides aminés ciblés (excepté pour le résidu P1) par la mutagenèse repose sur le fait que leur chaîne latérale ou chaîne principale est impliquée dans les contacts suivants :

- une interaction directe ou *via* une molécule d'eau avec la base endommagée (ex : mutants sur E2, E5, M75, E76, S217, I219, T221 et Y222, Figure 40).
- une interaction avec un résidu qui est lui-même en interaction avec la base endommagée (ex : mutants sur Y173 et S217, Figure 40).
- une interaction avec les phosphates de l'ADN bordant le nucléoside endommagé (ex : R260, Figure 41).
- une interaction avec la cytosine opposée au dommage (ex : R109, Figure 41).

La plupart des résidus ciblés par la mutagenèse sont conservés parmi les structures primaires de *Fpg* (Figure 22, « Introduction »). En se basant sur les études structurales précédentes de complexes *Fpg*/ADN lésé, chaque substitution d'acide aminé vise à supprimer les contacts spécifiques entre un résidu de l'enzyme et l'ADN endommagé. Brièvement, une analyse de nos résultats indique que l'on peut classer en quatre catégories les substitutions que l'on a introduites dans *LlFpg* :

- (i) Celles qui n'ont aucun effet (en rouge dans le Tableau 9, mutants M75L, S217A, T221A et R260I).
- (ii) Celles qui affectent uniquement l'activité ADN glycosylase et pas la fixation à l'ADN, mais pour lesquelles les mutants conservent tout de même une légère activité (en orange dans le Tableau 9, mutants P1G et E2D).
- (iii) Celles qui abolissent totalement l'activité ADN glycosylase, mais pas la capacité de fixation à l'ADN (en vert dans le Tableau 9, mutants E2N, E2R, E2T, E2L, E2M, E2Q, S217Q et Y222A).

- (iii) Celles qui affectent la capacité de fixation à l'ADN et l'activité ADN glycosylase (en bleu dans le Tableau 9, mutants E5Q, M75C, E76Y, R109I, R109S, Y173F, I219D et R260L).

Afin de ne pas trop alourdir la discussion, je ne discuterai ici que les mutations affectant les résidus P1, E2, R260 et Y222.

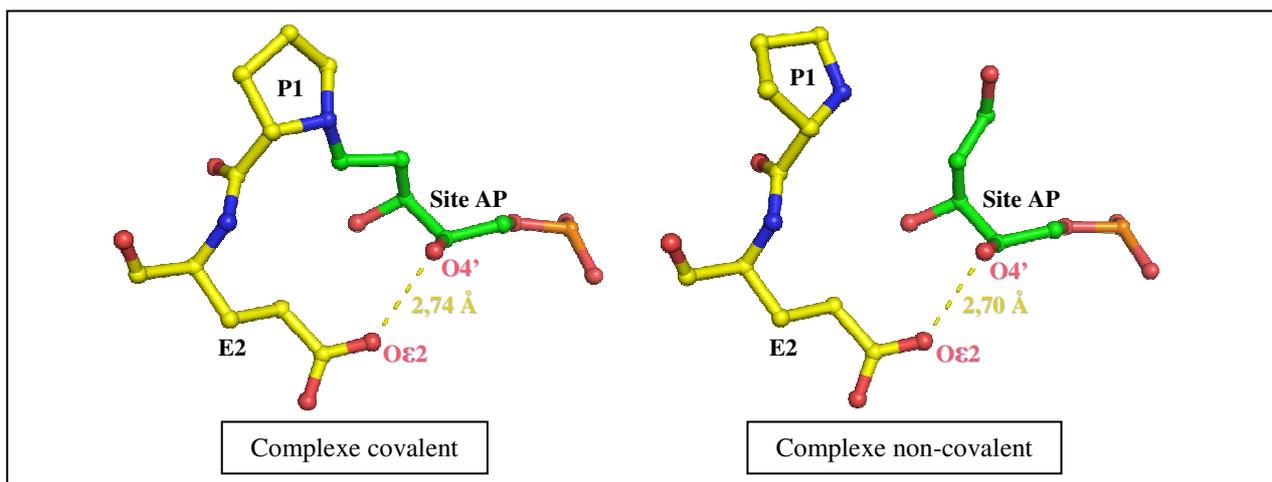
Nous savions déjà que l'enzyme *L/Fpg* mutante P1G était fortement déficiente en activité (Serre *et al.*, 2002; Pereira de Jésus *et al.*, 2005), ce qui est en accord avec nos essais sur les extraits bactériens. Par ailleurs, des études antérieures avaient démontré que le résidu P1 était impliqué *via* son amine terminale dans l'attaque nucléophile sur le C1' de la lésion, aboutissant à la formation d'un intermédiaire covalent de type base de Schiff (chapitre I.7.3 de l'introduction, Tchou *et al.*, 1995; Zharkov *et al.*, 1997; Sidorkina *et al.*, 2000). Notre analyse du mutant P1G de *L/Fpg* indique clairement que l'amine N-terminale d'une glycine n'est pas capable de catalyser efficacement cette attaque nucléophile. De plus, nous pouvons proposer que la haute nucléophilie de l'amine cyclique de la proline N-terminale de Fpg est nécessaire à une attaque nucléophile efficace par Fpg en C1' des lésions, étant donné que l'on sait que la nucléophilie de l'amine de la proline est très supérieure à celle de l'amine de tous les autres acides aminés (Brotzel *et al.*, 2007).

Certaines substitutions différentes sur un même résidu peuvent affecter différemment l'enzyme *L/Fpg*. C'est notamment le cas des mutations R260I et R260L (Tableau 9), qui sont respectivement associées à aucun effet et à une perte d'activité et de fixation à l'ADN. Les études structurales ont montré que le résidu R260 contacte les deux phosphates de part et d'autre du nucléoside endommagé *via* l'un des azotes de sa chaîne latérale (Figure 41). Nous nous attendions donc à ce que les protéines pour lesquelles R260 est substituée (par des résidus dont la chaîne latérale ne contient aucun donneur/accepteur efficace de liaisons hydrogène) amènent une forte baisse de la capacité de fixation à l'ADN de Fpg, et par conséquent de l'activité enzymatique. C'est en fait le cas pour la mutation R260L, mais de façon surprenante ce n'est pas le cas pour la mutation R260I. Ceci pourrait suggérer que la chaîne latérale de l'isoleucine serait en fait capable de réaliser des interactions avec les phosphates entourant la lésion, ce qui est pourtant peu probable étant donné que les chaînes latérales de l'arginine et de l'isoleucine ont des tailles différentes, et que le méthyle terminal de la chaîne latérale de l'isoleucine est un très mauvais donneur de liaison hydrogène. En l'absence de données structurales sur un complexe formé entre *L/Fpg* mutante R260I et un ADN lésé, il nous est donc impossible de confirmer ou infirmer cette dernière hypothèse.

Le résidu Y222 de la boucle  $\alpha$ F- $\beta$ 9 semble aussi être un résidu important pour l'activité de *LIFpg*. En effet, si on regarde les alignements de séquence de cette région de la protéine (Figure 22 de l'introduction, et Figure 5B de Coste *et al.*, 2008), on voit que l'on a toujours une tyrosine ou une phénylalanine à cette position de la protéine. Cela indique que la présence d'un résidu aromatique à cette position est probablement importante pour l'activité de Fpg. De plus, le mutant Y222A de *LIFpg* que nous avons étudié est totalement déficient en activité ADN glycosylase, mais est toujours capable de se fixer sur l'ADN lésé. Cela tend donc à confirmer que ce résidu est important uniquement pour l'activité de la protéine. Cependant, en l'état actuel de nos connaissances, nous ne savons pas quel est le rôle de ce résidu.

Le résidu le plus ciblé lors de notre étude de mutagenèse est le résidu E2 (mutants E2N, E2R, E2T, E2L, E2M, E2D et E2Q). Nous avons choisi de faire une étude plus approfondie sur le rôle de E2 car de nombreuses données structurales et fonctionnelles sont disponibles dans la littérature sur le rôle de E2 pour les Fpgs de types sauvages ou mutantes E2Q (Lavrukhin *et al.*, 2000; Gilboa *et al.*, 2002; Fromme *et al.*, 2003a; Perlow-Poehnelt *et al.*, 2004). De la même façon que pour notre étude avec *LIFpg*, la mutation E2Q pour les Fpg de *E. coli* et *B. stearothermophilus* affecte l'activité enzymatique sans pour autant perturber la fixation à l'ADN. La *BstFpg* mutante E2Q a d'ailleurs été utilisée pour résoudre la structure d'un complexe abortif avec un ADN contenant une 8-oxo-G (Fromme *et al.*, 2003a). Les premières évidences d'un rôle important de E2 ont été fournies par l'étude structurale par cristallographie de complexes covalents de type base de Schiff réduite entre Fpg et un ADN lésé (Gilboa *et al.*, 2002; Fromme *et al.*, 2002), et d'un complexe non-covalent entre *BstFpg* et un duplexe d'ADN contenant un site abasique réduit (Fromme *et al.*, 2002) (Figure 46). Ces structures révèlent clairement que E2 établit une liaison hydrogène avec le C4'-OH du désoxyribose sous forme ouverte lors d'un état de transition de la catalyse.

Il a été proposé par les auteurs (Gilboa *et al.*, 2002; Fromme *et al.*, 2002) que cette liaison hydrogène participerait au maintien d'une conformation du sucre ouvert optimale pour la catalyse. Cette hypothèse serait d'ailleurs en mesure d'expliquer la forte baisse d'activité ADN glycosylase de *LIFpg* mutante E2D, étant donné que le raccourcissement de la chaîne latérale entre un glutamate et un aspartate est susceptible de rendre difficile voire impossible cette liaison hydrogène cruciale.

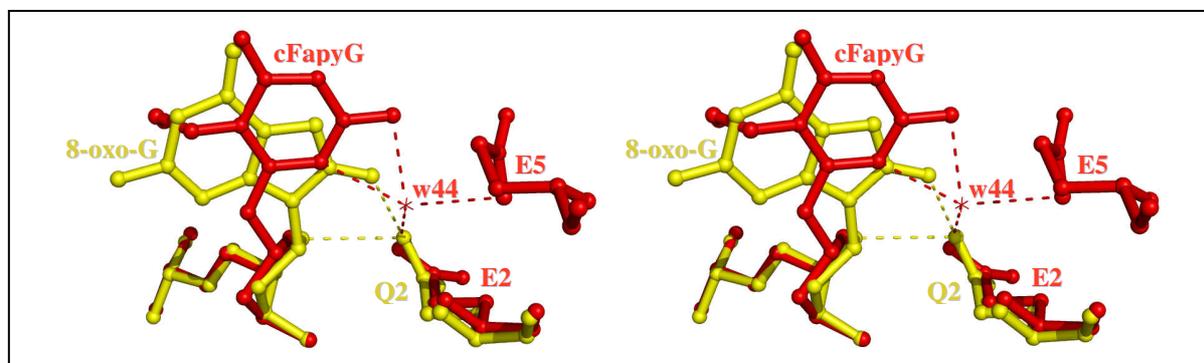


**Figure 46 : Stabilisation de la forme ouverte du site AP par une liaison hydrogène avec le résidu E2 de Fpg.** Les carbones de la protéine et du site AP sont respectivement colorés en jaune et en vert, les azotes sont colorés en bleu, les oxygènes en rose, et les phosphates en orange. Deux structures de complexes formés entre *Bst*Fpg et le site AP sont présentées (Fromme *et al.*, 2002), et correspondent à un complexe covalent de type base de Schiff réduite (à gauche, code PDB 1L1Z) et à un complexe non-covalent avec le site AP réduit (à droite, code PDB 1L1T). La liaison hydrogène importante pour la stabilisation du désoxyribose sous forme ouverte est représentée par un pointillé jaune.

Cependant, cette interaction *via* E2 n'est pas en mesure d'expliquer pourquoi la mutation E2Q de l'enzyme induit une perte totale de l'activité ADN glycosylase de Fpg, étant donné qu'un résidu glutamine devrait toujours être capable de réaliser cette liaison hydrogène importante. D'autre part, le mutant E2Q de *Ec*Fpg est affecté seulement d'un facteur 2 dans son activité AP lyase (Lavrukhin *et al.*, 2000). Le résidu E2 de Fpg semble donc essentiel uniquement pour l'activité ADN glycosylase de Fpg, et pas pour son activité AP lyase. Il est donc difficile de s'imaginer que la seule implication de E2 serait de maintenir une bonne conformation du désoxyribose, étant donné que dans ce cas la mutation E2Q devrait aussi toucher l'efficacité de l'activité AP lyase de Fpg (Figure 33 de l'introduction). Finalement, ces données structurales ne permettent donc pas d'expliquer quel est le rôle joué par E2 dans les activités ADN glycosylase et AP lyase de Fpg.

Par contre, l'hypothèse que nous avons proposée dans la publication (Coste *et al.*, 2008) concernant l'implication de E2 et d'une molécule d'eau dans l'activation du résidu P1 est en parfait accord avec nos données de mutagenèse. Dans ce mécanisme, la molécule d'eau est un co-catalyseur qui *via* E2 arrache un proton à P1. L'action de cette molécule d'eau dans la catalyse revient donc à baisser significativement le pKa réel de l'amine du résidu P1 (normalement 10,6). Dans ces conditions, le résidu P1 déprotoné devient un bon nucléophile à pH physiologique, et peut ainsi effectuer l'attaque en C1' du ribose initiant l'activité ADN glycosylase de Fpg. En effet, le remplacement d'un résidu E2 (chaîne latérale acide de pKa = 4) par un résidu dont la chaîne latérale est neutre ou basique est susceptible d'empêcher

l'activation du résidu P1. Cette hypothèse explique donc très bien l'absence d'activité observée pour les versions mutantes de *L/Fpg* autres que E2D (E2N, E2R, E2T, E2L, E2M, E2Q), étant donné qu'aucun de ces mutants ne possède une chaîne latérale acide. Elle est également en bon accord avec le fait que le mutant E2D de *Fpg* soit encore légèrement actif. En effet, l'aspartate possède une chaîne latérale acide, ce qui permettrait l'activation de P1. Par contre, la chaîne latérale moins longue pour un aspartate que pour un glutamate pourrait rendre difficile la déprotonation de P1, et donc permettrait d'expliquer la moindre activité du mutant E2D par rapport à la *Fpg* de type sauvage. En outre, notre structure de *L/Fpg* complexée au cFapyG (code PDB 1XC8, Coste et al., 2008) supporte l'hypothèse de l'intervention d'une molécule d'eau dans la catalyse. En effet, dans cette structure, on voit une molécule d'eau structurale (notée w44 dans le fichier PDB) idéalement positionnée pour être activée par E2 en vue d'aller déprotoner P1 (Figure 47). Par contre, aucune molécule d'eau n'est présente dans la structure de *BstFpg* mutante E2Q complexée à un ADN contenant une 8-oxo-G (Fromme *et al.*, 2003a) (Figure 47). Dans cette structure, les interactions avec la molécule d'eau sont remplacées par des interactions directes entre le résidu muté Q2 et les groupements O8 et O4' de la lésion 8-oxo-G. De plus, l'analyse structurale comparative indique que le cFapyG est stabilisé dans une conformation *anti*, alors que la 8-oxo-G est stabilisée dans une conformation *syn*.



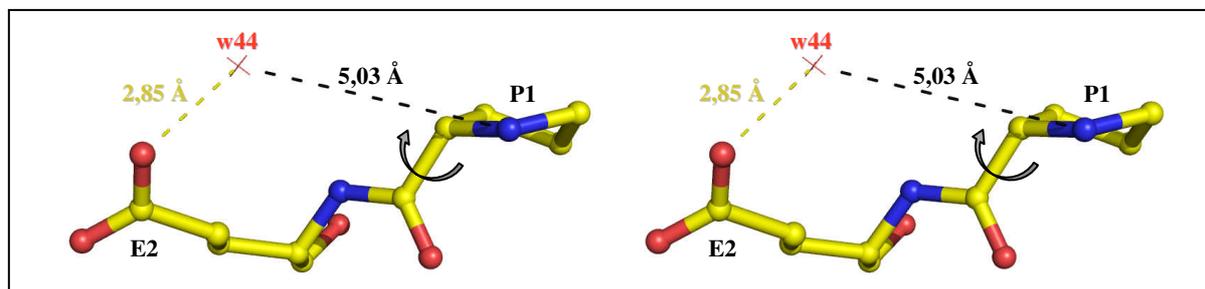
**Figure 47 :** Réseaux de liaisons hydrogène respectivement établis par les résidus E2 ou Q2 des structures *L/Fpg*WT/cFapyG et *BstFpg*E2Q/8-oxo-G (codes PDB 1XC8 et 1R2Y, respectivement). La structure *L/Fpg*WT/cFapyG est représentée en rouge, et la structure *BstFpg*E2Q/8-oxo-G est représentée en jaune. Les liaisons hydrogène sont représentées par des pointillés, et la molécule d'eau importante dans la structure *L/Fpg*WT/cFapyG est représentée par une croix rouge.

Dans ce cadre, on peut très bien se demander si la structure du complexe *BstFpg*E2Q/8-oxo-G ne serait pas biaisée par la présence de la mutation E2Q ou par l'utilisation d'une enzyme issue d'une bactérie thermophile. En effet, plusieurs arguments tendraient à nous rendre critiques quant à cette structure. Premièrement, plusieurs publications de dynamique moléculaire (Perlow-Poehnelt *et al.*, 2004 ; Zaika *et al.*, 2004) ont indiqué que la stabilisation

de la 8-oxo-G dans le site actif de Fpg est probablement possible en *syn* et en *anti*, mais que la forme *anti* est relativement plus stable avec l'enzyme de *E. coli* de type sauvage (Perlow-Poehnelt *et al.*, 2004). Les auteurs ont proposé que la structure de Fromme et collaborateurs était biaisée par l'utilisation du mutant E2Q, qui aurait tendance à favoriser la forme *syn* de la 8-oxo-G. D'autre part, nous avons aussi réalisé au laboratoire des essais de cristallographie sur un complexe *LI*FpgWT/c8-oxo-G (donc avec l'enzyme de type sauvage et un analogue carbocyclique de 8-oxo-G). Le modèle structural obtenu est de bonne qualité (données non présentées), hormis pour la partie base nucléique de la lésion pour laquelle la densité électronique n'est pas définie. Cela indique que cette partie du modèle est mobile. Ce résultat semble être en bon accord avec les conclusions précédentes concernant la possibilité de stabilisation de la 8-oxo-G en *syn* et *anti*, amenant donc un argument supplémentaire pour dire que la structure de Fromme était biaisée par l'utilisation du mutant E2Q (probable) ou par l'utilisation d'une protéine de thermophile (moins probable). De ce point de vue, il serait intéressant de résoudre la structure du mutant E2Q de *LI*Fpg complexé au duplexe c8-oxo-G/C, car dans le cas où l'on ne verrait que la forme *syn*, cela nous permettrait de conclure qu'il y a bien un biais dans la structure de Fromme provenant de l'utilisation du mutant E2Q de Fpg. Finalement, toutes ces informations nous amènent à proposer l'hypothèse que la forme *syn* stabilisée dans la structure de Fromme serait une forme non clivable de la 8-oxo-G, et que la forme clivable serait la forme *anti*. Cette proposition serait tout à fait en accord avec notre hypothèse selon laquelle une molécule d'eau serait nécessaire à l'activation de P1. En effet, la stabilisation de la 8-oxo-G en *anti* rendrait possible la stabilisation de cette molécule d'eau dans le site actif *via* un réseau de liaisons hydrogène similaire à celui observé avec le cFapyG. En outre, la stabilisation en *syn* de la 8-oxo-G complexée au mutant E2Q de Fpg serait aussi en mesure d'expliquer pourquoi ce mutant est inactif, car dans ce cas l'absence de la molécule d'eau cruciale rendrait impossible l'activation de P1, et donc rendrait impossible l'activité ADN glycosylase de Fpg. La stabilisation en *syn* de la 8-oxo-G serait ainsi une conformation catalytiquement inactive.

Un autre point peut cependant être discuté quant à l'hypothèse de l'activation du résidu P1 par le résidu E2 *via* une molécule d'eau. En effet, si on regarde plus en détail la structure du complexe *LI*Fpg/cFapyG (Figure 48), la courte distance séparant le carbonyle de E2 et la molécule d'eau w44 (2,85 Å) est en accord avec un possible réaction acido-basique entre ces deux partenaires. Par contre, la déprotonation de P1 par l'ion hydroxyle généré semble être difficile à concevoir dans cet état conformationnel, étant donné que la distance entre ces

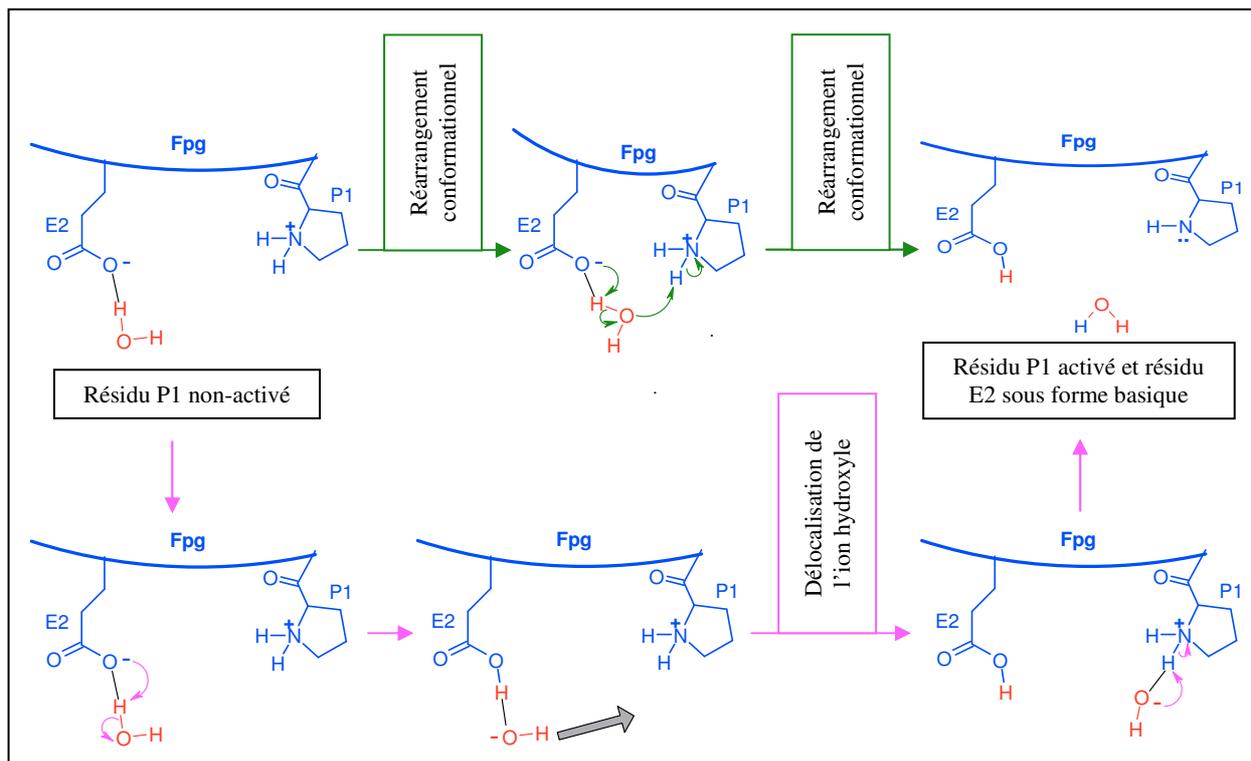
partenaires est relativement élevée (5,03 Å), et que l'ion hydroxyle généré reste probablement lié par liaison hydrogène à E2.



**Figure 48 :** Vue stéréoscopique de l'extrémité N-terminale de Fpg dans la structure de Fpg complexée au cFapyG (code PDB 1XC8).

Les carbones sont représentés en jaune, les azotes en bleu, et les oxygènes en rouge. La molécule d'eau 44 de ce fichier PDB est notée **w44** et est représentée par une croix rouge. La liaison hydrogène entre E2 et la molécule d'eau 44 est représentée par un pointillé jaune, et sa longueur est indiquée. La distance entre la molécule d'eau 44 et l'azote de P1 est localisée par un pointillé noir, et sa valeur est indiquée.

Plusieurs hypothèses sont plausibles pour expliquer comment la déprotonation de P1 est possible. La première, qui à mon sens est la plus probable, est que l'on ait un léger réarrangement conformationnel après ou pendant la déprotonation de la molécule d'eau par le résidu E2 de Fpg, réarrangement qui permettrait le rapprochement entre l'ion hydroxyle nouvellement créé et le résidu P1 de Fpg (Figure 49) (par exemple *via* une rotation du résidu P1 de Fpg autour de la liaison C $\alpha$ -C $\beta$ , voir la Figure 48). La seconde est que l'ion hydroxyle généré ne soit pas si fortement stabilisé qu'on pourrait le penser, rendant possible sa diffusion pour aller attaquer le résidu P1 (Figure 49). Enfin, une troisième possibilité, qui me semble moins probable, est qu'une autre molécule d'eau plus mobile soit responsable de cette activation, mais qu'on ne la voit pas dans la structure. Finalement, on peut indiquer que ces hypothèses seront difficiles à confirmer ou infirmer, étant donné que les structures cristallographiques que nous avons représentées un état « figé » des complexes Fpg/ADN, alors que nos hypothèses impliquent l'intervention d'une dynamique moléculaire. Dans ce cadre, la RMN pourrait être un outil de choix pour quantifier la dynamique associée aux résidus N-terminaux de Fpg (détermination de paramètres d'ordre S<sup>2</sup>). Cependant, l'attribution des résonances pour une protéine de plus de 200 résidus restant à l'heure actuelle difficile (d'autant plus si on parle d'un complexe protéine/ADN), cette approche par RMN sur nos complexes n'est pas pour le moment envisagée.



**Figure 49 : Hypothèses quant au mécanisme permettant l'activation par déprotonation du résidu P1 de Fpg.**

Les deux hypothèses proposées correspondant à un réarrangement conformationnel de Fpg ou à une délocalisation d'un ion hydroxyle sont respectivement indiquées par des flèches vertes et magenta. La molécule d'eau permettant l'activation de P1 est représentée en rouge, la protéine Fpg est représentée en bleu. Les liaisons hydrogène sont représentées par des pointillés noirs.

Un dernier point que nous pouvons discuter quant à l'intervention d'une molécule d'eau est que cette hypothèse est tout à fait en accord avec le mécanisme réactionnel de l'activité ADN glycosylase proposé à la fin de l'introduction du manuscrit. En effet, dans ce cas, après avoir arraché un proton à la molécule d'eau, le résidu E2 se trouve sous sa forme d'acide conjugué (protonée), et peut donc être l'acide responsable de la protonation de l'oxygène de l'hétérocycle du site AP qui permet sa stabilisation sous forme ouverte (Figure 32, « Introduction »).

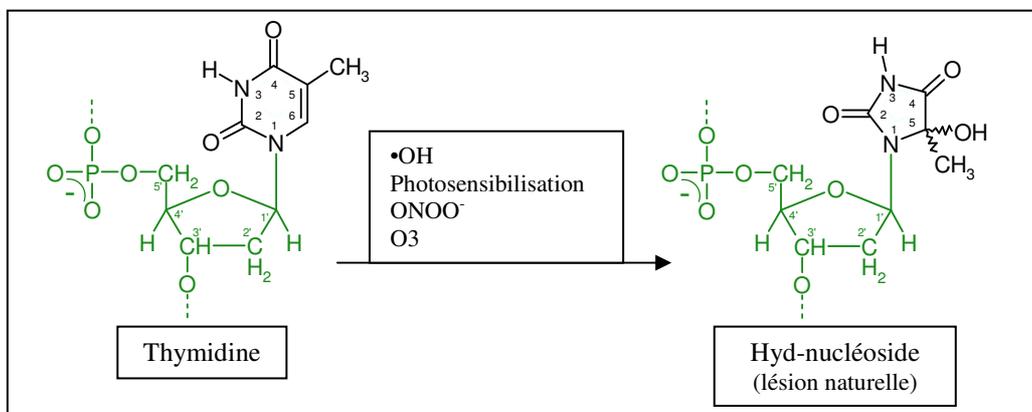
En conclusion, les études structurales et fonctionnelles que nous avons réalisées nous ont permis de faire avancer notre façon de comprendre l'activité ADN glycosylase de Fpg. Pour affiner notre modèle mécanistique, il faudrait faire une étude par cristallographie des rayons X de protéines Fpg mutantes qui seraient incapables ou capables de piéger la molécule d'eau dans le site actif de l'enzyme. Les mutantes E2Q et E2D de Fpg sont respectivement de bons candidats pour ces études. D'autre part, nous avons définis que certains de nos mutants protéiques étaient fortement touchés en activité, mais en conservant tout de même une légère

activité (*LIFpg* P1G et E2D). Ainsi, on peut se demander à partir de quel niveau une baisse d'activité de Fpg serait physiologiquement significative. Pour répondre à cette question, un test de complémentation d'une souche de *E. coli* délétée en *EcFpg* de type sauvage par nos formes mutantes de *LIFpg* est en cours de mise au point au laboratoire. Ce test nous permettra de quantifier l'activité de nos mutants non plus *in vitro*, mais *in vivo* en regardant des effets phénotypiques.

## II.2. Nouveaux éléments structuraux et fonctionnels sur le comportement de la Fpg de *Lactococcus lactis* (LlFpg) vis-à-vis d'une pyrimidine oxydée, la 5-Hydroxy-5-Méthyle-Hydantoïne (Hyd)

### II.2.1. Etude structure-fonction de la reconnaissance et de la métabolisation de l'Hyd par LlFpg

Bien que la protéine Fpg ait été décrite comme une ADN glycosylase impliquée dans l'excision des purines oxydées, de nombreux travaux indiquent que certaines pyrimidines oxydées peuvent être substrats de la Fpg de *E. coli* (*EcFpg*) (Hatahet *et al.*, 1994; Jurado *et al.*, 1998; Purmal *et al.*, 1998; Bourdat *et al.*, 1999; D'Ham *et al.*, 1999; Gasparutto *et al.*, 2000; Zhang *et al.*, 2000; Fromme *et al.*, 2003a; Simon *et al.*, 2006; Gasparutto *et al.*, 2009) (Figure 30, « Introduction »). En particulier, il a été montré récemment que la 5-Hydroxy-5-Méthyle-Hydantoïne (Hyd) est un substrat de *EcFpg* (Gasparutto *et al.*, 2000; Gasparutto *et al.*, 2009). Or cette lésion est plutôt décrite dans la littérature comme un substrat physiologique des ADN glycosylases procaryotes Nth et Nei (Breimer *et al.*, 1984; Dizdaroglu *et al.*, 1993; David *et al.*, 1998; Gasparutto *et al.*, 2000; Dizdaroglu *et al.*, 2005; Gasparutto *et al.*, 2009). L'Hyd est une lésion issue de l'oxydation de la thymine (Figure 50) (Halliwell *et al.*, 1991; Gasparutto *et al.*, 2000), ce qui fait qu'elle apparaît dans l'ADN en face d'une adénine dans les conditions physiologiques. L'Hyd est une lésion potentiellement létale *in vivo*, car elle bloque les ADN polymérase *in vitro* (Gasparutto *et al.*, 2000). Dans les cas où les ADN polymérase arrivent à passer la lésion, une adénine est majoritairement incorporée en face de l'Hyd *in vitro* (Gasparutto *et al.*, 2000). En conséquence, seul l'appariement de l'Hyd avec une adénine (Hyd/A) est physiologiquement relevant.



**Figure 50 : Formation de la 5-Hydroxy-5-Méthyle-Hydantoïne dans l'ADN.**

Adapté de Gasparutto *et al.*, 2000. La partie base nucléique de chaque nucléoside est indiquée en noir, et les squelettes phosphates/désoxyriboses sont indiqués en vert.

Dans ce chapitre, nous présenterons une analyse fonctionnelle et structurale visant à mettre en évidence les déterminants structuraux et fonctionnels de la reconnaissance et de la métabolisation de l'Hyd par *LIFpg*. Une analyse fonctionnelle comparative de l'action de *LIFpg* et de Nth de *E. coli* (*EcNth*) sur cette lésion sera aussi présentée.

### **II.2.1.1. Excision de l'Hyd par *LIFpg* et comparaison avec l'endonucléase III de *E. coli* (*EcNth*)**

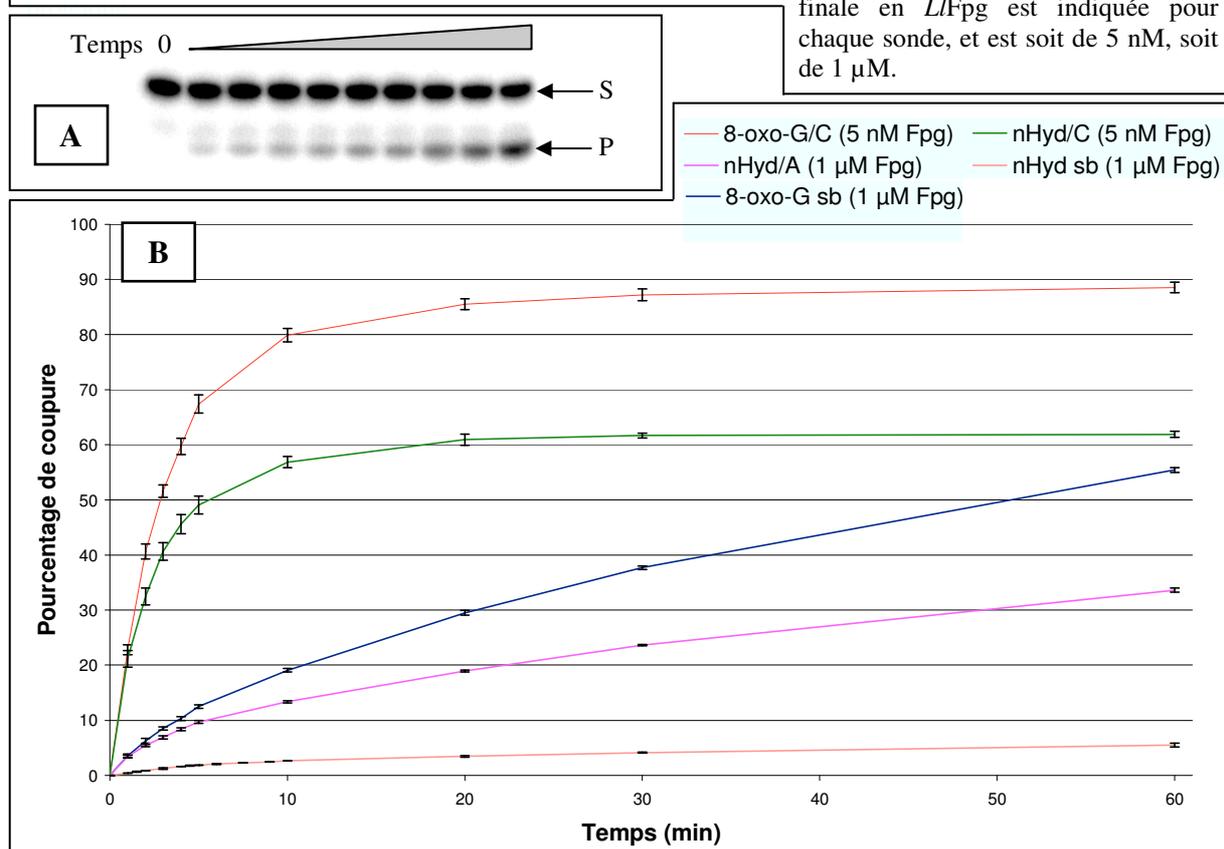
Dans un premier temps, nous avons examiné la capacité d'excision de l'Hyd par *LIFpg* sur de petits oligonucléotides synthétiques contenant une Hyd. Des cinétiques d'excision de l'Hyd ont été réalisées en tenant compte de la nature mono- ou bi-caténaire de l'ADN lésé, ainsi que de la base opposée au dommage. En parallèle, le même type d'expériences a été réalisé avec un oligonucléotide de même séquence contenant une 8-oxo-G comme substrat de référence (voir le « Matériel et méthodes »). En pratique, le brin contenant la lésion (Hyd ou 8-oxo-G) est radiomarké et chaque sonde oligonucléotidique est incubée avec une concentration d'enzyme donnée. Des prélèvements sont réalisés au cours du temps et analysés en gel de séquence afin de suivre l'avancement de la réaction (voir le « Matériel et méthodes »).

La présence d'un produit de clivage du brin lésé avec l'Hyd témoigne que cette lésion est bel et bien excisée par *LIFpg* (Figure 51A). Les résultats de quantification des expériences de cinétique d'activité sont donnés dans la Figure 51B (une attention particulière doit être donnée aux concentrations en *LIFpg*, car celles-ci ne sont pas identiques pour chaque sonde oligonucléotidique utilisée). Quel que soit le dommage considéré (8-oxo-G ou Hyd), l'activité ADN glycosylase de *LIFpg* est très peu efficace sur l'ADN simple-brin (courbes bleue et saumon, respectivement, Figure 51B), même si l'excision de la 8-oxo-G est plus efficace que celle de l'Hyd. Par contre, lorsque la lésion (8-oxo-G ou Hyd) est placée en face d'un C dans un ADN double brin, elle constitue un excellent substrat pour l'enzyme (courbes rouge ou verte, respectivement, Figure 51B), avec une activité ADN glycosylase de *LIFpg* légèrement meilleure pour la 8-oxo-G que pour l'Hyd. Il est cependant intéressant de noter que l'excision de l'Hyd dans l'ADN double-brin est fortement dépendante de la base qui lui est appariée. En effet, l'Hyd est un excellent substrat de l'enzyme quand elle est appariée à un C, et un très mauvais quand elle est appariée à un A (courbes verte et rose, respectivement, Figure 51B). En face d'un A, l'Hyd est même moins bien excisée que la 8-oxo-G dans un contexte simple-brin (courbes rose et bleue, respectivement, Figure 51B), ce qui témoigne de la très faible activité de *LIFpg* sur ce substrat.

**Figure 51 : Cinétique d'excision de la 8-oxo-G et de l'Hyd dans différents contextes oligonucléotidiques.**

(A) : Exemple de gel de séquence obtenu avec *L/Fpg* et le duplexe Hyd/C. Les notations **S** et **P** correspondent respectivement au substrat et au produit de réaction de l'enzyme. Les temps d'incubation correspondant à chaque piste sont de 0 ; 1 ; 2 ; 3 ; 4 ; 5 ; 10 ; 20 ; 30 et 60 minutes, respectivement. La concentration en protéine Fpg est de 5 nM. (B) : Courbes récapitulatives obtenues pour nos différents sondes oligonucléotidiques. Les notations /C ou /A font référence à des duplexes d'ADN avec respectivement un C ou un A en face de la lésion, la notation **sb** fait référence à des oligonucléotides simple-brins. La concentration

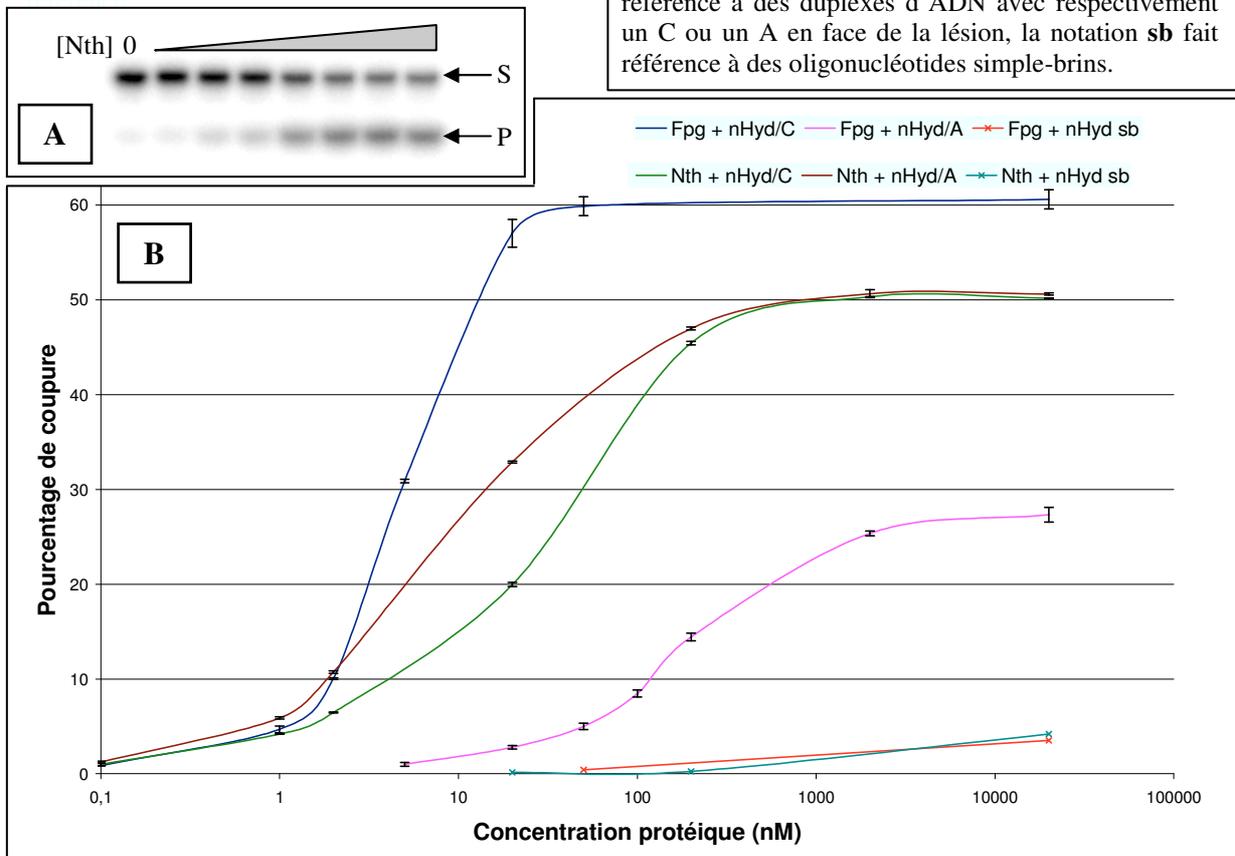
finale en *L/Fpg* est indiquée pour chaque sonde, et est soit de 5 nM, soit de 1  $\mu$ M.



Pour aller plus loin dans cette étude, nous avons ensuite comparé l'efficacité d'excision de l'Hyd par *L/Fpg* et par *EcNth*, une ADN glycosylase connue pour exciser efficacement cette lésion (Breimer *et al.*, 1984; Dizdaroglu *et al.*, 1993; David *et al.*, 1998; Gasparutto *et al.*, 2000; Dizdaroglu *et al.*, 2005; Gasparutto *et al.*, 2009). La présence d'un produit de clivage du brin lésé avec *Nth* témoigne que l'Hyd est, comme attendu, excisée par cette enzyme (Figure 52A).

**Figure 52 : Comparaison des activités Hyd-ADN glycosylase de *LIFpg* et *EcNth*.**

(A) : Exemple de gel de séquence obtenu avec *EcNth* et le duplexe Hyd/A. Les notations **S** et **P** correspondent respectivement au substrat et au produit de réaction de l'enzyme. Les concentrations en enzyme pour chaque piste sont de 0 ; 0,1 ; 1 ; 2 ; 20 ; 200 ; 2000 et 20000 nM, respectivement. (B) : Courbes récapitulatives obtenues pour *LIFpg* et *EcNth* avec nos différentes sondes oligonucléotidiques. Les notations /C ou /A font référence à des duplexes d'ADN avec respectivement un C ou un A en face de la lésion, la notation **sb** fait référence à des oligonucléotides simple-brins.

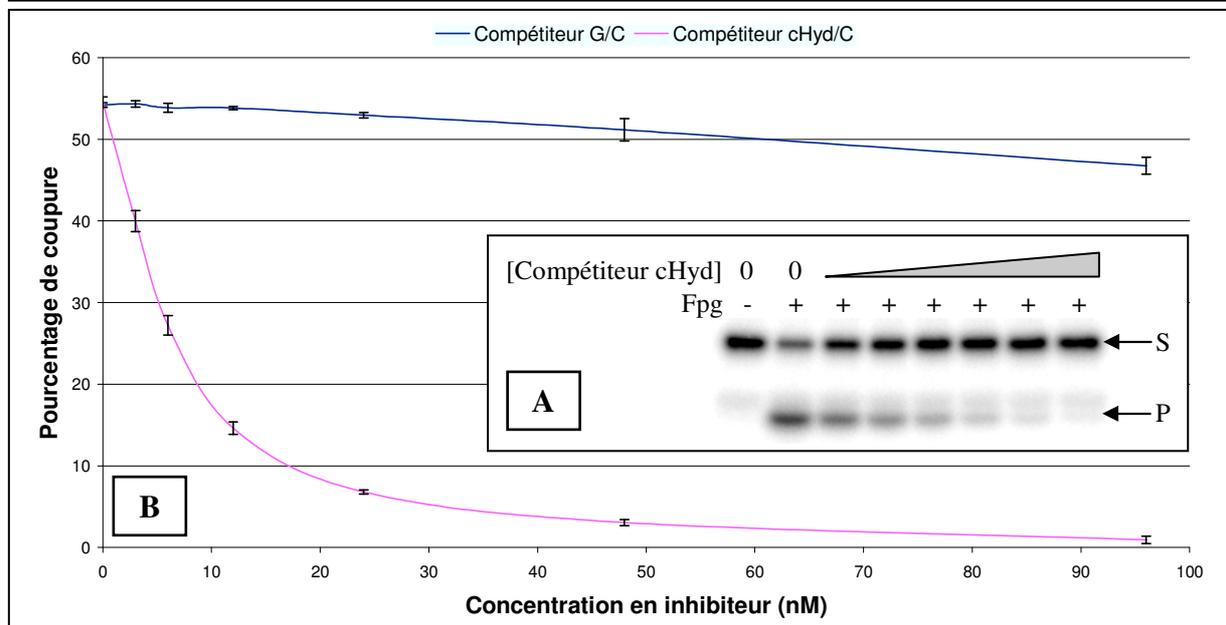


Quelle que soit l'enzyme considérée, on trouve à nouveau que l'Hyd est très mal excisée dans l'ADN simple-brin (courbes rouge et bleu-ciel, respectivement, Figure 52B). Comme nous l'avons observé sur les cinétiques d'excision (Figure 51), *LIFpg* excise beaucoup moins bien l'Hyd opposée à un A que l'Hyd opposée à un C (courbes rose et bleu-marine, respectivement, Figure 52B). Par contre, *EcNth* est nettement moins stricte quant à la préférence de base opposée à la lésion, même si elle montre une activité légèrement plus importante avec un A opposé à la lésion plutôt qu'un C (courbes marron et verte, respectivement, Figure 52B). On peut également noter que l'Hyd appariée à un C est un meilleur substrat pour Fpg qu'elle ne l'est lorsqu'elle est appariée à un A ou un C pour Nth (courbes bleue, marron et verte, respectivement, Figure 52B). De même, l'Hyd appariée à un A est un bien plus mauvais substrat pour Fpg que ne l'est l'Hyd appariée à un A ou un C pour Nth (courbes rose, marron et verte, respectivement, Figure 52B).



**Figure 54 : Inhibition de l'activité Hyd-ADN glycosylase de *LIFpg* par un oligonucléotide contenant un analogue cHyd ou par un oligonucléotide sans lésion.**

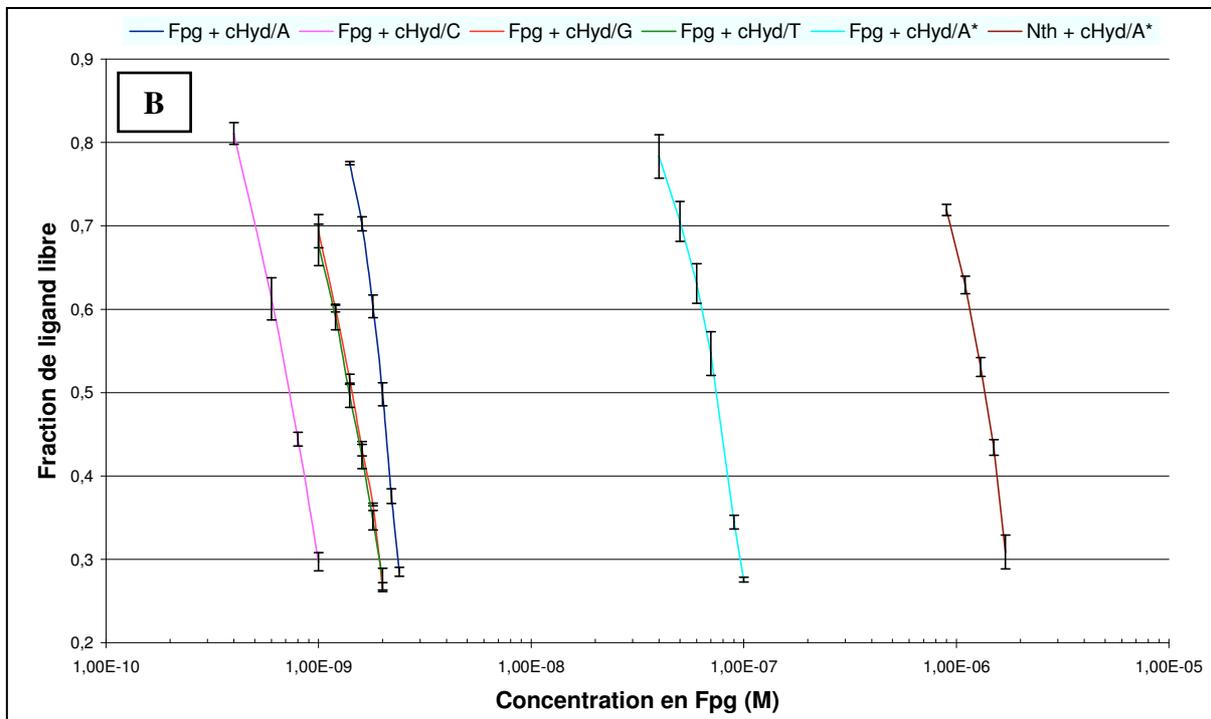
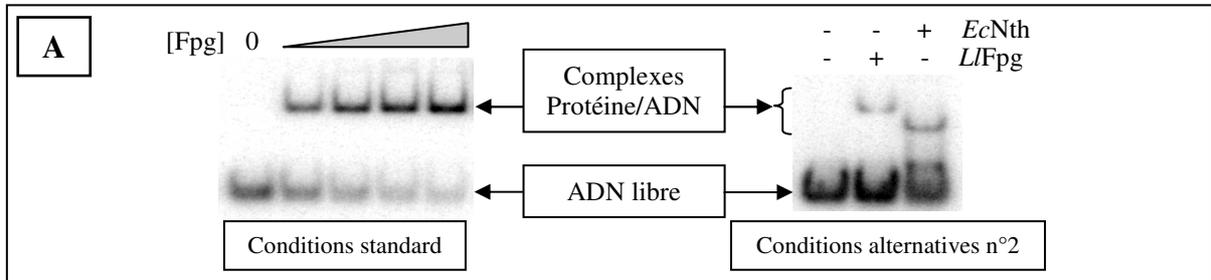
(A) : Exemple de gel de séquence obtenu avec *LIFpg*, le duplexe Hyd/C comme substrat, et le duplexe cHyd/C comme compétiteur. Les notations S et P correspondent respectivement au substrat et au produit de réaction de l'enzyme. La concentration en enzyme est de 12 nM. Les concentrations en ADN compétiteur cHyd/C pour les pistes 3 à 8 sont de 3 ; 6 ; 12 ; 24 ; 48 et 96 nM, respectivement. (B) : Courbes récapitulatives des inhibitions de l'activité Hyd-ADN glycosylase de *LIFpg* obtenues avec les ADN compétiteurs G/C et cHyd/C.



Afin de quantifier l'interaction de *LIFpg* avec le cHyd, nous avons déterminé par électrophorèse en gel-retard les constantes de dissociation apparentes ( $K_{D\text{ app}}$ ) de l'enzyme pour l'Hyd dans différents contextes d'appariement (conditions standard, voir le « Matériel et méthodes ») (Figure 55). De plus, nous avons aussi réalisé des expériences comparatives de ce type sur *EcNth* et *LIFpg* (conditions alternatives n°2, voir le « Matériel et méthodes ») (Figure 55). Quelle que soit la base opposée à l'analogue cHyd, les constantes de dissociation apparentes de *LIFpg* sont toutes de l'ordre du nanomolaire (Figure 55C), ce qui indique que tous ces duplexes d'ADN sont de bons ligands pour l'enzyme. Le meilleur ligand (duplexe cHyd/C, courbe rose, Figure 55B) est environ trois fois mieux reconnu par *LIFpg* que le moins bon ligand (duplexe cHyd/A, courbe bleu-ciel, Figure 55B) ( $K_{D\text{ app}}$  de 0,73 et 1,98 nM, respectivement, Figure 55C). On observe donc une discrimination négative par *LIFpg* pour un A opposé à l'analogue de lésion par rapport à un C pour la fixation à l'ADN, ce qui va dans le même sens que l'activité Hyd-ADN glycosylase de l'enzyme (Figures 51 et 52). L'effet de la base opposée à la lésion sur l'activité Hyd-ADN glycosylase de *LIFpg* est cependant plus important que dans le cas de l'interaction de l'enzyme avec le cHyd (respectivement plusieurs centaines de fois, Figures 51 et 52, contre environ 3 fois, Figure 55C).

**Figure 55 : Analyse de l'interaction de *LIFpg* et *EcNth* avec l'analogue carbocyclique cHyd.**

Les notations /A, /C, /G et /T font référence à des duplexes d'ADN avec respectivement un A, un C, un G et un T en face du cHyd, la notation \* fait référence aux conditions expérimentales alternatives n°2 utilisées pour comparer *EcNth* et *LIFpg* (voir le « Matériel et méthodes »). (A) : Exemples de gel-retards obtenus avec *LIFpg* sur le duplexe cHyd/A dans les conditions standard (concentrations protéiques de 1,8 ; 2 ; 2,2 et 2,4 nM pour les pistes 2 à 5, respectivement), et avec *LIFpg* et *EcNth* sur le duplexe cHyd/A dans les conditions alternatives n°2 (concentrations protéiques de 20 et 300 nM, respectivement). (B) : Courbes de titrage. (C) : Constantes de dissociations apparentes obtenues à partir des courbes précédentes.



**C**

Sonde oligonucléotidique	$K_{D\ app}$ (nM)			
	cHyd/C	cHyd/T	cHyd/G	cHyd/A
<i>LIFpg</i>	0,73	1,40	1,44	1,98 74*
<i>EcNth</i>	ND	ND	ND	1360*

Pour ce qui est de la comparaison de *LIFpg* et *EcNth*, la migration électrophorétique des complexes obtenus est en accord avec les masses moléculaires des enzymes (31 et 23,4 kDa, respectivement, Figure 55A). De plus, *LIFpg* est nettement plus affine du duplexe cHyd/A que *EcNth* ( $K_{D\ app}$  de 1,36  $\mu$ M et 74 nM, respectivement, valeurs notées avec une \*, Figure 55C).

Or, sachant que le duplexe Hyd/A est nettement mieux réparé par *EcNth* que par *LIFpg* (Figure 52), cela indique qu'une bonne fixation n'est pas forcément associée à une bonne activité ADN glycosylase, et inversement.

### II.2.1.3. Bases structurales de la reconnaissance de l'Hyd par LIFpg

#### II.2.1.3.1. Résolution et validation de la structure cristallographique du complexe abortif entre LIFpg et un duplexe d'ADN contenant l'analogue de substrat cHyd opposé à une cytosine

Des cristaux de complexes entre LIFpg et un duplexe cHyd/C ont été obtenus à 20°C (voir le « Matériel et méthodes »). La qualité de diffraction des cristaux a été analysée au laboratoire et les meilleurs candidats ont été sélectionnés pour une collecte de données de diffraction sous rayonnement synchrotron. Un jeu complet de données de diffraction a été enregistré sur la ligne ID 23-1 de l'ESRF (Grenoble) à une longueur d'onde de 0,979 Å avec un cristal qui a diffracté à une résolution maximale de 1,9 Å (Tableau 12). Les statistiques obtenues pour la collection des données de diffraction sont présentées dans le Tableau 11. Les données correspondant au cristal sélectionné sont complètes à 99,5 % (Tableau 12), la part des données de diffraction perdue est donc négligeable.

Source de radiation	ESRF ID 23-1
Longueur d'onde (Å)	0,97935
Nombre total de réflexions	406749
Nombre de réflexions uniques	48244
Groupe d'espace	P 4 <sub>1</sub> 2 <sub>1</sub> 2
Paramètres du cristal	
Longueurs (Å)	a = b = 91,58 c = 142,79
Angles (°)	α = β = γ = 90
Données complètes (%) <sup>a</sup>	99,5 (99,1)
Redondance <sup>a</sup>	8,4 (8,6)
<i>I</i> / $\sigma$ ( <i>I</i> ) moyen <sup>a</sup>	25,9 (7,7)
$R_{\text{sym}}$ (%) <sup>a, b</sup>	5,5 (29,9)

<sup>a</sup> Les valeurs entre parenthèses correspondent aux données de résolution maximale.

<sup>b</sup>  $R_{\text{sym}} = 100 \times \sum |I_i - \langle I \rangle| / \sum I_i$ , où  $I_i$  est l'intensité observée pour une réflexion  $i$ , et  $\langle I \rangle$  est l'intensité moyenne pour plusieurs réflexions liées par symétrie.  $R_{\text{sym}}$  est calculé sur des intensités obtenues avec un même cristal.

**Tableau 11 : Statistiques de collection des données de diffraction.** (complexe non-covalent).

Le problème des phases a été résolu par la méthode de remplacement moléculaire en utilisant la structure de *LIFpg* de type sauvage complexée à un ADN contenant un site cFapyG (code PDB 1XC8). La carte de densité électronique obtenue après remplacement moléculaire

s'est avérée être d'excellente qualité. La chaîne principale de la protéine *LIFpg* a nécessité peu de repositionnements manuels. Par contre, certaines chaînes latérales ont du être ajustées manuellement grâce au programme TURBO-FRODO. Le programme REFMAC a ensuite été utilisé pour l'affinement de la position x, y, z de chacun des atomes ainsi que du facteur B d'agitation thermique. On peut noter que, contrairement aux autres atomes, l'affinement du facteur B du zinc a été réalisé de façon anisotrope car les cartes de densité électronique résiduelles  $mF_o - DF_c$  ont montré une forte anisotropie autour de cet atome. Les valeurs des facteurs R obtenues à la fin de l'affinement du modèle moléculaire (respectivement 16,3 % et 19,5 % pour  $R_{work}$  et  $R_{free}$ ) indiquent que le modèle atomique obtenu est en bonne corrélation avec les données expérimentales, et la faible différence entre ces valeurs (environ 3 %) confirme la validité de notre procédure d'affinement (Tableau 12).

Résolution (Å)	29.85-1.90
Nombre de réflexions utilisées	45759
$R_{work}$ (%) <sup>a</sup>	16.3
$R_{free}$ (%) <sup>a</sup>	19.5
Etude du diagramme de Ramachandran <sup>b</sup>	
Résidus dans les régions favorables (%)	87,18
Résidus dans les régions additionnellement autorisées (%)	12,82
Résidus dans les régions autorisées par extension (%)	0
Résidus dans les régions interdites (%)	0
Valeurs moyennes du facteur B :	
Sur tous les atomes (Å <sup>2</sup> )	20,08
Sur les atomes de la protéine (Å <sup>2</sup> )	15,70
Sur les atomes de l'ADN (Å <sup>2</sup> )	26,26
Sur les atomes de l'eau (Å <sup>2</sup> )	33,52
Ecart type par rapport à l'idéal	
Sur la longueur des liaisons (Å)	0,012
Sur les angles de liaisons (°)	1,527
Sur les angles de torsion (°)	5,4
Nombre d'atomes (atomes d'hydrogènes non comptés) :	
Protéine	2123
ADN	566
Eau	428
Glycérol	6
Total	3123

<sup>a</sup>  $R = 100 \times \frac{\sum |F_{obs} - k| F_{calc}}{\sum |F_{obs}|}$ , avec  $R_{work}$  calculé sur les réflexions utilisées pour l'affinement, et  $R_{free}$  calculé sur les 5 % de réflexions non-utilisées. k est un facteur d'échelle entre  $F_{obs}$  et  $F_{calc}$ .

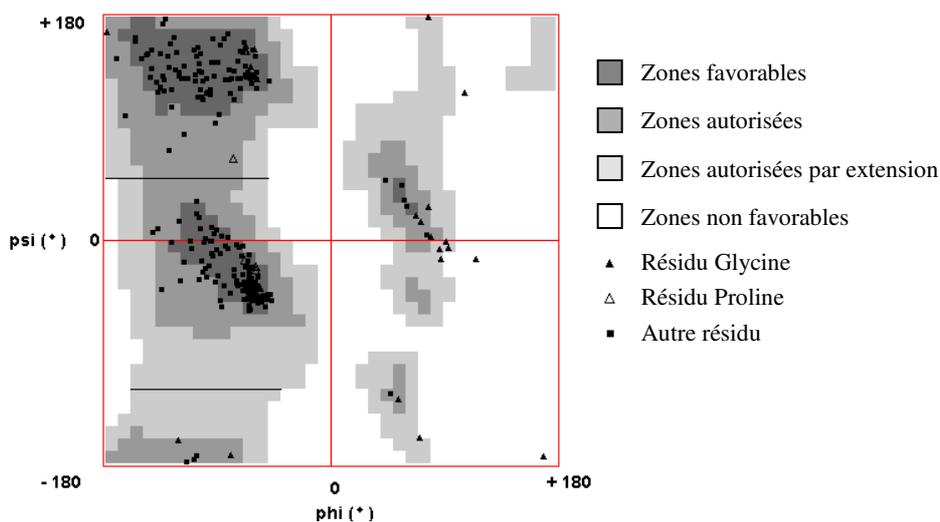
<sup>b</sup> Les données présentées pour le diagramme de Ramachandran ne tiennent pas compte des prolines et glycines de l'enzyme.

**Tableau 12 : Statistiques sur l'affinement et sur le modèle moléculaire.** (complexe non-covalent).

Le modèle atomique final obtenu contient 3123 atomes, dont 2123 atomes correspondant à la protéine, 566 à l'ADN, 428 aux oxygènes de l'eau, et 6 à une molécule de glycérol. Les chaînes latérales des résidus 15, 31, 42, 65, 88, 130, 150, 213, 230 et 271 n'ont pas pu être reconstruites totalement à cause de la mobilité de certains des atomes qui les constituent. Le modèle final obtenu comprend les 28 nucléotides de l'ADN. Par contre, les résidus 219 à 223 de la protéine ne sont pas visibles dans ce modèle car ils sont localisés dans

une boucle de la protéine connue pour être mobile (boucle  $\alpha$ F- $\beta$ 9, Coste *et al.*, 2008), et n'ont donc pas pu être reconstruits.

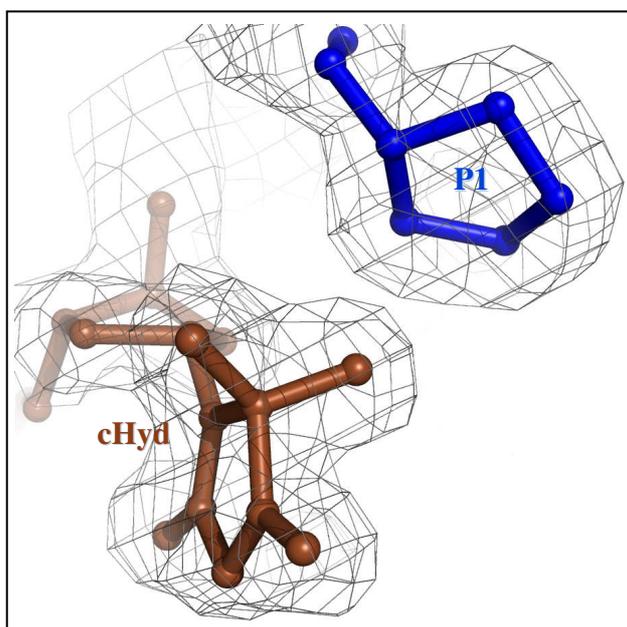
Le diagramme de Ramachandran (Figure 56) obtenu à partir de notre structure de *LIFpg* correspond à un modèle structural protéique de bonne qualité étant donné que l'on trouve environ 87,2 % des résidus dans les zones favorables de ce diagramme, et 100 % dans les zones favorables ou autorisées (les prolines et glycines étant exclues).



**Figure 56 :** Diagramme de Ramachandran sur la structure de Fpg dans le complexe *LIFpg/cHyd* non-covalent. Diagramme déterminé grâce au programme RP (2.0), Gopalakrishnan *et al.*, 2007.

Le modèle structural ne contient aucun résidu avec des angles ou des distances significativement différentes de la valeur idéale (Tableau 12).

La carte de densité électronique  $2F_o - F_c$  contourée à  $1\sigma$  pour la partie base nucléique du *cHyd* et pour le résidu catalytique proline N-terminal de la protéine est bien définie (Figure 57).

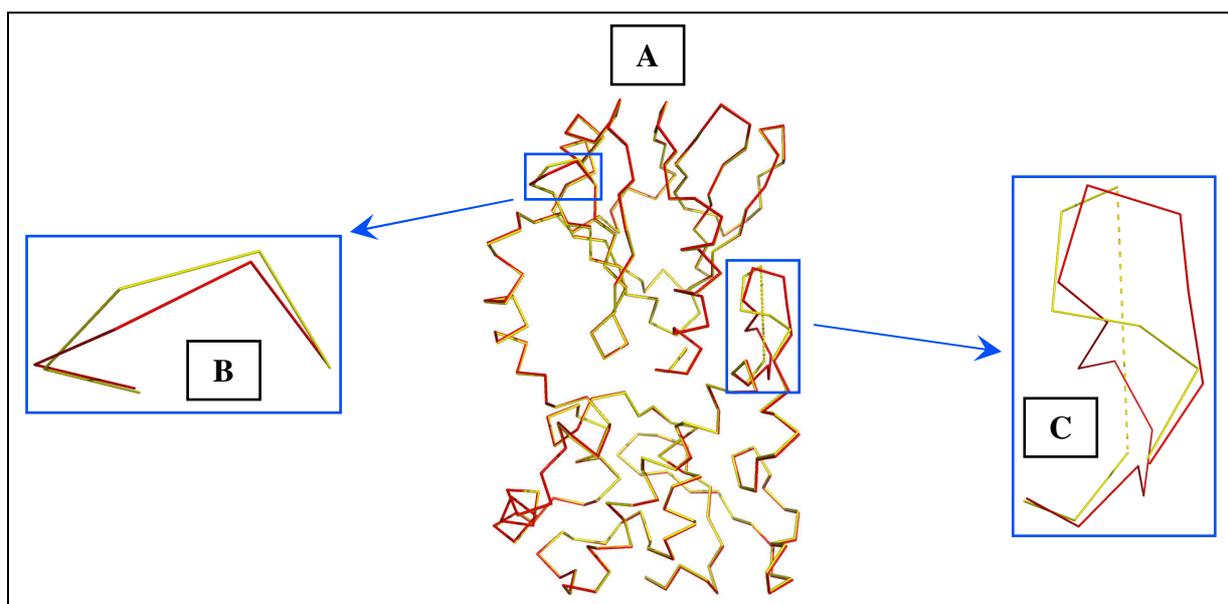


**Figure 57 :** Représentation du modèle structural non-covalent obtenu entre *LIFpg* et le duplexe *cHyd/C*, montrant la lésion 5-OH-5-Me-Hydantoïne et le résidu P1 de Fpg.

Les atomes de notre modèle sont en représentation bâtons, et sont respectivement colorés en marron et en bleu pour la lésion et le résidu P1 de Fpg. La carte de densité électronique  $2F_o - F_c$  contourée à  $1\sigma$  de cette zone est représentée par un maillage gris.

### II.2.1.3.2. Description et analyse de la structure du complexe non-covalent formé entre *L*Fpg et le duplexe d'ADN cHyd/C

Si on compare notre structure à celle de la *Tt*Fpg libre (Sugahara *et al.*, 2000), on ne voit que peu de différences, ce qui indique qu'il n'y a que peu de réarrangements conformationnels de Fpg lors de sa fixation à l'ADN. Les seules variations significatives sont localisées exactement aux mêmes endroits qu'elles avaient été observées dans les structures précédentes (Figure 25, « introduction »). Par contre, la détermination de l'amplitude de ces différences *via* le calcul de l'écart type sur la position des atomes (RMSD) est impossible sur ces deux structures. En effet, celles-ci ayant été réalisées avec des Fpg provenant de différentes bactéries (*Thermus thermophilus* et *Lactococcus lactis*), le nombre et le type de résidus sont différents dans les structures, ce qui rend de tels calculs impossibles. Par contre, on peut comparer notre structure à celle de *L*Fpg complexée à un cFapyG (code PDB 1XC8, Coste *et al.*, 2008). Dans ce cas, l'utilisation d'une Fpg provenant de la même bactérie (*Lactococcus lactis*) pour les deux structures nous permet de réaliser des superpositions plus précises (Figure 58). Ainsi, les structures des squelettes de Fpg complexée à un cFapyG ou un cHyd sont quasi identiques (Figure 58A), avec seulement quelques différences au niveau d'une boucle du domaine N-terminal (résidus 64 à 66, boucle  $\beta$ 3- $\beta$ 4, Figure 58B) et au niveau de la boucle  $\alpha$ F- $\beta$ 9 de la région interdomaine (résidus 217 à 228, Figure 58C), dont certains résidus sont absents de la structure avec le cHyd (résidus 219 à 223).



**Figure 58 : Superposition des squelettes protéiques de *L*Fpg complexée à un cFapyG ou à un cHyd.**

La structure avec le cFapyG est en rouge, celle avec le cHyd est en jaune. La superposition des structures a été réalisée sur les 264 Ca communs entre les deux structures (les résidus 219 à 223 sont absents de la structure avec le cHyd, et les résidus 73 et 121 ont deux conformations dans cette structure). (A) : Vue globale des squelettes protéiques issus des deux structures. (B) : Résidus 63 à 67 de la boucle  $\beta$ 3- $\beta$ 4 issus des deux structures. (C) : Résidus 217 à 229 de la boucle  $\alpha$ F- $\beta$ 9 issus des deux structures.

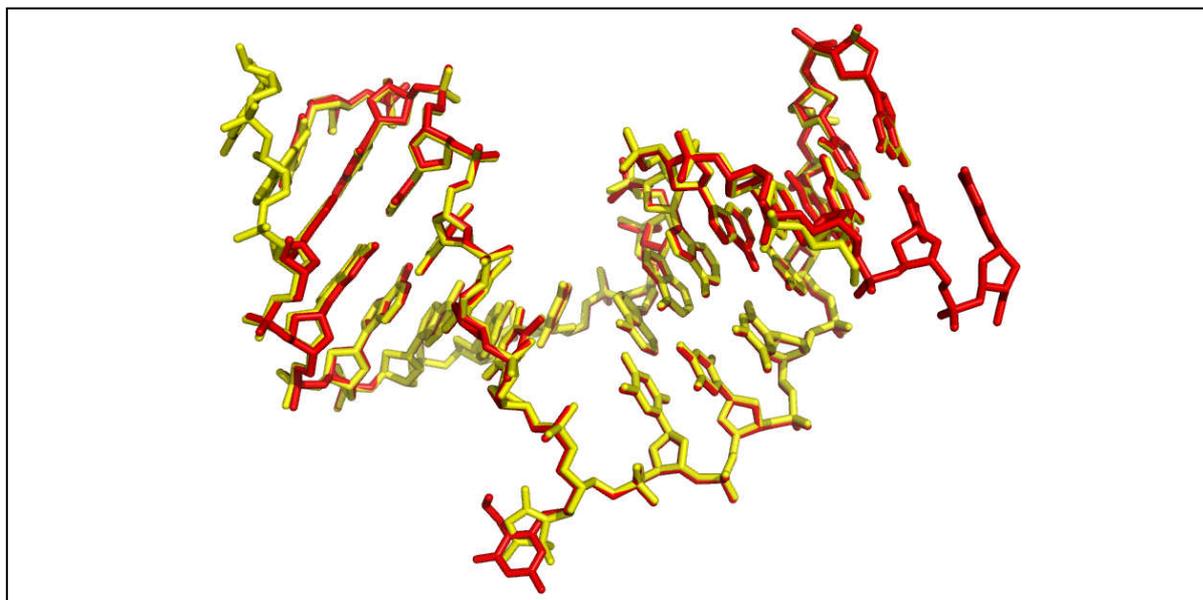
En effet, si on calcule les RMSD après différentes superpositions, on obtient :

- RMSD = 0,597 Å sur les 264 atomes C $\alpha$  comparables. On ne prend pas en compte les résidus 219 à 223, qui sont absents de la structure avec le cHyd, ni les résidus 73 et 121, qui ont deux conformations possibles dans la structure avec le cHyd contre une seule dans celle avec le cFapyG.
- RMSD = 0,152 Å sur 257 atomes C $\alpha$ . Ce calcul est identique au précédent, sauf que l'on ne prend pas en compte les résidus de la région interdomaine dont la position diffère entre les deux structures (résidus 217, 218 et 224 à 228).
- RMSD = 0,136 Å sur 254 atomes C $\alpha$ . Ce calcul est identique au précédent, sauf que l'on ne prend pas non plus en compte les résidus 64 à 66 de la boucle N-terminale dont la position diffère entre les deux structures.

Les principales différences entre les positions des C $\alpha$  de nos deux structures se situent donc bien au niveau des deux boucles évoquées précédemment (Figure 58). En ne tenant pas compte des résidus de ces boucles dans notre calcul, la valeur du RMSD diminue fortement, passant d'une valeur qui correspond à des structures proches (0,597 Å) à une valeur correspondant à deux structures quasi-identiques (0,136 Å). De plus, la plus grande différence par rapport au RMSD global est observée quand on supprime uniquement la région interdomaine (passage de 0,597 Å à 0,152 Å), ce qui indique que cette boucle est la région où la position des C $\alpha$  est la plus différente entre les deux structures. On peut indiquer ici que de légères différences de position de la boucle N-terminale incluant les résidus 64 à 66 ont déjà été observées dans les structures de Fpg obtenues précédemment au laboratoire (Serre *et al.*, 2002; Coste *et al.*, 2004; Pereira de Jésus *et al.*, 2005; Coste *et al.*, 2008). Ces différences ne sont probablement pas liées à la fixation de Fpg sur différents substrats, car cette région ne fait aucune interaction avec l'ADN, et se situe à une distance importante de celui-ci. Par contre, ces différences peuvent probablement être attribuées à l'exposition de cette boucle au solvant ainsi qu'à la présence d'une glycine en position 64, induisant une plus grande mobilité de cette région. D'autre part, des différences de positionnement du squelette protéique dans la boucle  $\alpha$ F- $\beta$ 9 de la région interdomaine ont déjà été observées pour toutes les structures de Fpg obtenues à ce jour. En l'occurrence, certains résidus de cette région sont absents dans de nombreuses structures, indiquant une mobilité forte de ceux-ci. Les seules structures dans lesquelles la totalité des résidus de cette région sont présents sont celles de *Tt*Fpg libre (code PDB 1EE8, Sugahara *et al.*, 2000), de *Bst*FpgE2Q complexée à une 8-oxo-G ou une DHU (codes PDB 1R2Y et 1R2Z, respectivement, Fromme *et al.*, 2003), de *Lt*Fpg de type sauvage ou *Lt*FpgP1G complexée à un THF (codes PDB 1PM5 et 1PJJ, respectivement, Pereira de Jésus

*et al.*, 2005) et la structure de *L/Fpg* complexée à un cFapyG (code PDB 1XC8, Coste *et al.*, 2008).

Les structures des ADN contenant les lésions cFapyG ou cHyd sont très similaires, avec un RMSD de seulement 0,301 Å sur 516 atomes (soit tous les atomes communs entre les deux duplexes d'ADN, sauf la base nucléique lésée) (Figure 59).

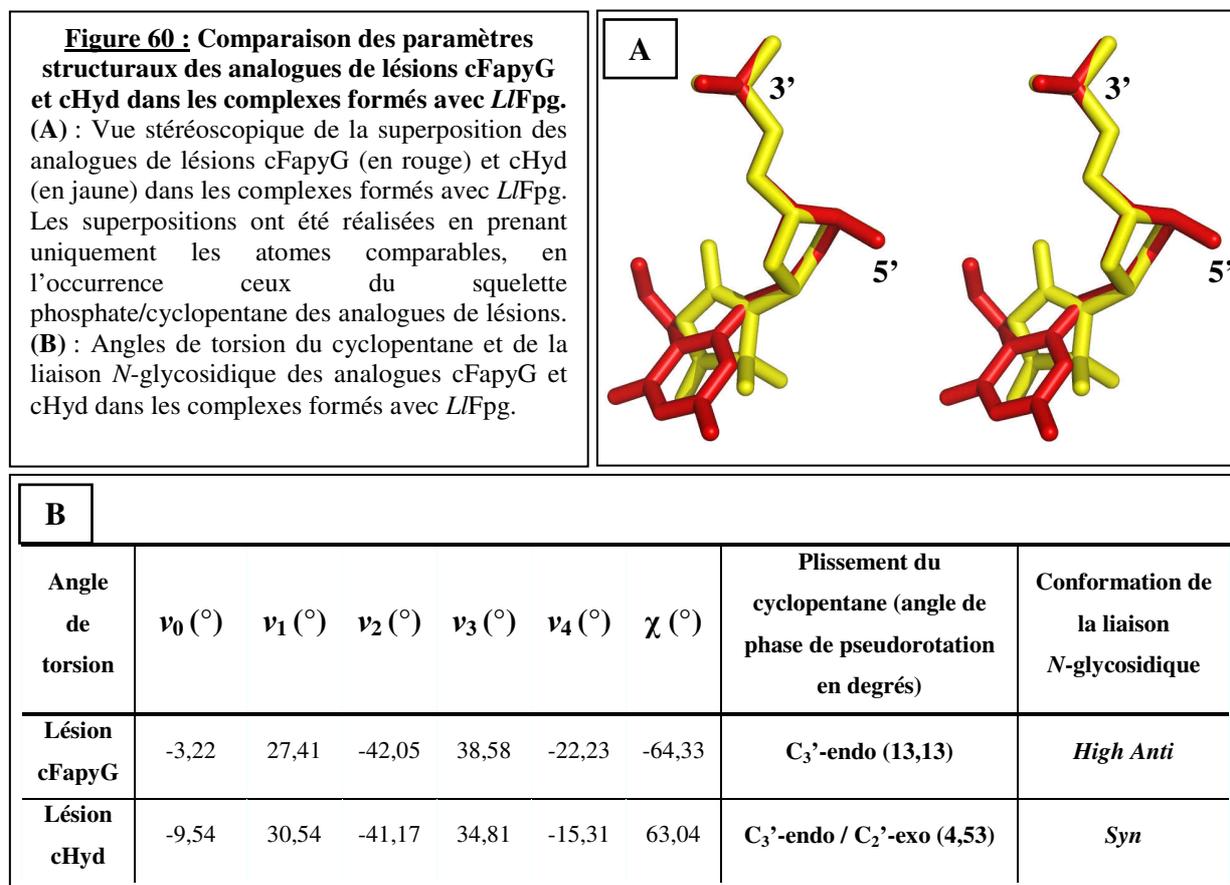


**Figure 59 : Superposition des ADNs contenant les lésions cFapyG (en rouge) et cHyd (en jaune) issues des structures des complexes *L/Fpg/cFapyG* et *L/Fpg/cHyd*.**

Ces superpositions ont été réalisées en prenant les atomes de l'ADN communs entre les deux structures.

Dans les deux structures, l'ADN a subi une forte courbure centrée sur la lésion, et celle-ci est en position extrahélicale. Les seules différences majeures que l'on peut observer correspondent à la partie base nucléique des lésions d'une part, et aux extrémités du brin opposé à la lésion d'autre part. Les différences pour les extrémités de l'ADN ne sont pas biologiquement pertinentes, étant donné qu'il n'y a aucune interaction de ces régions avec la protéine. Au niveau des lésions, les squelettes phosphate/sucre des lésions sont dans une conformation quasi-identique, avec un RMSD de seulement 0,091 Å sur les 11 atomes de cette partie des lésions (Figure 60). De plus, les angles de torsion du cyclopentane (Figure 100, « Matériel et méthodes) dans les deux structures sont très proches, définissant cependant un plissement légèrement différent (voir « Matériel et méthodes »). Dans le cas du cFapyG, le cyclopentane adopte un plissement C<sub>3</sub>'-endo, contre un intermédiaire entre C<sub>3</sub>'-endo et C<sub>2</sub>'-exo avec le cHyd. D'autre part, si on considère la conformation de la liaison N-glycosidique (Définie grâce à l'angle  $\chi$ : Figure 100, « Matériel et méthodes), une différence majeure est

observée entre les deux analogues de lésions, avec une conformation *high anti* pour le cFapyG ( $\chi = -64,33^\circ$ ), et *syn* pour le cHyd ( $\chi = 63,04^\circ$ ) (Figure 60B).

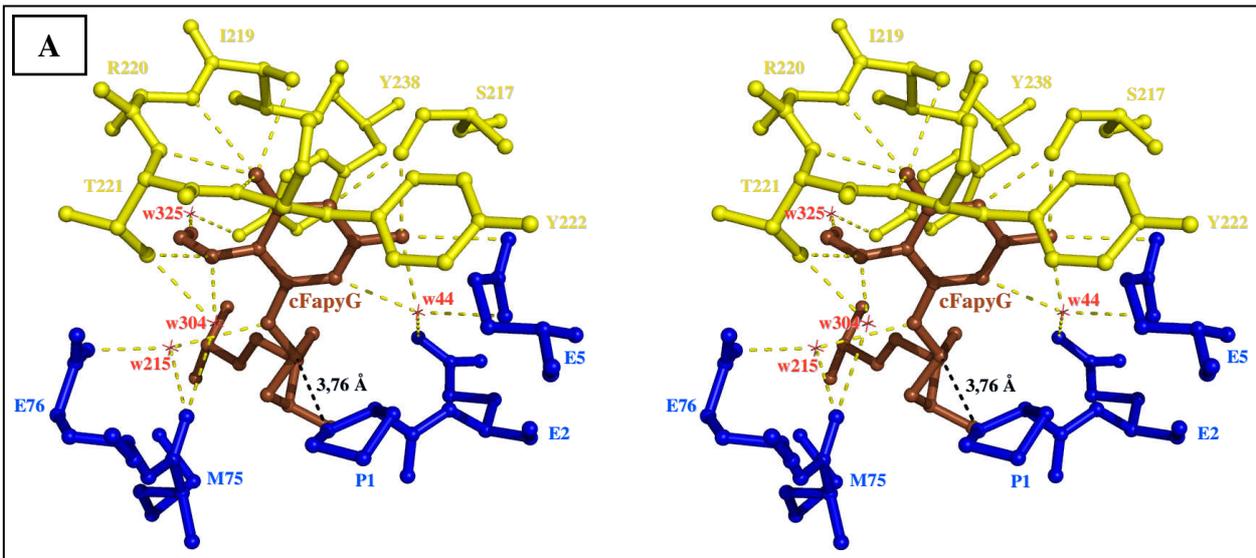


Les liaisons hydrogène présentes entre LIFpg et les ADN contenant les analogues de lésions cFapyG ou cHyd ont été répertoriées (voir le « Matériel et méthodes »). Ces réseaux d'interactions sont assez peu différents pour le cFapyG et le cHyd. Notamment, nous n'avons observé aucune différence significative entre les réseaux d'interactions présents entre Fpg et les squelettes phosphate/sucres des ADN, pas plus qu'au niveau de la reconnaissance par le résidu R109 de la cytosine opposée aux analogues de lésions (Serre *et al.*, 2002). Par contre, des différences significatives ont été vues pour les réseaux de liaisons hydrogène établis entre la protéine et la partie base nucléique des deux analogues de lésions (Figure 61 adaptée de Coste *et al.*, 2008, et Figure 62). Ainsi, 20 liaisons hydrogène sont dénombrées entre le cFapyG et la protéine, dont 8 directes entre des résidus de la protéine et l'analogue de lésion, et 12 participant à des interactions médiées par une seule molécule d'eau (Coste *et al.*, 2008). Aucune interaction *via* plus d'une molécule d'eau n'a été détectée. De plus, ces interactions font intervenir des résidus de domaines différents de la protéine, avec une majorité de résidus de la région interdomaine (6 résidus impliqués, en jaune dans la Figure G) par rapport aux

résidus du domaine N-terminal (4 résidus impliqués, en bleu dans la Figure G). Dans la structure du complexe *LlFpg/cHyd*, au total 20 liaisons hydrogène sont établies entre le cHyd et la protéine, dont une seule directe entre un résidu de la protéine et l'analogue de lésion, et 2 participant à des interactions médiées par une seule molécule d'eau. Toutes les autres liaisons sont médiées par plus d'une molécule d'eau. Comme précédemment, ces interactions font intervenir des résidus de domaines différents de la protéine, cependant ici on a une implication majoritaire de résidus du domaine N-terminal (6 résidus, en bleu dans la Figure 62) par rapport à la région interdomaine (2 résidus, en jaune dans la Figure 62). On peut donc noter que les résidus de Fpg et le réseau de molécules d'eau impliqués dans la reconnaissance de la base nucléique endommagée sont différents avec le cFapyG et le cHyd, et que seuls les résidus E5 et M75 de l'enzyme sont impliqués dans la reconnaissance des deux analogues de lésions.

**Figure 61 : Réseau de liaisons hydrogène établi entre *LF*Fpg et la partie base nucléique du cFapyG.**

D'après Coste *et al.*, 2000. Les résidus de la protéine sont soumis au même code couleur que pour les Figures 23 et 24 de l'introduction du manuscrit, suivant la définition de Sugahara *et al.*, 2000. Le cFapyG et les molécules d'eau sont représentés en marron et en rouge, respectivement. (A) : Vue stéréoscopique du réseau de liaisons hydrogène entre Fpg et la partie base nucléique du cFapyG. Les liaisons hydrogène sont représentées par des pointillés jaunes. La distance entre l'azote du résidu P1 de Fpg et le C1' du cFapyG est indiquée en noir, et est représentée par un pointillé noir. (B) : Liste des liaisons hydrogène établies entre Fpg et la partie base nucléique du cFapyG. Pour chaque liaison, la distance indiquée correspond à celle entre le donneur et l'accepteur de liaison hydrogène (D et A dans la Figure 103 du « Matériel et méthodes »).

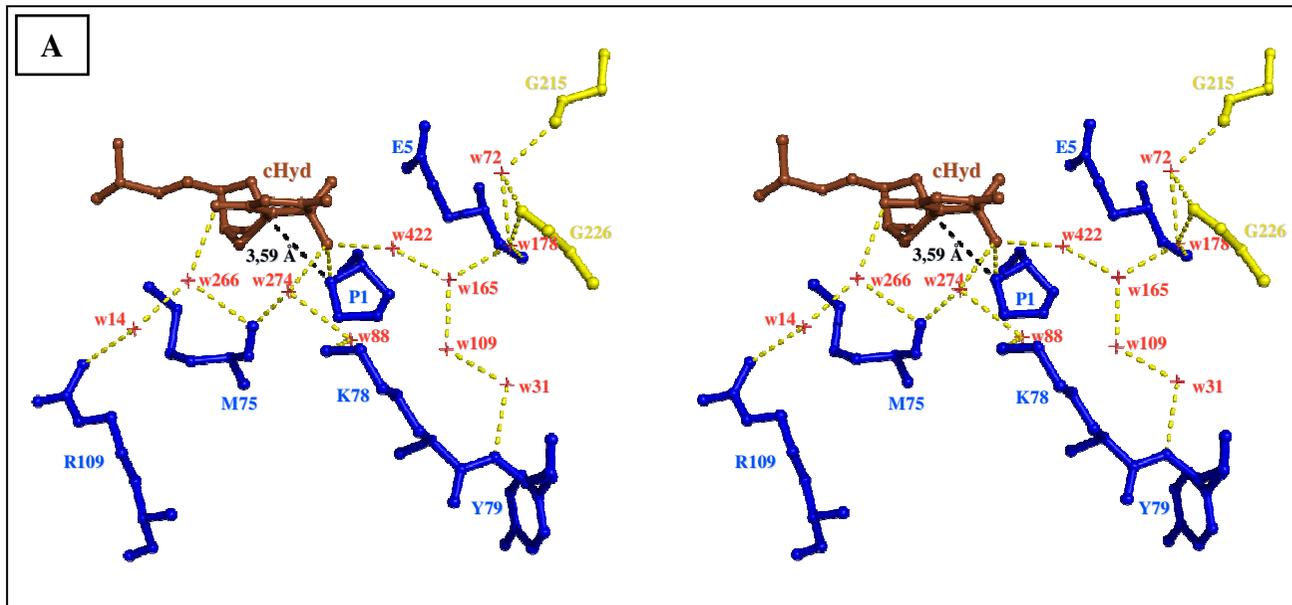


**B**

Résidu	Atome	Distance (Å)	Résidu	Atome	Distance (Å)	Résidu	Atome
Liaisons hydrogène directes entre Fpg et la lésion cFapyG :							
cFapyG	N1	2,66	S217	O			
cFapyG	N2	2,77	S217	O			
cFapyG	N2	3,06	E5	OE2			
cFapyG	O6	2,92	I219	N			
cFapyG	O6	2,89	R220	N			
cFapyG	O6	3,13	T221	N			
cFapyG	O6	3,05	Y222	N			
cFapyG	N7	2,86	T221	OG1			
Liaisons hydrogène entre Fpg et la lésion cFapyG médiées par une molécule d'eau :							
cFapyG	N2	2,72	w44	O	3,1	E5	OE1
cFapyG	N3	2,81	w44	O	2,85	E2	OE2
cFapyG	N7	2,73	w304	O	2,67	T221	OG1
					3,13	M75	O
cFapyG	N9	2,94	w215	O	2,89		
					2,70	E76	OE2
cFapyG	O8	2,72	w325	O	2,76	Y238	OH

**Figure 62 : Réseau de liaisons hydrogène établi entre L/Fpg et la partie base nucléique du cHyd.**

Les résidus de la protéine sont soumis au même code couleur que pour les Figures 23 et 24 de l'introduction du manuscrit, suivant la définition de Sugahara *et al.*, 2000. Le cHyd et les molécules d'eau sont représentés en marron et en rouge, respectivement. (A) : Vue stéréoscopique du réseau de liaisons hydrogène entre Fpg et la partie base nucléique du cHyd. Les liaisons hydrogène sont représentées par des pointillés jaunes. La distance entre l'azote du résidu P1 de Fpg et le C1' du cHyd est indiquée en noir, et est représentée par un pointillé noir. (B) : Liste des liaisons hydrogène établies entre Fpg et la partie base nucléique du cHyd. Pour chaque liaison, la distance indiquée correspond à celle entre le donneur et l'accepteur de liaison hydrogène (D et A dans la Figure 103 du « Matériel et méthodes »).



**B**

Résidu	Atome	Distance (Å)	Résidu	Atome												
cHyd	O5	2,68	P1	N												
cHyd	O2	3,01	w266	O	2,71	w14	O	2,96	R109	NH1						
					3,08	M75	O									
cHyd	O5	2,91	w274	O	2,70											
					2,93	w88	O	2,88	K78	NZ						
		2,72	w422	O	2,42	w165	O	2,64	w109	O	2,64	w31	O	2,79	Y79	N
								2,65			2,87	E5	O			
									w178	O	2,99	G226	O			
											2,80	w72	O	2,55	G226	O

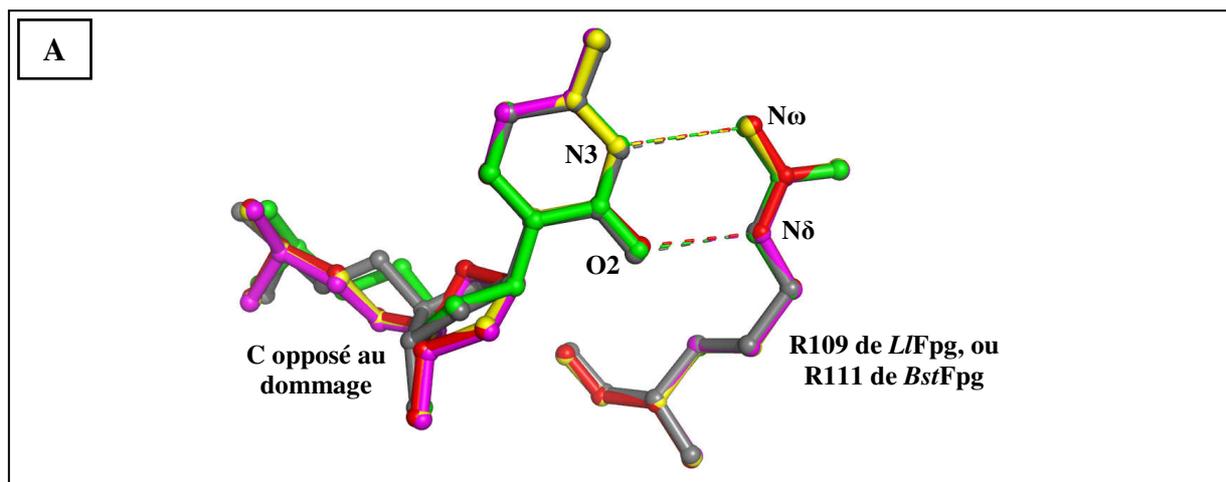
#### II.2.1.4. Conclusion et discussion

La structure cristallographique que nous avons résolue entre *L/Fpg* et le duplexe cHyd/C est la première structure obtenue d'une Fpg de type sauvage complexée à un ADN contenant une pyrimidine oxydée. En effet, une autre structure de Fpg en complexe avec un ADN contenant une Dihydrouracile (DHU) est déjà disponible dans la PDB, mais implique la *BstFpg* de type « mutant inactif » E2Q (code PDB 1R2Z, Fromme *et al.*, 2003a).

L'analyse fonctionnelle présente dans ce chapitre indique que les lésions Hyd et 8-oxo-G sont toutes les deux mieux excisée par *L/Fpg* (et *EcNth*) dans un ADN double-brin plutôt que dans ADN un simple-brin (Figures 51 et 52). Cela est tout à fait en accord avec les données de la bibliographie, qui indiquent que ces ADN-glycosylases bifonctionnelles sont spécifiques de l'excision des lésions contenues dans des ADNs double-brins (Breimer *et al.*, 1984; Tchou *et al.*, 1991). Cela n'est en fait pas étonnant sachant que la voie de réparation BER nécessite un brin intact opposé au brin endommagé pour permettre la resynthèse de l'information génétique par une ADN polymérase (chapitre I.5, Introduction). Sachant que certaines lésions sont mutagènes, beaucoup d'ADN glycosylases possèdent une spécificité quant aux bases appariées aux lésions qu'elles réparent, afin d'éviter de fixer de façon définitive une mutation. Ainsi, lors de notre analyse fonctionnelle sur l'Hyd, nous avons observé un effet fort de la base opposée au dommage sur la métabolisation de la lésion par *L/Fpg*. En effet, le duplexe d'ADN Hyd/A étant un substrat nettement moins bon pour l'enzyme que le duplexe Hyd/C (Figures 51 et 52). Cette observation est d'ailleurs corrélée aux données de fixation de l'enzyme sur l'analogue de lésion cHyd (Figure 55), avec cependant un effet de la base opposée au dommage plus faible sur la fixation que sur l'activité Hyd-ADN glycosylase. Ces données sont en accord avec celles de la littérature, qui indiquent qu'une lésion opposée à un C est toujours mieux excisée par Fpg que lorsqu'elle est opposée à un A, et cela quelle que soit la lésion considérée (Tableau 7 de l'introduction, Tchou *et al.*, 1991; Castaing *et al.*, 1993; Tchou *et al.*, 1994; Asagoshi *et al.*., 2000; Leipold *et al.*, 2000; Speina *et al.*, 2001; Wiederholt *et al.*, 2002a; Wiederholt *et al.*, 2003; Wuenschell *et al.*, 2003; Zaika *et al.*, 2004). Fonctionnellement, cela est tout à fait en accord avec le rôle principal émis pour Fpg *in vivo*, en l'occurrence lutter contre les effets mutagènes de la 8-oxo-G (chapitre I.6.1 de l'introduction sur système GO procaryote, Michaels *et al.*, 1992b).

On peut alors se demander quels sont les déterminants structuraux permettant à Fpg d'avoir une telle spécificité pour la base opposée aux lésions qu'elle excise. Un déterminant

structural important permettant à l'enzyme de reconnaître la base opposée au dommage a été découvert en 2002, et consiste en deux liaisons hydrogène entre l'Arginine 109 de *L/Fpg* et une cytosine opposée à un 1,3-propanediol (analogue non métabolisable de site abasique) (code PDB 1KFV, Serre *et al.*, 2002). La R109 de *L/Fpg* fait partie de la triade d'intercalation de l'enzyme (M75, R109, F111, Figure 27 de l'introduction), qui est une signature toujours conservée dans la sous-famille structurale des Fpgs. Quelle que soit l'origine de l'enzyme (*E. coli*, *L. lactis*, *B. stearothermophilus*) et quel que soit le dommage considéré, toutes les structures cristallographiques de complexes Fpg/ADN dans lesquelles la base nucléique lésée est opposée à un C révèlent ces mêmes interactions entre l'arginine conservée et la base opposée au dommage (Figure 63). Ainsi, la R109 de *L/Fpg* et la R111 de *BstFpg* réalisent chacune deux liaisons hydrogènes avec la cytosine opposée aux différentes lésions ou analogues de lésions (cHyd, cFapyG, Bz-cFapyG, 8-oxo-G et DHU, Figure 63). Dans tous les cas, la cytosine opposée au dommage est maintenue intrahélicale par ces interactions, et dans une conformation *anti*. La longueur et l'orientation des deux liaisons hydrogène, ainsi que les donateurs et accepteurs de proton impliqués dans ces interactions, sont toujours identiques (Figure 63B). Cependant, toutes ces structures ont été réalisées avec un ADN contenant une cytosine opposée au dommage. Ces structures ne sont donc pas informatives pour comprendre structurellement l'effet sur *L/Fpg* de la base opposée au dommage que nous avons observé pour l'activité Hyd-ADN glycosylase et pour l'interaction de l'enzyme avec les ADNcHyd/X (X = A, T, G, C).

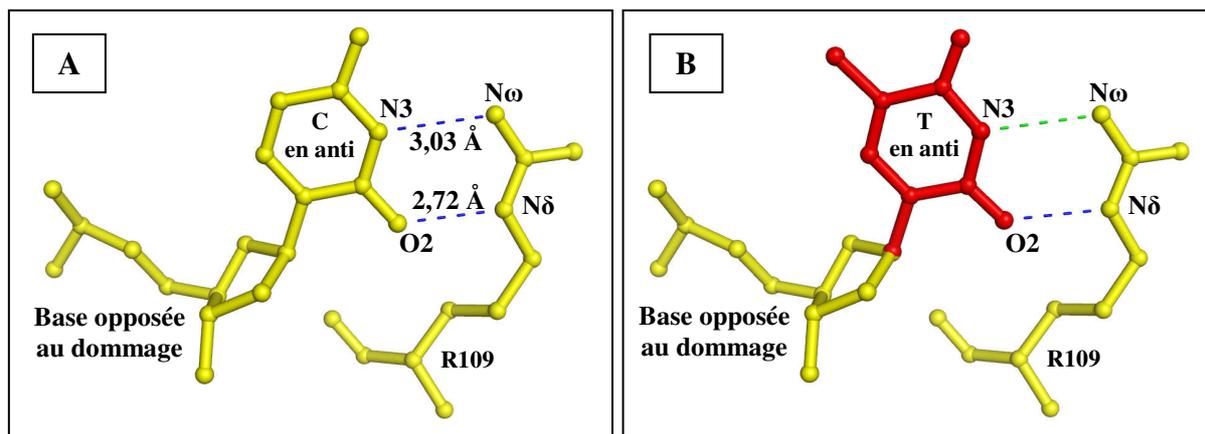


**B**

Structure	RMSD avec la structure de <i>LIFpgWT+cHyd</i> (Å)	Longueur de la liaison hydrogène entre Nϵ de l'Arg et N3 du C (Å)	Longueur de la liaison hydrogène entre Nδ de l'Arg et O2 du C (Å)
<i>LIFpgWT+cHyd</i>	0	3,03	2,72
<i>LIFpgWT+cFapyG</i>	0,042	3,08	2,70
<i>LIFpgWT+Bz-cFapyG</i>	0,054	3,08	2,74
<i>BstFpgE2Q+8-oxo-G</i>	0,117	2,98	2,72
<i>BstFpgE2Q+DHU</i>	0,125	3,03	2,78

**Figure 63 : Reconnaissance par une arginine de Fpg de la de la cytosine opposée au dommage.**  
Toutes les structures connues de Fpg complexée non covalamment à un dommage de base nucléique ont été superposées sur notre structure de *LIFpg* complexée non covalamment au cHyd, en ne superposant que la partie base nucléique de la cytosine, et le résidu arginine entier (soit 19 atomes au total). Les structures sont : *LIFpgWT+cHyd* (en jaune, ce travail), *LIFpgWT+cFapyG* (en rouge, code PDB 1XC8, Coste *et al.*, 2008), *LIFpgWT+Bz-cFapyG* (en magenta, code PDB 3C58, Coste *et al.*, 2008), *BstFpgE2Q+8-oxo-G* (en vert, code PDB 1R2Y, Fromme *et al.*, 2003a), *BstFpgE2Q+DHU* (en gris, code PDB 1R2Z, Fromme *et al.*, 2003a). (A) : Représentation des superpositions précédentes. Les liaisons hydrogènes sont représentées par des pointillés. (B) : Tableau donnant les RMSD pour chaque superposition précédente avec notre structure de *LIFpgWT/cHyd*, et donnant les longueurs des liaisons hydrogène entre le Nϵ de l'Arg et le N3 du C, et entre le Nδ de l'Arg et le O2 du C.

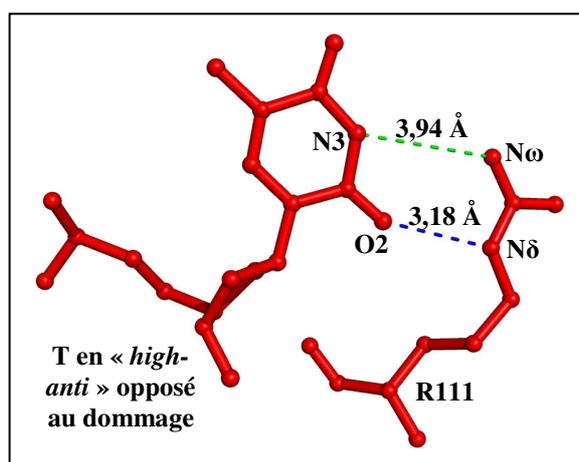
Afin de tenter d'évaluer les contraintes structurales (stériques et électrostatiques) qui pourraient être associées à ces effets sur la fixation et la catalyse, nous avons réalisé des reconstructions de bases nucléiques à la place de la cytosine opposée au dommage dans notre structure de *LIFpg* fixée sur un duplexe cHyd/C. Ainsi, la modélisation d'une thymine en *anti* ( $\chi = -137,62^\circ$ ) en face du dommage ne pose pas de problème stérique particulier pour la reconnaissance par la R109 de l'enzyme (Figure 64).



**Figure 64 : Reconstruction d'une thymine en *anti* dans notre structure de *LFpg/cHyd*.**

Les atomes issus de notre structure de *LFpg/cHyd* sont représentés en jaune, ceux de la thymine reconstruite sont représentés en rouge. La cytosine opposée au dommage est en *anti* dans notre structure de *LFpg/cHyd* ( $\chi = -137,62^\circ$ ), et on a reconstruit une thymine dans une conformation identique. Les liaisons hydrogènes sont représentées par des pointillés bleus, une interaction de type Van Der Waals est représentée par un pointillé vert. (A) : Cytosine opposée au dommage et R109 issus de notre structure de *LFpg/cHyd*. (B) : Vue identique au (A), mais avec reconstruction d'une thymine en *anti* à la place de la cytosine opposée au dommage.

Par contre, il est probable que l'interaction avec une thymine opposée au dommage soit nettement moins forte qu'avec une cytosine à cette position. En effet, dans ce cas on a disparition d'une des deux liaisons hydrogènes au profit d'une interaction de type Van Der Waals (beaucoup plus faible), avec même une possibilité de répulsion entre les deux NH présents en N3 de la thymine et en N $\omega$  de l'arginine (Trait pointillé vert, Figure 64B). Ce modèle de complexe entre *LFpg* et un duplexe *cHyd/T* est d'ailleurs en parfait accord avec la structure cristallographique de *BstFpg* complexée à un duplexe d'ADN contenant un site AP réduit opposé à une thymine (code PDB 1L2C, Fromme *et al.*, 2002) (Figure 65). Dans cette structure expérimentale, la thymine opposée au site AP réduit est stabilisée en *high-anti* ( $\chi = -101,77^\circ$ ), et est en interaction avec *BstFpg* via l'Arg 111 par une liaison hydrogène et une interaction de Van Der Waals. La répulsion entre les NH présents en N3 de la thymine et en N $\omega$  de l'arginine induit cependant un léger éloignement des deux partenaires en interaction (Figures 63 et 65). De plus, les données structurales indiquent que cette thymine présente une agitation thermique (facteur B) nettement plus importante que pour le reste de la structure, et pour la cytosine opposée au même dommage dans la structure de *BstFpg* complexée au duplexe AP-red/C (code PDB 1L1T, Fromme *et al.*, 2002). Ceci témoigne de la faible intensité de l'interaction entre l'arginine de *Fpg* et une thymine opposée au dommage par rapport à une cytosine à cette position.

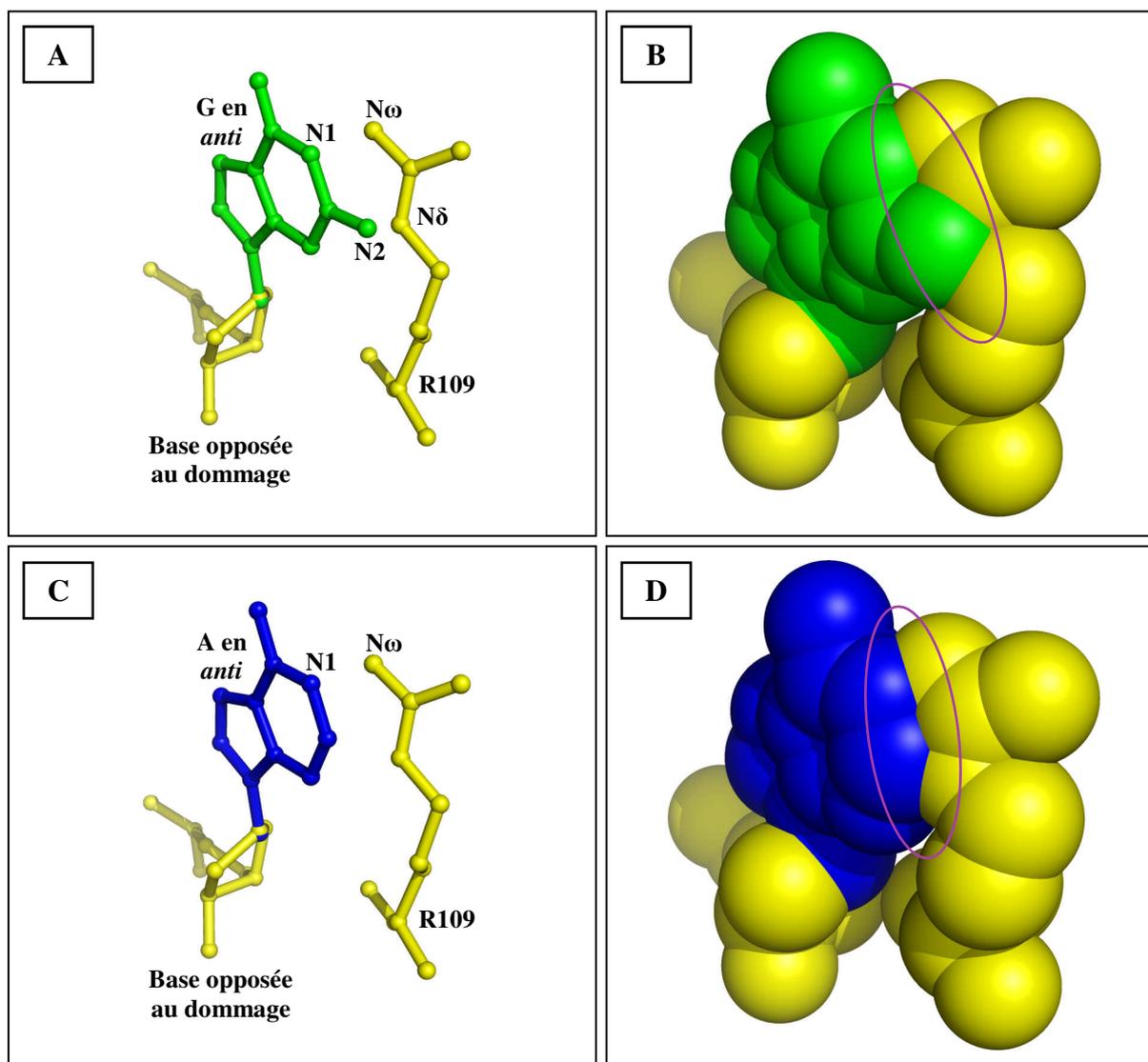


**Figure 65 : Représentation de la thymine en *anti* dans la structure de *BstFpg* complexée à un site AP réduit opposé à une thymine.**

Issu de Fromme *et al.*, 2002, code PDB 1L2C. La thymine opposée au site AP réduit et l'Arg 111 qui la contacte sont représentés en rouge dans une vue similaire à celle de la Figure 64. La thymine opposée au dommage est en « *high-anti* » dans la structure ( $\chi = -101,77^\circ$ ). La liaison hydrogène est représentée par un pointillé bleu, l'interaction de type Van Der Waals est représentée par un pointillé vert, et leur longueur respective est indiquée.

La moins forte stabilisation par Fpg d'un T que d'un C en face du dommage peut être corrélée à la différence de constante de dissociation apparente que nous avons mesurée, avec une valeur environ deux fois plus importante pour le duplexe cHyd/T par rapport au duplexe cHyd/C (Figure 55). Par contre, on pourrait s'attendre à ce que cela soit associé à une activité moins forte de l'enzyme sur les lésions opposées à une thymine par rapport aux mêmes lésions opposées à une cytosine. Cependant ce n'est pas toujours le cas (Tableau 7 de l'introduction, Tchou *et al.*, 1991; Tchou *et al.*, 1994; Asagoshi *et al.*, 2000; Speina *et al.*, 2001; Zaika *et al.*, 2004). En fait, outre des effets structuraux lors de la stabilisation de la base en conformation extrahélicale, il n'est pas exclu que des effets de l'environnement séquentiel du dommage et/ou des effets dus à la nature du dommage lui-même (jouant sur la déformabilité/stabilité de la paire de base au niveau du dommage ou dépendant de la présence de déterminants structuraux du dommage dans l'un des sillons de l'ADN, etc) se surajoutent. D'autre part, on peut préciser que l'accommodation d'un T en *syn* est impossible, car dans cette conformation il n'y aurait pas de groupement donneur/accepteur de liaison hydrogène possible pour faire l'interaction avec l'arginine conservée de Fpg. La conformation *anti* de la thymine opposée au dommage est donc indiscutablement la forme qui est stabilisée par Fpg.

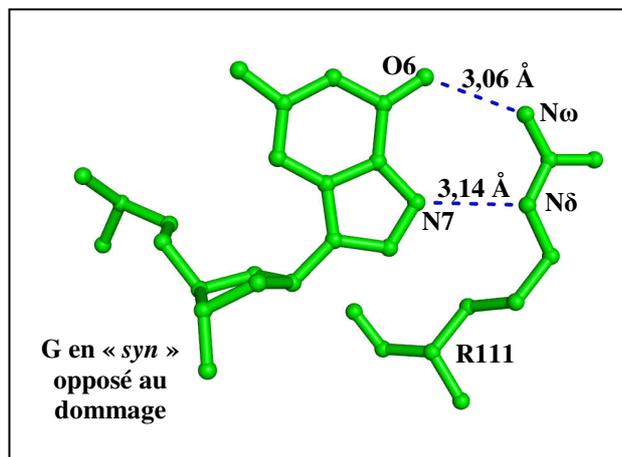
Contrairement à ce que nous venons de voir, la reconstruction d'une purine (A ou G) en *anti* ( $\chi = -137,62^\circ$ ) dans notre structure de *LIFpg* complexée au duplexe cHyd/C est associée à un problème stérique avec R109 (Figure 66).



**Figure 66 : Reconstruction d'une guanine ou d'une adénine en *anti* dans notre structure de *LIFpg/cHyd* : apparition d'un problème stérique.**

Les atomes issus de notre structure de *LIFpg/cHyd* sont représentés en jaune, ceux de la guanine reconstruite sont représentés en vert, et ceux de l'adénine reconstruite sont représentés en bleu. Les bases nucléiques opposées au dommage ont été reconstruites en *anti* dans une conformation identique à celle que l'on avait pour la cytosine opposée au dommage dans notre structure de *LIFpg* complexée au *cHyd*. La localisation d'un problème stérique est indiquée en magenta. **(A)** : Représentation en bâtonnets (facilite la vision de la molécule). La distance entre le N1 du G et le N $\omega$  de R109 est d'environ 1,6 Å. La distance entre le N2 du G et le N $\delta$  de R109 est d'environ 0,9 Å. **(B)** : Vue identique au (A), mais en représentation en sphères dont le rayon correspond au rayon de covalence de chaque atome (facilite la vision de l'emplacement du problème stérique). **(C)** : Représentation en bâtonnets. La distance entre le N1 du G et le N $\omega$  de R109 est d'environ 1,5 Å. **(D)** : Vue identique au (C), mais en représentation en sphères.

Dans ce cas, la stabilisation en *syn* des purines opposées aux dommages par *LIFpg* serait plus favorable que leur stabilisation en *anti*. Cette hypothèse est d'ailleurs en parfait accord avec une structure de *BstFpg* complexée à un duplexe d'ADN contenant un site AP réduit opposé à une guanine (code PDB 1L2D, Fromme *et al.*, 2002) (Figure 67).

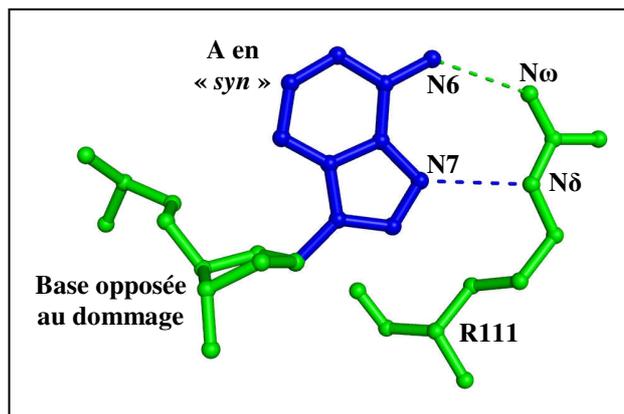


**Figure 67 :** Représentation de la guanine en *syn* dans la structure de *BstFpg* complexée à un site AP réduit opposé à une guanine.

Issu de Fromme *et al.*, 2002, code PDB 1L2D. La guanine opposée au site AP réduit et l'Arg 111 qui la contacte sont représentés en vert dans une vue similaire à celle de la Figure 64. La thymine opposée au dommage est en « *syn* » dans la structure ( $\chi = 61,84^\circ$ ). Les liaisons hydrogène sont représentées par des pointillés bleus, et leur longueur est indiquée.

Dans cette structure expérimentale, la guanine opposée au site AP réduit est stabilisée en *syn* ( $\chi = 61,84^\circ$ ) par *BstFpg*, avec création de deux liaisons hydrogène entre le O6 de la guanine et le N $\omega$  de R111 de l'enzyme (longueur de 3,06 Å), et entre le N7 de la guanine et le N $\delta$  de R111 de l'enzyme (longueur de 3,14 Å). Ces interactions permettent donc une stabilisation forte de la guanine opposée au site AP réduit par *BstFpg*. A ce niveau, il est donc probable que *L/Fpg* stabilise aussi la guanine en *syn* via deux liaisons hydrogènes avec R109, et cela sans générer de problème stérique. Par contre, la stabilisation forte par *Fpg* d'une guanine opposée à une lésion, et faible d'une thymine opposée à une lésion, devrait induire une meilleure réparation par l'enzyme d'un duplexe lésion/G par rapport à un duplexe lésion/T. C'est en fait rarement le cas (Tableau 7 de l'introduction, Tchou *et al.*, 1991; Castaing *et al.*, 1993; Tchou *et al.*, 1994; Asagoshi *et al.*, 2000; Speina *et al.*, 2001). Comme précédemment, il est évident que d'autres effets jouent sur la spécificité de base opposée à la lésion de *Fpg* en plus de la stabilisation *via* cette arginine.

Enfin, une adénine opposée au dommage pourrait aussi être stabilisée en *syn* par la R109 de *L/Fpg* (ou la R111 de *BstFpg*) (Figure 68).



**Figure 68 :** Reconstruction d'une adénine en *syn* dans la structure de *BstFpg* complexée à un ADN contenant un site AP réduit opposé à une guanine.

Les atomes issus de la structure de *BstFpg*/(APred/G) (code PDB 1L2D, Fromme *et al.*, 2002) sont représentés en vert, ceux de l'adénine reconstruite sont représentés en bleu. L'adénine reconstruite opposée au dommage est en *syn* ( $\chi = 61,84^\circ$ ) dans une conformation identique à celle de la guanine initialement présente dans la structure. Une probable liaison hydrogène est représentée par un pointillé bleu, une probable interaction de type Van Der Waals est représentée par un pointillé vert.

Cependant, on a un NH<sub>2</sub> en position 6 de l'adénine, d'où le probable remplacement de la liaison hydrogène avec le N $\omega$  de R111 de *BstFpg* par une interaction faible de type Van Der Waals. Comme nous l'avons vu avec une thymine opposée au dommage (Figure 65), il est probable qu'il y ait en fait une répulsion entre le NH<sub>2</sub> en position 6 de l'adénine et le NH en position  $\omega$  de R111. Cette répulsion serait associée à un éloignement des deux partenaires en interaction. L'interaction de Fpg avec l'adénine est donc probablement très faible, ce qui peut expliquer nos résultats concernant la faible activité de *L/Fpg* sur le duplexe Hyd/A par rapport au duplexe Hyd/C. De même, cela peut expliquer la faible activité sur les lésions opposées à un A par rapport à celles opposées à un C ou un G que l'on retrouve dans la littérature (Tableau 7 de l'introduction et références correspondantes).

Par contre, l'interaction de l'arginine de Fpg étant probablement aussi faible avec un T en *anti* qu'ici avec un A en *syn*, les duplexes lésion/A et lésion/T devraient être des substrats aussi mauvais l'un que l'autre pour Fpg, et devraient toujours être les deux plus mauvais substrats de l'enzyme. Or, les données de la littérature montrent qu'un T opposé à une lésion est toujours un substrat nettement meilleur qu'un A opposé à la même lésion (Tableau 7 de l'introduction et références correspondantes). La meilleure activité de Fpg sur les duplexes lésion/T par rapport aux duplexes lésion/A ne peut donc s'expliquer structurellement uniquement par l'interaction avec R109 de *L/Fpg* ou R111 de *BstFpg*. Comme évoqué précédemment, d'autres paramètres jouent probablement sur la spécificité de base opposée à la lésion de Fpg. Cela illustre que l'analyse structurale ne permet pas toujours des

extrapolations réalistes relatives aux propriétés fonctionnelles d'une enzyme, d'autant plus si ces données structurales sont incomplètes. En effet, la résolution d'une structure cristallographique de Fpg complexée sur un ADN avec un A en face d'une lésion permettrait peut-être de pointer du doigt d'autres déterminants structuraux expliquant la forte discrimination négative de l'enzyme pour une adénine opposée au dommage. Cependant, à ce jour, tous les essais de notre équipe pour cristalliser de tels complexes ont avorté. Il faut noter également que toute l'analyse structure/activité que nous venons de présenter pour comprendre l'effet de la base opposée à la lésion suppose dans tous les cas que la base lésée a été extrudée hors de la double-hélice d'ADN et stabilisée en position extrahélicale dans la poche du site actif de Fpg. Ainsi, dans le cas d'une adénine opposée à la lésion, il n'est pas exclu que l'enzyme soit incapable d'extruder le dommage hors de la double hélice d'ADN (ou que cette extrusion soit très inefficace), et par conséquent incapable de le métaboliser par attaque nucléophile en C1' du dommage extrahélicale. Une telle situation a été suggérée pour la 8-oxo-G appariée à une adénine. Dans ce cas, la 8-oxo-G adopte une conformation *syn* dans la double hélice d'ADN, lui permettant de former une paire de base stable entre sa face Hoogsteen et la face Watson-Crick de l'adénine en *anti* (Figure 18A de l'introduction) (Kouchakdjian *et al.*, 1991; McAuley-Hecht *et al.*, 1994). Dans le cas d'une paire 8-oxo-G/C, les deux partenaires sont en conformation *anti*, et sont tous deux en interaction *via* leur face Watson-Crick (Figure 18A de l'introduction) (Oda *et al.*, 1991; Lipscomb *et al.*, 1995). Ainsi, les déterminants de la 8-oxo-G (N7 protoné et C8 oxydé) sont exposés dans le grand sillon de l'ADN pour la paire 8oxo-G/C. Par contre, le groupement 8-oxo de la paire 8-oxo-G/A se retrouve exposé dans le petit sillon de l'ADN, et le groupement N7-H se retrouve engagé dans une liaison hydrogène de type Hoogsteen avec l'adénine. Une hypothèse avait alors été émise, qui indiquait que Fpg pouvait sélectionner les 8-oxo-G à extruder ou non *via* la présence ou l'absence dans le grand sillon des déterminants structuraux précédents (Tchou *et al.*, 1994; Lipscomb *et al.*, 1995). En conséquence, modéliser un A en face de la lésion extrudée en se basant sur nos structures est peut-être purement théorique et n'a pas forcément de sens du point de vue physiologique.

Contrairement à *LI*Fpg, la base opposée au dommage influence peu la métabolisation de l'Hyd par *Ec*Nth. Le duplexe d'ADN Hyd/A est un substrat légèrement meilleur pour l'enzyme qu'un duplexe Hyd/C (Figure 52). Nos résultats sont en accord avec les données de la littérature, qui indiquent qu'une lésion opposée à un A est toujours mieux excisée par Nth

que quand celle-ci est opposée à un C, et cela quelle que soit la lésion considérée (Tableau 13).

**Tableau 13 : Différences d'activité de Nth en fonction de la base opposée à la lésion pour différentes lésions.**

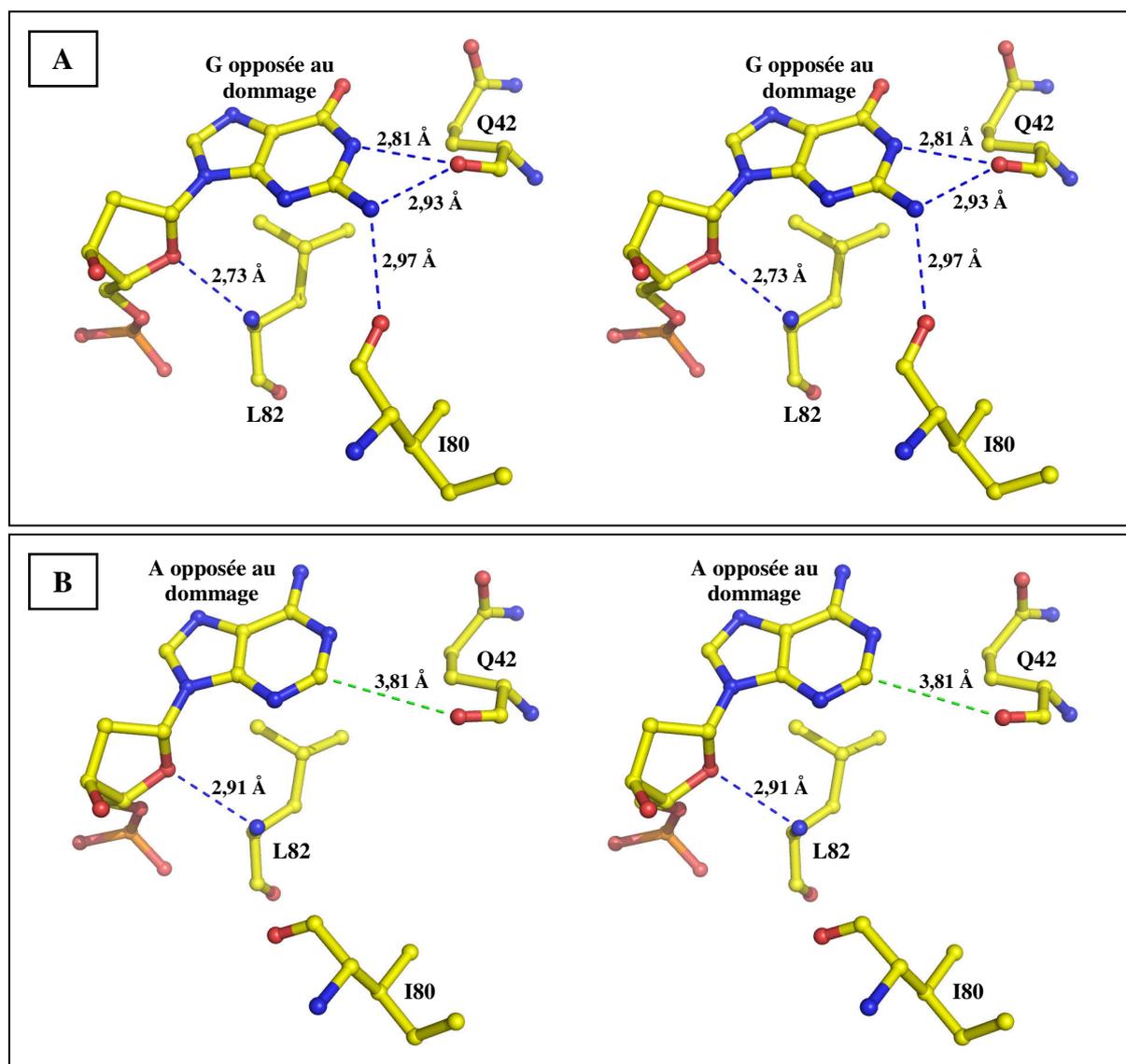
La notation peu explicite de certaines lésion ou analogues de lésions, ainsi que l'origine des lésions est :  $N^3$ - (2-acide hydroxyisobutyrique)urée (Ac. HI-Urée, lésion issue de thymine); Xanthine (lésion issue de guanine); 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG, lésion issue de guanine); 7,8-dihydro-8-oxo-Guanine (8-oxo-G, lésion issue de guanine); 4,6-diamino-5-formamidopyrimidine (FapyA, lésion issue d'adénine).

Lésion	Ordre de préférence pour la base opposée à la lésion	Publication de référence
Ac. HI-Urée	$dG > dA \geq dT \geq dC$ , donc $dA > dC$	Jurado <i>et al.</i> , 1998
FapyG	$dG \geq dA > dT \geq dC$ , donc $dA > dC$	Wiederholt <i>et al.</i> , 2005
8-oxo-G	$dG > dA > dT \geq dC \approx 0$ , donc $dA > dC$	Matsumoto <i>et al.</i> , 2001
FapyA	$dG > dA > dC > dT$ , donc $dA > dC$	Wiederholt <i>et al.</i> , 2005

D'une façon plus générale, une lésion appariée à une purine est toujours un meilleur substrat pour Nth qu'une lésion appariée à une pyrimidine. Fonctionnellement, cela est tout à fait en accord avec la spécificité se substrat de Nth, car cette enzyme est décrite comme étant spécifique de l'excision des lésions pyrimidiques (Dizdaroglu *et al.*, 1993; Hatahet *et al.*, 1994; Fromme *et al.*, 2003b; Dizdaroglu *et al.*, 2005; Huffman *et al.*, 2005), qui sont naturellement appariées à des purines. De plus, les données d'activité de Nth (Tableau 13) sont en accord avec des données d'efficacité de fixation de *Bst*Nth sur un ADN contenant un site THF (analogue non métabolisable de site abasique), qui indiquent que l'ordre de préférence de base opposée au dommage est  $dG \geq dA \gg dC \approx dT$  (Fromme *et al.*, 2003b).

Un déterminant structural important permettant à l'enzyme de reconnaître la base opposée au dommage a été découvert récemment grâce à la résolution des structures cristallographiques de *Bst*Nth complexée covalamment à des ADNs contenant un site AP opposé à un G ou un A (Fromme *et al.*, 2003b) (Figure 69). Ainsi, une guanine en face de la lésion est bien reconnue par Nth *via* quatre liaisons hydrogènes. Par contre, une adénine opposée au dommage est reconnue, mais nettement moins bien qu'une guanine, avec formation d'une seule liaison hydrogène et d'une interaction faible de type Van Der Waals) (Figure 69). Cela peut expliquer qu'une guanine en face d'une lésion soit un cas plus favorable à l'activité de Nth qu'une adénine à cette position (Tableau 13). Par contre, la stabilisation forte par Nth d'une pyrimidine en face d'une lésion paraît difficile même si il n'y a pas de problème stérique. En effet, la distance entre la base nucléique opposée au dommage

et les résidus Q42 et I80 est trop importante, ce qui rend impossible les interactions entre une pyrimidine et ces résidu sans avoir de réarrangement conformationnel majeur (Figure 69).



**Figure 69 : Reconnaissance de la base opposée au dommage par *Bst*Nth.**

Issu des structures de *Bst*Nth complexé covalentement à des ADNs contenant un site AP par réduction de la base de Schiff (Fromme *et al.*, 2003b). Les carbones sont représentés en jaune, les oxygènes en rouge, les azotes en bleu, et les phosphates en orange. Les liaisons hydrogène sont représentées par des pointillés bleus, et une interaction de type Van Der Waals par un pointillé vert. (A) : Complexe avec une guanine opposée à la lésion. (B) : Complexe avec une adénine opposée à la lésion.

Cette observation pourrait d'ailleurs être associée au fait que l'Hyd est légèrement mieux réparée par *Ec*Nth en face d'une adénine qu'en face d'une cytosine (Figure 55). De plus, cela permet d'expliquer qu'une pyrimidine en face d'une lésion est plus défavorable à l'activité de l'enzyme qu'une purine à cette position (Tableau 13). Par ailleurs, l'interaction entre le carbonyle de la chaîne principale du résidu L82 de Nth et le O4' du désoxyribose

portant la base opposée au dommage est une interaction totalement non spécifique qui permet à l'enzyme de stabiliser à cette position n'importe quelle base nucléique. En outre, sachant qu'une adénine opposée au dommage ne réalise qu'une interaction de Van Der Waals en plus de la liaison hydrogène avec L82 (Figure 69), on peut penser qu'une cytosine opposée à une lésion ne serait pas beaucoup moins bien stabilisée par l'enzyme qu'une adénine. Cela permettrait d'expliquer que l'effet d'un A ou d'un C opposé à l'Hyd est plus faible sur l'activité Hyd-ADN glycosylase de *EcNth* que sur celle de *LIFpg* (Figure 52). En effet, *LIFpg* reconnaît les bases opposées aux dommages uniquement *via* des interactions spécifiques (Figures 63 et 64), qui sont cependant nettement plus faibles avec une adénine qu'avec une cytosine.

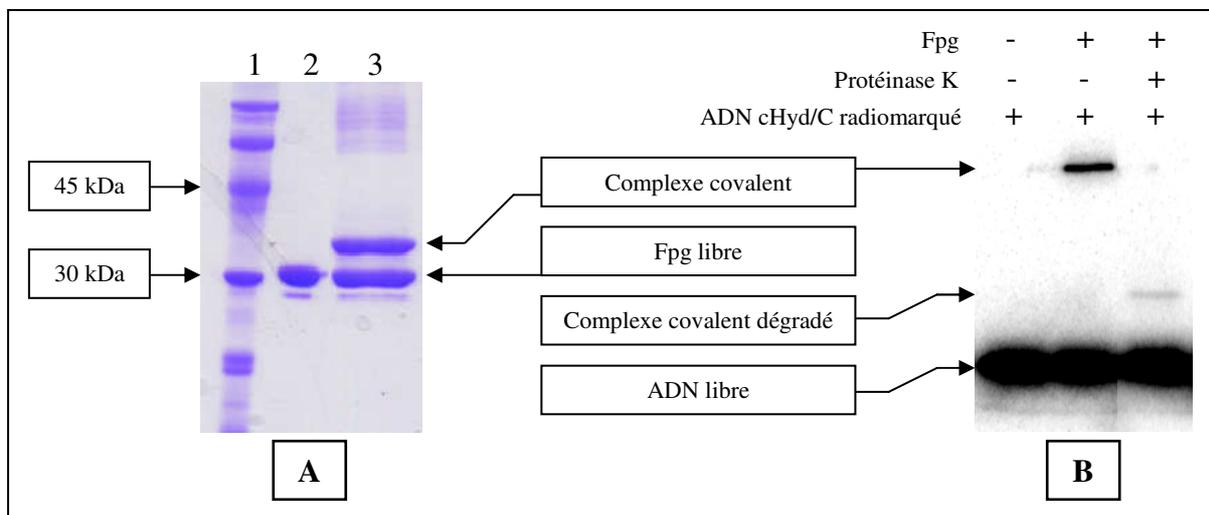
Enfin, nos données indiquent que le duplexe Hyd/A est nettement moins bien réparé par *LIFpg* que par *EcNth* (Figure 52). Or seul l'appariement de l'Hyd avec une adénine est physiologiquement relevant, car cette lésion est non mutagène et est issue de l'oxydation de la thymine (Figure 50; Halliwell *et al.*, 1991; Gasparutto *et al.*, 2000). Cela semblerait indiquer que la réparation de l'Hyd *in vivo* est probablement assurée par Nth (et possiblement aussi Nei) plutôt que par Fpg. Cependant, des données bibliographiques ont montré que cet effet est très réduit si on compare la réparation d'un duplexe Hyd/A par *EcFpg* et *EcNth* (différence de  $k_{cat}/K_M$  de seulement 1,6 fois, Gasparutto *et al.*, 2009). Ainsi, il semble que les deux enzymes puissent être impliquées dans la réparation de l'Hyd *in vivo* chez *E. coli*. Ces différences impliquent donc que pour réellement pouvoir conclure quant à l'enzyme majoritairement responsable de l'excision de l'Hyd *in vivo* chez *L. lactis*, il aurait bien évidemment fallu que l'on compare les Fpg et Nth provenant de cette bactérie. Par contre, nos résultats peuvent vouloir dire qu'il y a une différence de spécificité de substrat entre *EcFpg* et *LIFpg*. De plus, l'activité de Fpg sur l'Hyd appariée à un A, même si faible, peut être utile *in vivo* en cas de défaillance de l'enzyme Nth (et Nei) ou en cas d'apparition d'une trop grande quantité de lésions de ce type.

D'autre part, notre étude fonctionnelle a montré que le duplexe Hyd/A est beaucoup moins bien réparée par *LIFpg* que le duplexe 8-oxo-G/C (Figure 51). Cependant, des données bibliographiques ont montré que cet effet est très réduit chez *EcFpg* (différence de  $k_{cat}/K_M$  de seulement 1,5 fois, Gasparutto *et al.*, 2009). Cela peut à nouveau vouloir dire qu'il y a une différence de spécificité de substrat entre *EcFpg* et *LIFpg*, avec par exemple une plus forte discrimination de la base opposée au dommage par *LIFpg*.

## II.2.2. Formation d'un complexe covalent abortif entre l'Hyd et LIFpg

### II.2.2.1. Mise en évidence et caractérisation préliminaire du complexe covalent

Lors des études structurales sur le complexe formé entre *LIFpg* et le duplexe *cHyd/C*, nous avons obtenu des cristaux à 20 et à 37°C (voir Matériel et méthodes). La structure du complexe abortif présentée précédemment dans le chapitre II.2.1.3 a été réalisée à partir des cristaux obtenus à 20°C. Bien qu'observables à des concentrations supérieures en agent cristallisant que celles utilisées à 20°C (voir Matériel et méthodes), les cristaux obtenus à 37°C présentent un faciès similaire à ceux obtenus à 20°C, avec cependant une taille bien supérieure (600×200×200 µm et 200×100×50 µm, respectivement). Du point de vue de l'état de la protéine, les cristaux obtenus à 20 et 37°C varient significativement. En effet, l'analyse SDS-PAGE du contenu de chaque type de cristal indique sans ambiguïté que ceux obtenus à 37°C contiennent une espèce protéique de poids moléculaire supérieur à celui de la protéine libre, et cela à hauteur d'environ 50% (Piste 3, Figure 70A).

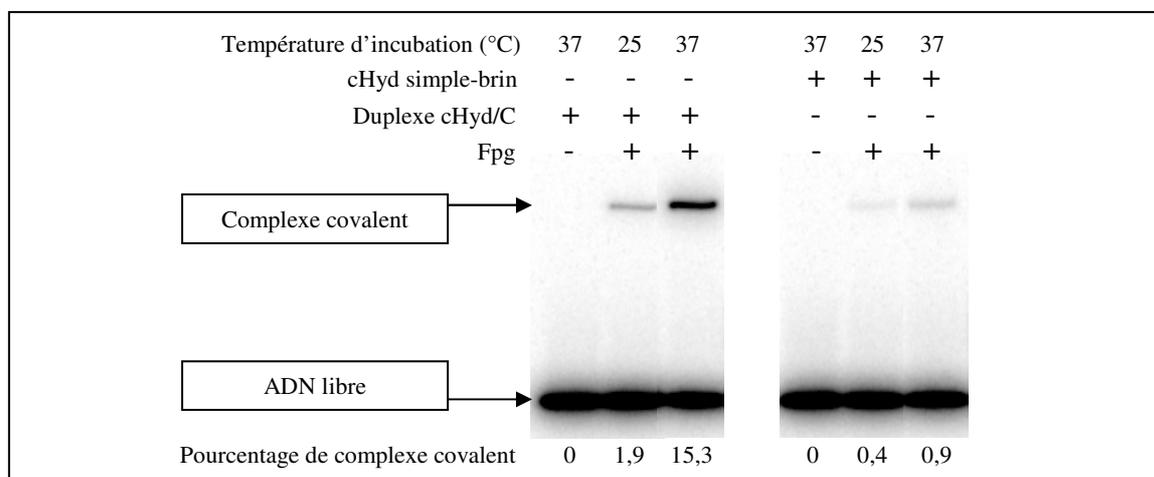


**Figure 70 :** Mise en évidence d'un complexe covalent entre *LIFpg* et le duplexe *cHyd/C*.

(A) : Analyse par SDS-PAGE des cristaux dissouts de *LIFpg* complexée au duplexe *cHyd/C*. (1) : Marqueurs de tailles. (2) : Cristal obtenu à 20°C dissout. (3) : Cristal obtenu à 37°C dissout. (B) : Analyse par SDS-PAGE-TRAP du complexe covalent formé entre *LIFpg* et le duplexe *cHyd/C* (voir Matériel et méthodes). Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). La concentration finale en *Fpg* pour les essais déposés sur les pistes 2 et 3 est de 12,5 µM. Les complexes ont été incubés 18 heures à 37°C, puis l'essai déposé dans la piste 3 a subi une digestion de 30 min à 37°C en présence de protéinase K.

D'après les marqueurs de poids moléculaires (Piste 1, Figure 70A), l'espèce protéique de plus grande taille ( $\approx 35$  kDa) pourrait correspondre à un complexe entre *LIFpg* ( $\approx 31$  kDa) et le brin d'ADN contenant la lésion ( $\approx 4,2$  kDa). Ensuite, nous avons démontré que cette espèce de haut poids moléculaire correspondait bien à un complexe nucléoprotéique par incubation de protéine *LIFpg* en présence d'ADN radiomarqué (Figure 70B). Ainsi, l'analyse par SDS-PAGE-TRAP (autoradiographie de gels SDS-PAGE, voir le Matériel et méthodes) du mélange d'incubation indique clairement que c'est le brin lésé qui est associé covalamment à l'enzyme (seul ce brin est radiomarqué) (Piste 2, Figure 70B), et la sensibilité du complexe à la protéinase K confirme l'intervention d'un partenaire protéique dans ce complexe (Piste 3, Figure 70B).

Comme nous l'avons observé lors des essais de cristallogenèse (Figure 70A), ce complexe covalent est fortement favorisé lorsque les deux partenaires du complexe sont incubés à 37°C plutôt qu'à 25°C (Figure 71). On n'observe d'ailleurs plus du tout de complexe lors d'une incubation à 4°C (données non présentées). De plus, l'analyse par SDS-PAGE-TRAP indique que la réaction de pontage covalent est favorisée par la structure bicaténaire de l'ADN, et cela quelle que soit la température de formation du complexe (Figure 71).

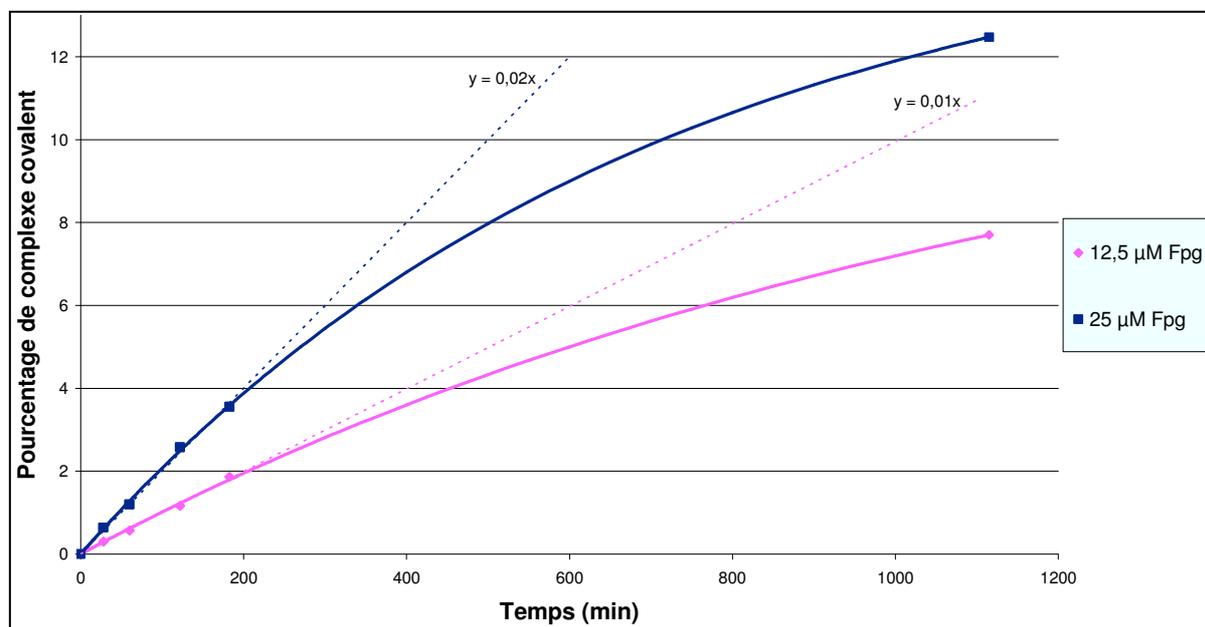


**Figure 71 : Effet de la température d'incubation et de la nature mono- ou bi-caténaire de l'ADN contenant le cHyd sur la formation du complexe covalent avec *LIFpg*.**

Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 25 ou 37°C, et la concentration finale en Fpg est de 12,5  $\mu$ M. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes).

Ainsi, on observe environ 4,8 fois plus de complexe covalent avec un ADN double-brin plutôt que simple-brin à 25°C, et 17 fois plus à 37°C (Figure 71). D'autre part, pour le duplexe cHyd/C, la réaction de pontage covalent est environ 8 fois plus efficace à 37°C qu'à 25°C (Figure 71).

Même si la réaction de pontage n'est pas très véloce, la formation de complexe covalent est directement dépendante du temps d'incubation et de la concentration en enzyme, au moins sur les trois premières heures de réaction (Figure 72).



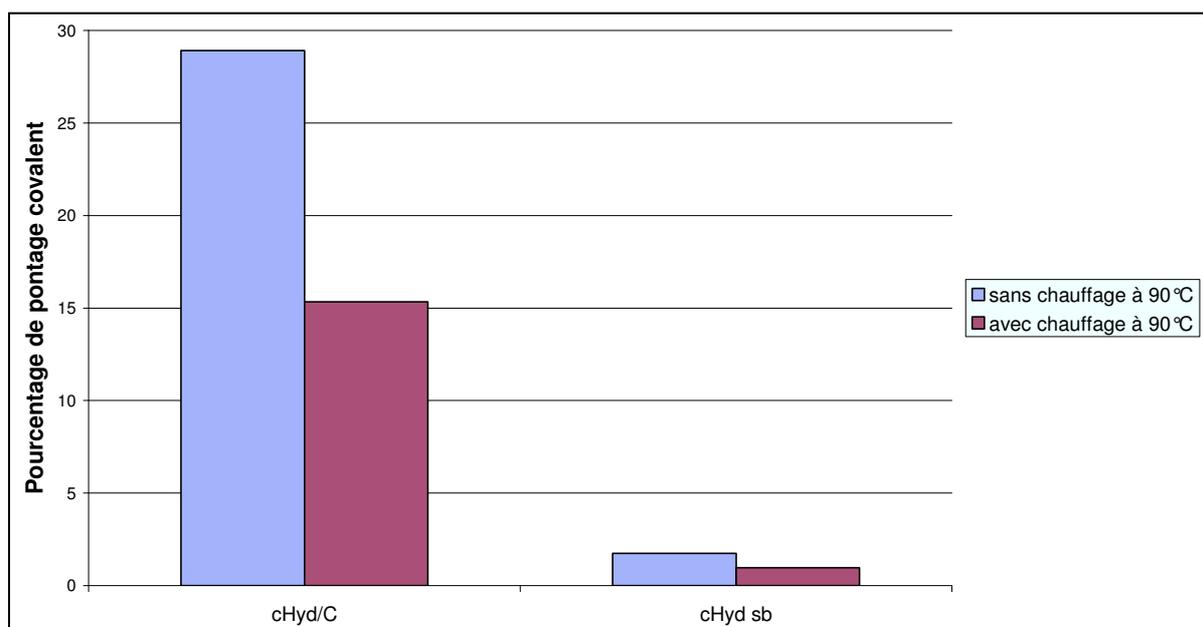
**Figure 72 : Cinétiques de formation de complexe covalent entre *LIFpg* et le duplexe *cHyd/C* à 37°C.**

Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Deux cinétiques ont été réalisées à 37°C, avec des concentrations finales en Fpg de 12,5 et 25 µM. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes). Des ajustements linéaires sur les données allant jusqu'à 183 minutes de réaction ont été ajoutés au graphique représentatif des résultats, et sont représentés par des pointillés. Les équations correspondant à ces ajustements sont indiquées.

En effet, la vitesse initiale de formation de complexe covalent est multipliée par 2 lorsqu'on double la concentration en *LIFpg* dans les essais (traits pointillés, Figure 72). On est donc bien en présence d'une réaction catalysée par l'enzyme.

Ensuite, nous nous sommes intéressés à la caractérisation de la stabilité de nos complexes covalents. Ainsi, nous avons vu que ceux-ci sont clairement sensibles aux hautes températures (Figure 73).

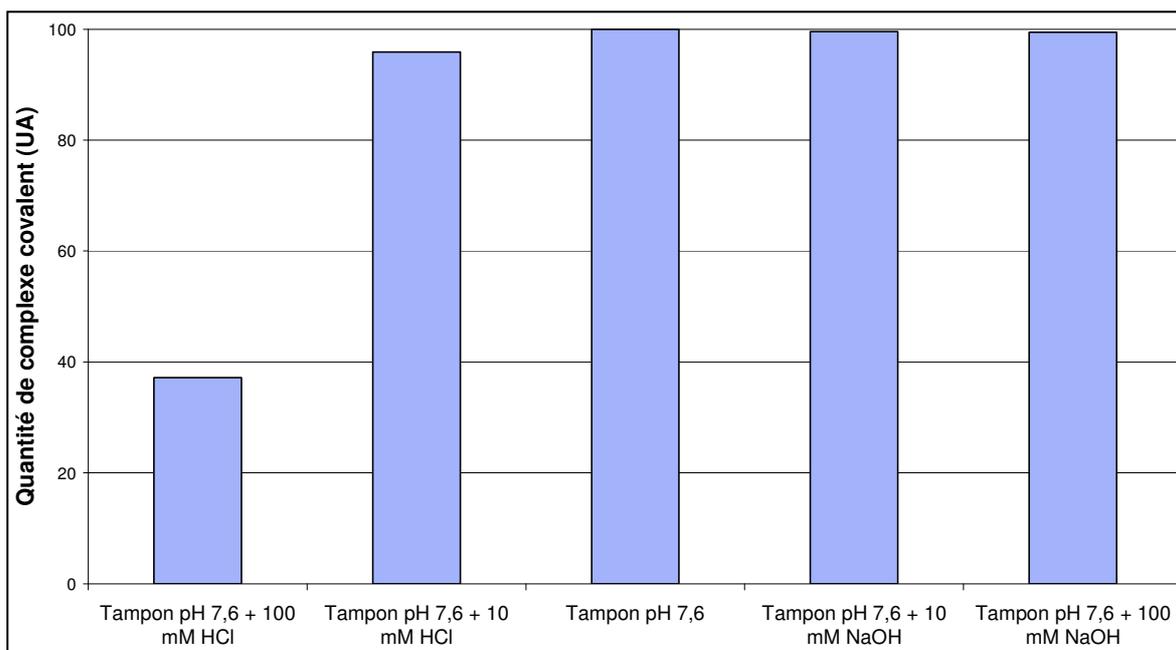
En effet, si on supprime l'étape de dénaturation thermique de notre protocole de SDS-PAGE-TRAP (voir Matériel et méthodes), on détecte environ 1,8 fois plus de complexe que lorsque l'on réalise cette étape, et cela que ce soit la nature mono- ou bicaténaire de la sonde oligonucléotidique contenant le *cHyd* (Figure 73).



**Figure 73 : Effet de la présence ou non de l'étape de dénaturation thermique du SDS-PAGE-TRAP sur la stabilité du complexe covalent formé entre *LIFpg* et le duplexe cHyd/C ou le cHyd simple-brin.** Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 37°C, et la concentration finale en Fpg est de 12,5 µM. L'analyse des résultats est faite par SDS-PAGE-TRAP, mais en effectuant ou non l'étape de dénaturation thermique des échantillons pendant 2 min à 90°C (voir Matériel et méthodes).

Afin de s'assurer que tous les complexes observés sont bien covalents, on a testé l'effet de l'absence de cette étape de dénaturation thermique sur le complexe non-covalent (incubé à 4°C pendant 15 minutes). Ainsi, la présence de SDS et la migration électrophorétique à température ambiante suffisent à dissocier les partenaires du complexe non-covalent (données non présentées). En conséquence, tous les complexes observés par SDS-PAGE-TRAP sont covalents, et ceci même sans réaliser l'étape de dénaturation thermique des échantillons. Cela nous permet donc de confirmer que notre complexe covalent est fortement instable à haute température, et il est probable qu'une incubation de plus de deux minutes à 90°C finirait par dissocier 100 % de nos complexes.

Outre la température, le pH influe aussi fortement sur la stabilité du complexe covalent, celui-ci étant clairement acido-labile (Figure 74). En effet, l'ajout de 100 mM final en HCl dans une solution (pourtant tamponnée avec environ 25 mM de tampons Tris et HEPES à pH 7,6 : voir Matériel et méthodes) de complexe covalent néoformé se traduit par une dissociation d'environ 63 % du complexe (Figure 74). Par contre, le complexe covalent est stable à pH physiologique et en présence de fortes concentrations de soude (Figure 74).



**Figure 74 : Effet du pH sur la stabilité du complexe covalent formé entre *LIFpg* et le duplexe cHyd/C.**

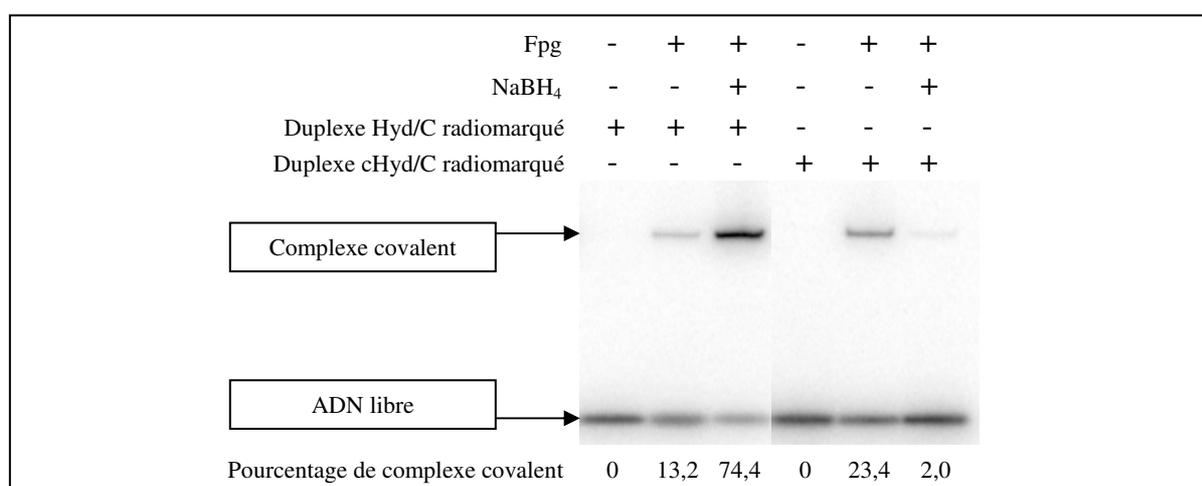
Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 37°C, et la concentration finale en Fpg est de 25 µM. Une quantité variable d'acide ou de base est alors ajoutée à chaque essai, puis on fait une incubation supplémentaire de 30 min à 50°C. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes). La quantité de complexe covalent est normalisée par rapport à celle que l'on a dans l'essai sans ajout d'acide ou de base, pour lequel on fixe une valeur de 100 unités arbitraires (UA).

Cette acido-labilité de notre complexe covalent est d'ailleurs fort intéressante, car elle explique pourquoi on n'a pas pu déterminer la masse de notre complexe par spectrométrie de masse (MALDI-TOF). En effet, la forte dilution et les conditions acides d'analyse par cette technique ont certainement eu pour conséquence de dissocier 100 % de notre complexe covalent. Par contre, cette analyse par spectrométrie de masse nous a permis d'observer qu'après dissociation du complexe covalent, l'oligonucléotide portant le cHyd était intact. Cette observation indique immédiatement que la réaction de pontage covalent doit correspondre à une simple condensation de l'enzyme sur l'ADN lésé cHyd sans qu'il y ait une réelle métabolisation de la lésion.

#### **II.2.2.2. Analyse comparative de la spécificité de la réaction de pontage et de l'activité Hyd-ADN glycosylase de *LIFpg***

La formation, dans des conditions physiologiques (125 mM NaCl, 37°C, pH 7,6), d'un complexe covalent entre *LIFpg* et le duplexe cHyd/C nous conduit à nous demander si cette réaction est possible avec la lésion Hyd naturelle (greffée sur un désoxyribose plutôt qu'un

cyclopentane). Pour examiner cette question, nous avons réalisé des incubations de *LIFpg* avec un duplexe cHyd/C ou un duplexe Hyd/C. La quantité de complexe covalent formé avec le cHyd est nettement plus importante qu'avec l'Hyd (environ 1,7 fois plus, pistes 2 et 5, Figure 75). D'autre part, au cours de la catalyse, on sait que Fpg réalise un intermédiaire transitoire covalent de type base de Schiff entre son amine N-terminale et le C1' d'un désoxyribonucléoside endommagé tel que l'Hyd. Cet intermédiaire peut d'ailleurs être piégé de façon irréversible sous la forme d'une base de Schiff réduite par réduction au borohydrure de sodium ( $\text{NaBH}_4$ ) ou au cyanoborohydrure de sodium ( $\text{NaCNBH}_3$ ) (Tchou *et al.*, 1995). Nous avons réalisé ce type d'essai sur nos duplexes Hyd/C et cHyd/C, et analysé les résultats par SDS-PAGE-TRAP (Figure 75).



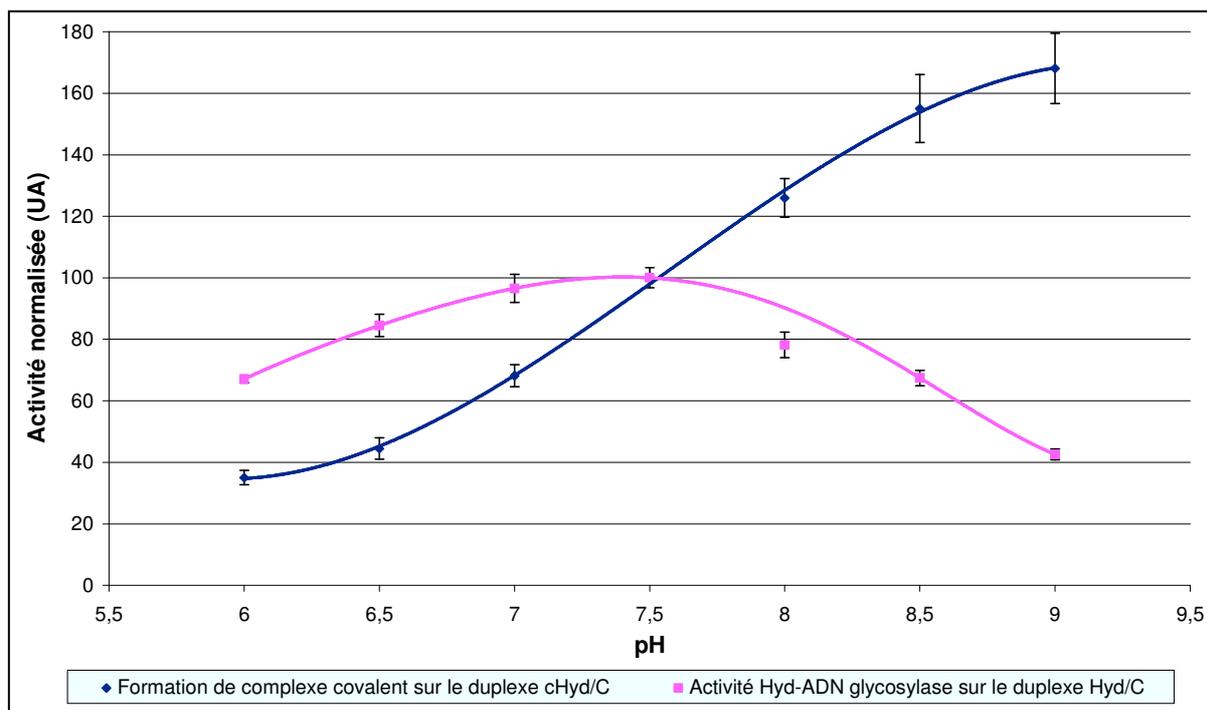
**Figure 75 : Le complexe covalent formé entre *LIFpg* et la 5-Hydroxy-5-Méthyle-Hydantoïne n'est pas un intermédiaire de type base de schiff.**

Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 37°C pour les essais sans  $\text{NaBH}_4$ , et de 10 min à 37°C pour les essais avec  $\text{NaBH}_4$ . La concentration finale en Fpg est de 25  $\mu\text{M}$ , et celle du  $\text{NaBH}_4$  est de 180 mM. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes).

Comme nous l'avons démontré dans le chapitre II.2.1, le duplexe Hyd/C est un bon substrat pour l'activité ADN glycosylase de *LIFpg* (Figures 51 et 52). Ainsi, il est possible de piéger de façon irréversible l'intermédiaire de type base de Schiff formé entre *LIFpg* et ce duplexe d'ADN par réduction au borohydrure de sodium (Piste 3, Figure 75). Cependant, il est intéressant de noter que l'on observe un complexe covalent entre *LIFpg* et le duplexe Hyd/C même en l'absence de  $\text{NaBH}_4$  (Piste 2, Figure 75). Sachant qu'aucun complexe covalent de type base de Schiff réduite n'est observable en l'absence de  $\text{NaBH}_4$  pour d'autres lésions que l'Hyd (ex : 8-oxo-G, Tchou *et al.*, 1995), on peut proposer que le complexe covalent borohydrure-indépendant observé avec l'Hyd est tout à fait similaire à celui observé avec l'analogue cHyd (Pistes 2 et 5, respectivement, Figure 75), et ne correspond pas à une

base de Schiff réduite. Ainsi, la réaction Hyd-ADN glycosylase et la réaction de pontage covalent borohydrure-indépendante sont toutes les deux possibles avec l'Hyd. Même si la formation de base de Schiff semble être plus favorable (Pistes 2 et 3, Figure 75), la réaction borohydrure indépendante et l'activité Hyd-ADN glycosylase sont vraisemblablement en compétition dans le cas du duplexe Hyd/C. Dans le cas du cHyd, l'ouverture du carbocycle portant la lésions étant impossible (contrairement aux lésions naturelles ayant un désoxyribose), il est de ce fait impossible de piéger un intermédiaire covalent de type base de Schiff entre *L/Fpg* et le duplexe cHyd/C. Cependant, une légère quantité de complexe covalent est observable avec le cHyd en présence de NaBH<sub>4</sub> (Piste 6, Figure 75). En fait, ce complexe covalent est vraisemblablement le même que celui observé sans borohydrure, la plus faible quantité de complexe observée sans NaBH<sub>4</sub> étant uniquement due à un temps d'incubation plus court (2 % et 23,4 %, respectivement, Figure 75). Ainsi, seule la réaction de pontage covalent borohydrure-indépendante étant possible avec le cHyd (et pas l'activité Hyd-ADN glycosylase) le duplexe cHyd/C peut être considéré comme un substrat suicide pour *L/Fpg* dans les conditions de pH, de température et de force ionique physiologiques.

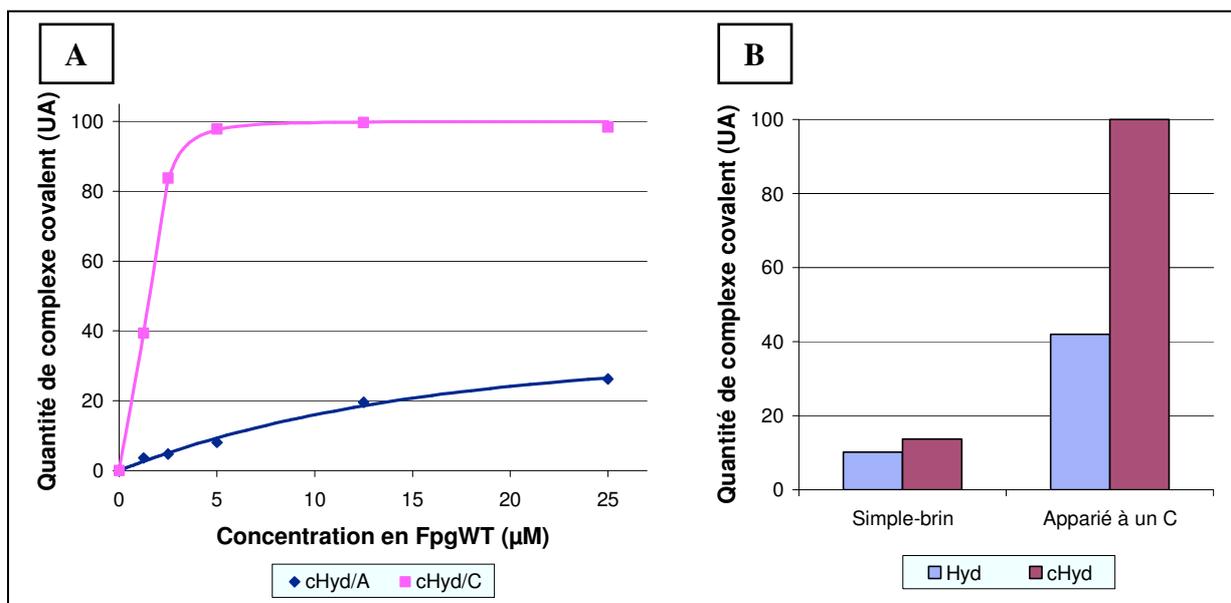
Nous avons vu que le pH influait fortement sur la stabilité du complexe covalent formé entre *L/Fpg* et le duplexe cHyd/C (Figure 74). Nous avons donc examiné l'effet du pH sur la formation de ce complexe covalent avec le duplexe cHyd/C, ainsi que sur l'activité Hyd-ADN glycosylase de l'enzyme sur le duplexe Hyd/C. Ces deux réactions se sont avérées être très sensibles au pH du milieu expérimental, mais possèdent une signature clairement différente lorsqu'on compare leur dépendance vis-à-vis du pH (Figure 76). En effet, dans la gamme de pH étudiée (pH 6 à pH 9), l'efficacité de la réaction suicide augmente de façon quasi-linéaire entre pH 6,5 et pH 8,5, cette linéarité s'amortissant de part et d'autre de ces bornes (courbe sigmoïde bleue, Figure 76). Entre pH 6 et 9, l'efficacité de la réaction suicide est multipliée par environ 4,8 fois, avec une valeur intermédiaire à pH physiologique. La dépendance vis-à-vis du pH de l'activité Hyd-ADN glycosylase est par contre très différente, et décrite par une courbe en cloche dont l'optimal d'activité est observé au voisinage de pH 7,5 (en rose, Figure 76). Ainsi, une diminution ou une augmentation de 1,5 unité pH par rapport à cet optimal amènent respectivement une perte de environ 33 % et 57 % de l'activité Hyd-ADN glycosylase de *L/Fpg*.



**Figure 76 :** Effet du pH sur l'activité Hyd-ADN glycosylase de *LIFpg* pour le duplexe Hyd/C, et sur la formation de complexe covalent entre *LIFpg* et le duplexe cHyd/C.

Les tampons d'incubation et de dilution d'enzyme utilisés sont les n°4 à différents pH (voir Annexe). Le temps d'incubation est de quatre jours à 37°C pour les tests de formation de complexe covalent, et de 30 min à 4°C pour les tests d'activité Hyd-ADN glycosylase. La concentration finale en Fpg est de 20 µM pour les tests de formation de complexe covalent, et de 5 nM pour les tests d'activité Hyd-ADN glycosylase. L'analyse des résultats est faite par SDS-PAGE-TRAP pour les tests de formation de complexe covalent, et sur gel dénaturant l'ADN pour les tests d'activité Hyd-ADN glycosylase (voir Matériel et méthodes). Les résultats ont été normalisés en fixant une valeur de 100 unités arbitraires (UA) pour le résultat obtenu à pH physiologique (7,5), que ce soit pour les tests de formation de complexe covalent ou pour les tests d'activité Hyd-ADN glycosylase.

Comme nous l'avons montré dans le chapitre II.2.1 pour l'activité Hyd-ADN glycosylase (Figures 51 et 52), la réaction de pontage covalent avec l'Hyd ou le cHyd est aussi fortement dépendante de la base opposée à la lésion, ainsi que de la nature mono- ou bicaténaire de la sonde oligonucléotidique (Figure 77). Pour la région de la courbe où la formation de complexe covalent est directement proportionnelle à la concentration d'enzyme (jusqu'à 2,5 µM en *LIFpg*, Figure 77A), la réaction suicide est environ 16 fois plus efficace avec le duplexe cHyd/C qu'avec le duplexe cHyd/A. D'autre part, la réaction de pontage covalent est aussi fortement stimulée par la nature bicaténaire de la sonde oligonucléotidique, que ce soit pour la lésion Hyd ou son analogue cHyd (environ 4,1 fois pour l'Hyd et 7,3 fois pour le cHyd, Figure 77B). De plus, on retrouve à nouveau que la réaction de pontage covalent est plus efficace avec le cHyd qu'avec l'Hyd (Figure 77B), ce qui est en accord avec notre conclusion précédente concernant la présence d'une compétition entre cette réaction et l'activité Hyd-ADN glycosylase pour l'Hyd, et pas pour le cHyd.



**Figure 77 : Effet du contexte d'appariement et de la nature désoxyribonucléique ou carbocyclique de la 5-Hydroxy-5-Méthyle-Hydantoïne sur l'efficacité de la réaction de pontage covalent avec *LIFpg*.**

Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 37°C. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes). Les résultats ont été normalisés en fixant une valeur de 100 unités arbitraires (UA) pour la quantité de pontage maximale, obtenue avec le duplexe cHyd/C. (A) : Quantité de pontage covalent obtenu avec les duplexes cHyd/C et cHyd/A en fonction de la concentration en Fpg. (B) : Quantité de complexe covalent obtenu entre Fpg et l'Hyd sous forme simple- ou double-brin, et sous forme naturelle ou carbocyclique. La concentration finale en Fpg est de 12,5 µM.

### II.2.2.3. La proline N-terminale de *LIFpg* est le nucléophile impliqué dans la réaction suicide avec le cHyd

#### II.2.2.3.1. Évidence structurale

Pour déterminer les atomes impliqués dans la liaison covalente entre *LIFpg* et le duplexe cHyd/C, nous avons résolu la structure cristallographique du complexe à partir des cristaux obtenus à 37°C (chapitre II.2.2.1 et « Matériel et méthodes »). L'obtention de ce type de cristal dépend essentiellement de la température de cristallisation. Les hautes concentrations en agent cristallisant et en complexe nucléoprotéique ont uniquement pour but de contrebalancer la plus grande solubilité des complexes à 37°C plutôt qu'à 20°C. Néanmoins, la qualité des cristaux obtenus pour des temps de cristallisation courts s'est avérée médiocre pour ce qui est de leur diffraction aux rayons X. Dans un grand nombre de cas, la carte de densité électronique obtenue était difficilement interprétable dans la zone d'intérêt, c'est-à-dire au niveau du cHyd. Cette difficulté d'interprétation de la carte de densité électronique pouvait être liée à un contenu hétérogène de ces cristaux, où pouvaient coexister

des complexes covalents et non-covalents (Figure 70A). Cette situation particulière est probablement rendue possible par le fait que le complexe non-covalent et le complexe covalent cristallisent dans le même groupe d'espace avec des paramètres de maille quasi-identiques (Tableaux 11 et 14, respectivement). Nous avons résolu ce problème expérimental en laissant croître ces cristaux plus de 15 jours à 37°C avant de les stocker dans l'azote liquide, ce temps de réaction long permettant de favoriser la formation de complexe covalent. Le cristal qui nous a permis de résoudre la structure du complexe covalent formé entre *LIFpg* et le duplexe *cHyd/C* a diffracté jusqu'à 1,9 Å de résolution. Un jeu complet de diffraction de ce cristal a pu être collecté sur la ligne de lumière ID 23-1 de l'ESRF (Grenoble) (Tableau 14).

Source de radiation	ESRF ID 23-1
Longueur d'onde (Å)	0,97935
Nombre total de réflexions	272948
Nombre de réflexions uniques	48939
Groupe d'espace	P 4 <sub>1</sub> 2 <sub>1</sub> 2
Paramètres du cristal	
Longueurs (Å)	a = b = 91,72 c = 143,24
Angles (°)	α = β = γ = 90
Données complètes (%) <sup>a</sup>	99,9 (99,9)
Redondance <sup>a</sup>	5,6 (5,6)
<i>I</i> / $\sigma$ ( <i>I</i> ) moyen <sup>a</sup>	13,8 (4,5)
<i>R</i> <sub>sym</sub> (%) <sup>a, b</sup>	9,3 (33,4)

<sup>a</sup> Les valeurs entre parenthèses correspondent aux données de résolution maximale.

<sup>b</sup>  $R_{sym} = 100 \times \sum |I_i - \langle I \rangle| / \sum I_i$ , où  $I_i$  est l'intensité observée pour une réflexion  $i$ , et  $\langle I \rangle$  est l'intensité moyenne pour plusieurs réflexions liées par symétrie.  $R_{sym}$  est calculé sur des intensités obtenues avec un même cristal.

**Tableau 14 : Statistiques de collection des données de diffraction.** (complexe covalent).

Le problème des phases a été résolu par la méthode de remplacement moléculaire en utilisant la structure de *LIFpg* de type sauvage complexée à un ADN contenant un site *cFapyG* (code PDB 1XC8).

La carte de densité électronique obtenue après remplacement moléculaire s'est avérée être d'excellente qualité. La chaîne principale de la protéine *LIFpg* a nécessité peu de repositionnements manuels, par contre, certaines chaînes latérales ont du être ajustées manuellement grâce au programme TURBO-FRODO.

Le programme REFMAC a ensuite été utilisé pour l'affinement de la position  $x$ ,  $y$ ,  $z$  de chacun des atomes ainsi que du facteur  $B$  d'agitation thermique. On peut noter que, contrairement aux autres atomes, l'affinement du facteur  $B$  du zinc a été réalisé de façon anisotrope car les cartes de densité électronique résiduelles  $mF_o - DF_c$  ont montré une forte anisotropie autour de cet atome. Les valeurs des facteurs  $R$  obtenues à la fin de l'affinement du modèle moléculaire (respectivement 16,2 % et 19,7 % pour  $R_{work}$  et  $R_{free}$ ) indiquent que le

modèle atomique obtenu est en bonne corrélation avec les données expérimentales, et la faible différence entre ces valeurs (de 3,5 %) confirme la validité de notre procédure d'affinement (Tableau 15).

Le modèle atomique final obtenu possède 3156 atomes, dont 2123 atomes correspondant à la protéine, 565 à l'ADN, 462 aux oxygènes de l'eau, et 6 à une molécule de glycérol. La structure de certaines chaînes latérales exposées au solvant n'a pas pu être reconstruite entièrement à cause de la mobilité des atomes correspondants (résidus 15, 21, 31, 42, 65, 88, 130, 213 et 230). De plus, le modèle obtenu comprend chacun des 28 nucléotides de l'ADN, ceci contrairement à la structure de la protéine, où les résidus 219 à 223 sont manquants. En effet, ces résidus sont localisés dans la boucle  $\alpha$ F- $\beta$ 9 de la protéine, qui est connue pour être mobile, d'où l'impossibilité de reconstruire ces résidus.

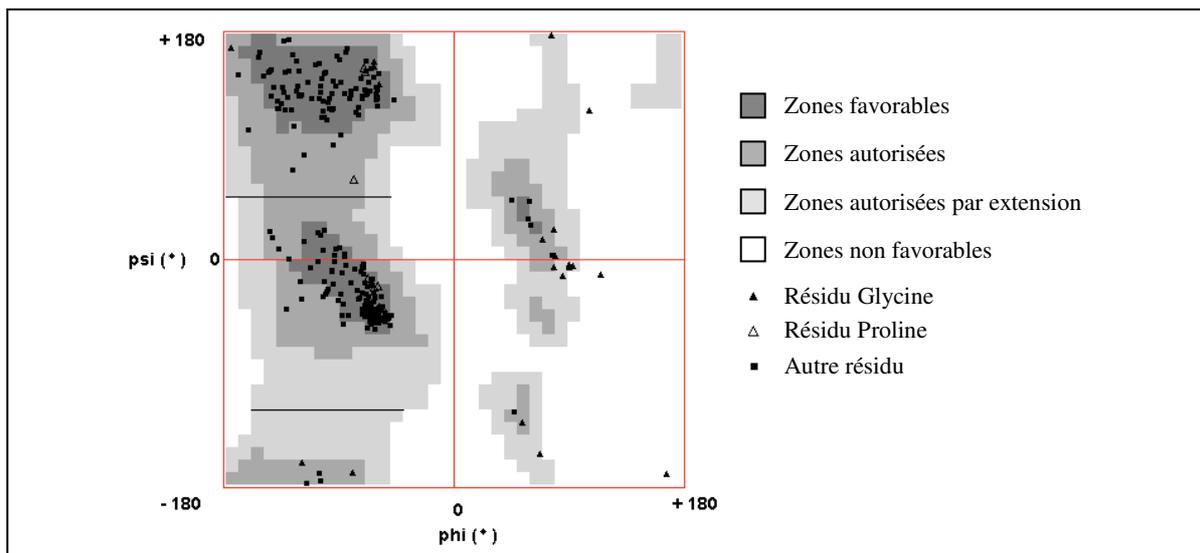
Résolution (Å)	45.88-1.90
Nombre de réflexions utilisées	46387
R <sub>work</sub> (%) <sup>a</sup>	16.2
R <sub>free</sub> (%) <sup>a</sup>	19.7
Etude du diagramme de Ramachandran <sup>b</sup>	
Résidus dans les régions favorables (%)	87,55
Résidus dans les régions additionnellement autorisées (%)	12,45
Résidus dans les régions autorisées par extension (%)	0
Résidus dans les régions interdites (%)	0
Valeurs moyennes du facteur B :	
Sur tous les atomes (Å <sup>2</sup> )	20,72
Sur les atomes de la protéine (Å <sup>2</sup> )	16,89
Sur les atomes de l'ADN (Å <sup>2</sup> )	26,67
Sur les atomes de l'eau (Å <sup>2</sup> )	34,02
Ecart type par rapport à l'idéal	
Sur la longueur des liaisons (Å)	0,013
Sur les angles de liaisons (°)	1,448
Sur les angles de torsion (°)	5,517
Nombre d'atomes (atomes d'hydrogènes non comptés) :	
Protéine	2123
ADN	565
Eau	462
Glycérol	6
Total	3156

<sup>a</sup>  $R = 100 \times \frac{\sum |F_{obs} - k| F_{calc}}{\sum |F_{obs}|}$ , avec R<sub>work</sub> calculé sur les réflexions utilisées pour l'affinement, et R<sub>free</sub> calculé sur les 5 % de réflexions non-utilisées. k est un facteur d'échelle entre F<sub>obs</sub> et F<sub>calc</sub>.

<sup>b</sup> Les données présentées pour le diagramme de Ramachandran ne tiennent pas compte des prolines et glycines de l'enzyme.

**Tableau 15 : Statistiques sur l'affinement et sur le modèle moléculaire.** (complexe covalent).

Le diagramme de Ramachandran (Figure 78) obtenu pour la protéine dans le complexe témoigne d'une structure protéique de bonne qualité étant donné que l'on trouve environ 87,5 % des résidus dans les zones favorables de ce diagramme, et 100 % dans les zones favorables ou autorisées (les prolines et glycines étant exclues).



**Figure 78 :** Diagramme de Ramachandran sur la structure de *LIFpg* dans le complexe covalent formé avec le duplexe cHyd/C.

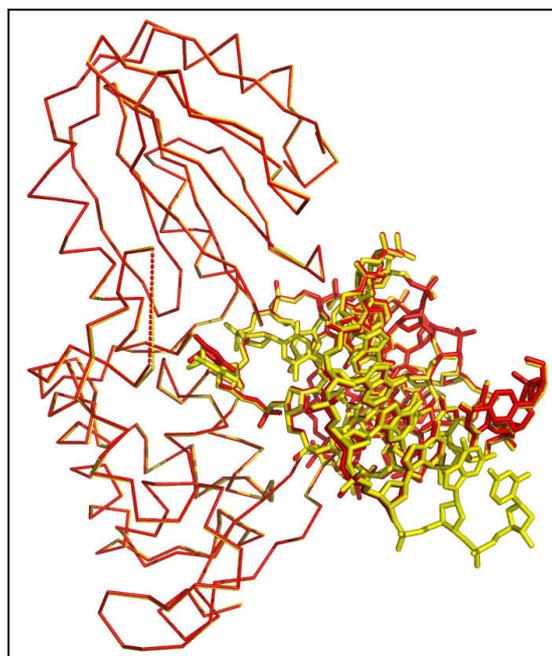
Diagramme déterminé grâce au programme RP (2.0), Gopalakrishnan *et al.*, 2007.

De plus, le modèle structural ne montre aucun résidu avec des angles ou des distances significativement différentes de l'idéal ce qui confirme que la structure est de bonne qualité.

La structure globale du complexe covalent formé entre *LIFpg* et le duplexe cHyd/C est très similaire à celle du complexe non-covalent (Figure 79).

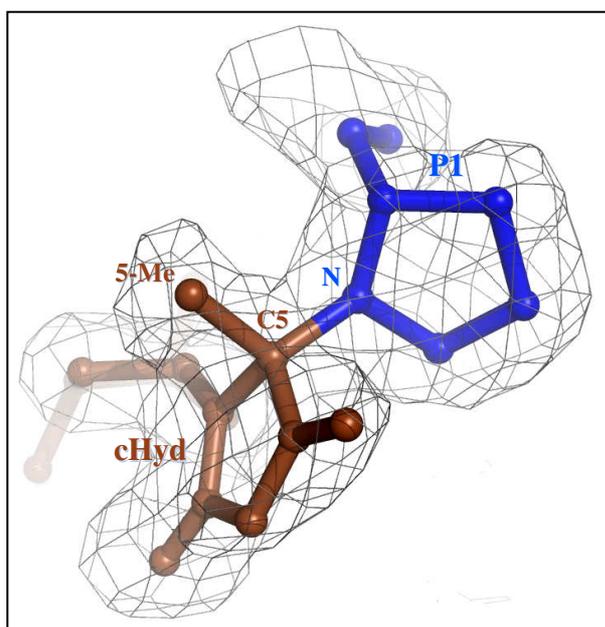
**Figure 79 :** Superposition des structures globales de *LIFpg* complexée covalamment ou non à un cHyd.

La structure avec le complexe covalent est en rouge, celle avec le complexe non-covalent est en jaune. La superposition des structures a été réalisée sur les 263 Ca communs entre les deux structures (les résidus 219 à 223 sont absents des structures avec le cHyd, et les résidus 66, 73 et 121 ont deux conformations dans l'une des structures), et sur les 524 atomes communs de l'ADN. Le RMSD correspondant à cette superposition est de 0,322 Å, et passe à 0,107 Å si on ne prend pas en compte les atomes de l'ADN, à 0,351 Å si on ne prend pas en compte les atomes de la protéine, et à 0,303 Å si on ne prend en compte ni les atomes de la protéine ni les atomes du cHyd.



Les différences observables au niveau des extrémités du brin opposé à la lésion proviennent du fait que nous avons utilisé un ADN à bord cohésif avec une base 5'-sortante

pour le complexe non-covalent, et 3'-sortante pour le complexe covalent. Ces différences n'ont aucun impact sur la structure du reste des complexes, car elles ne sont pas localisées dans des zones d'interaction entre la protéine et l'ADN. Dans le cas du complexe covalent, le cHyd a été extrait hors de la double hélice d'ADN et stabilisé en position extrahélicale dans la poche du site actif de l'enzyme comme dans le cas du complexe non-covalent (Figure 79). La différence majeure entre les structures des complexes covalent et non-covalent réside dans l'observation d'une continuité de densité électronique entre le cHyd et le résidu catalytique P1 de *L/Fpg* (Figure 80).

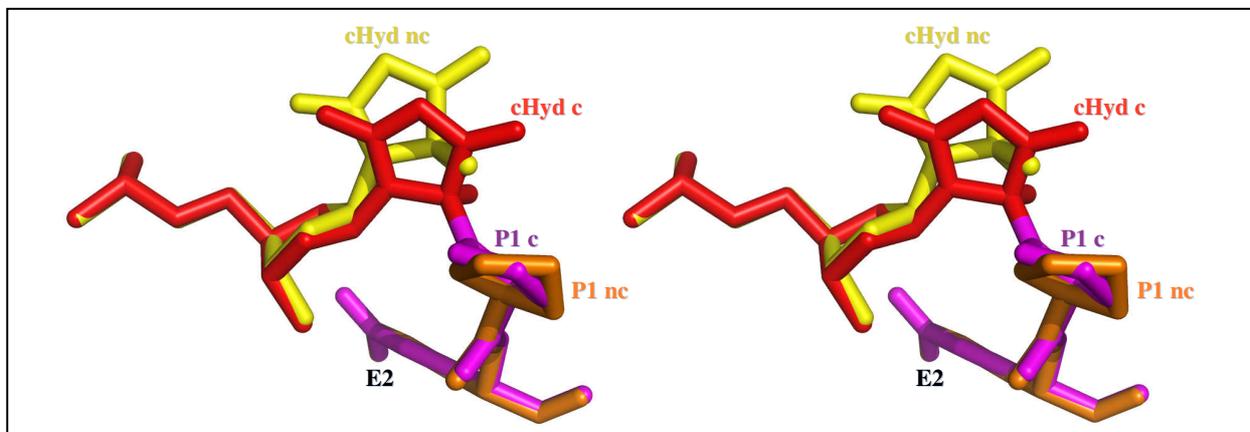


**Figure 80 : Représentation du modèle structural covalent obtenu entre *L/Fpg* et le duplexe cHyd/C, montrant la lésion 5-OH-5-Me-Hydantoïne et le résidu P1 de Fpg.**

Les atomes de notre modèle sont en représentation bâtons, et sont respectivement colorés en marron et en bleu pour la lésion et le résidu P1 de Fpg. La carte de densité électronique  $2F_o - F_c$  contourée à  $1\sigma$  de cette zone est représentée par un maillage gris.

Cette continuité de densité électronique témoigne indiscutablement de la présence d'une liaison covalente entre le carbone C5 du cHyd et l'azote N-terminal de la protéine. La longueur de cette liaison covalente (1,51 Å) est en très bon accord avec la longueur théorique d'une liaison carbone-azote (1,52 Å, calculé à partir des rayons de covalence des deux atomes). Comme le laissait déjà entrevoir l'analyse du complexe covalent par spectrométrie de masse, ce pontage covalent entre le duplexe cHyd/C correspond à une réaction de condensation (perte d'une molécule d'eau) associée au départ du groupement hydroxyle en C5 du cHyd. En effet, le groupement restant attaché en C5 du cHyd dans le complexe covalent est le groupement méthyle, car la longueur de la liaison covalente entre le C5 et l'atome central de ce groupement est de 1,55 Å, ce qui est en bien meilleur accord avec la longueur théorique d'une liaison carbone-carbone qu'avec celle d'une liaison carbone-oxygène (1,54 Å et 1,50 Å, respectivement, calculés à partir des rayons de covalence des atomes). La formation du complexe covalent résulte du rapprochement du résidu P1 de *L/Fpg*

et de la partie base nucléique lésée du cHyd (Figure 81). En effet, lors du passage du complexe non-covalent au complexe covalent, la distance entre le C5 du cHyd et l'azote de P1 passe de 3,68 Å à 1,52 Å.



**Figure 81 :** Vue stéréoscopique des différences structurales entre les complexes covalent ou non au voisinage de la lésion cHyd.

La superposition est faite sur les mêmes atomes que pour la Figure 79. Les analogues de lésion cHyd dans les complexes covalent ou non sont respectivement notés **cHyd c** et **cHyd nc**, et sont représentés en rouge et en jaune, respectivement. Les résidus P1 de Fpg dans les complexes covalent ou non sont notés **P1 c** et **P1 nc**, et sont représentés en magenta et en orange, respectivement. Le résidu E2 est représenté avec le même code couleur que pour le résidu P1.

Ce rapprochement est associé à de légères modifications structurales au niveau de la conformation du cyclopentane et de la partie base nucléique du cHyd (Tableaux 16), ainsi qu'au niveau du positionnement du cycle du résidu P1 de Fpg (Tableaux 17 et Figure 82). Ces légers réajustements structuraux ne sont cependant pas associés à des variations structurales ailleurs dans l'ADN ou dans la protéine. Entre le complexe non-covalent et le complexe covalent, le plissement du cyclopentane passe de C3'-endo / C2'-exo à C2'-exo (Tableau 16).

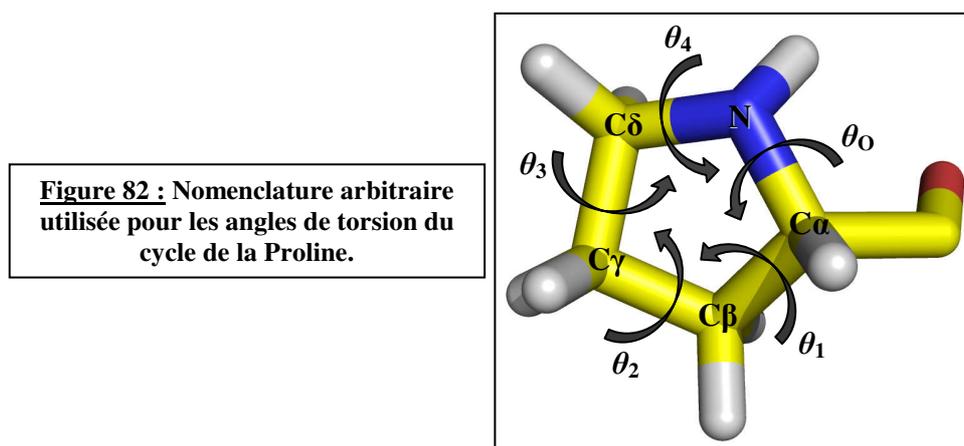
Angle de torsion	$\nu_0$ (°)	$\nu_1$ (°)	$\nu_2$ (°)	$\nu_3$ (°)	$\nu_4$ (°)	$\chi$ (°)	Plissement du cyclopentane (angle de phase de pseudorotation en degrés)	Conformation de la liaison N-glycosidique
Complexe non-covalent	-9,54	30,54	-41,17	34,81	-15,31	63,04	C <sub>3</sub> '-endo / C <sub>2</sub> '-exo (4,53)	Syn
Complexe covalent	-26,57	47,21	-49,18	34,50	-4,82	59,18	C <sub>2</sub> '-exo (337,52)	Syn

**Tableau 16 :** Angles de torsion du sucre et de la liaison N-glycosidique de la lésion cHyd dans les complexes covalent ou non.

La détermination de ces angles a été faite grâce au programme Swiss-PdbViewer version 4.0 (Guex *et al.*, 1997). La nomenclature des angles de torsion est indiquée dans la Figure 100 du « Matériel et méthodes ».

Cette différence subtile, alliée à une légère rotation autour de la liaison *N*-glycosidique (Tableau 16), amène une nette variation de position de l'atome C5 du cHyd, celui-ci étant plus proche du résidu P1 dans la structure du complexe covalent (Figure 81).

D'autre part, même si le positionnement du résidu P1 de *L/Fpg* semble peu différent dans les complexes covalent ou non (Figure 81), le plissement du cycle de P1 est modifié entre les deux structures (Tableau 17 et Figure 82). En effet, si on utilise le formalisme communément utilisé pour définir le plissement des désoxyriboses de l'ADN (voir Matériel et méthodes), on obtient des angles de phase de pseudorotation de 62,68° et 6,65° pour le résidu P1 dans les complexes covalent et non-covalent, ce qui nous donne respectivement des plissements de type C $\delta$ -exo et C $\beta$ -exo/C $\gamma$ -endo. Ainsi, le plissement du résidu P1 est très différent pour les deux types de complexes, ce qui permet un net rapprochement entre l'azote de P1 et le C5 du cHyd dans le complexe covalent par rapport au complexe non-covalent (Figure 81).



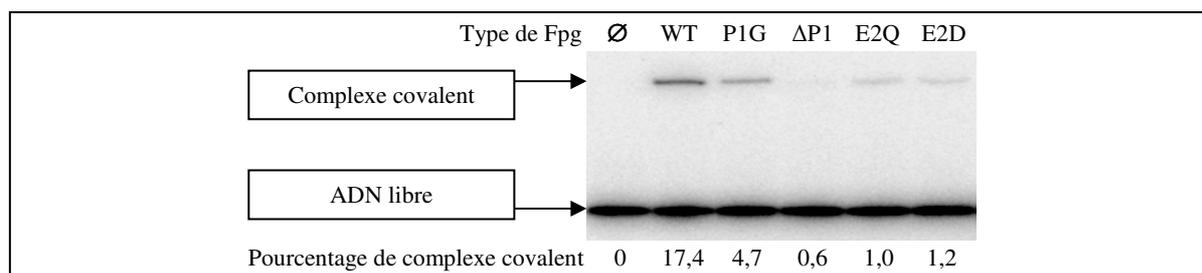
Angle de torsion	$\theta_0$ (°)	$\theta_1$ (°)	$\theta_2$ (°)	$\theta_3$ (°)	$\theta_4$ (°)
Complexe non-covalent	8,27	-30,78	42,16	-36,57	17,60
Complexe covalent	-21,14	4,22	13,57	-25,63	29,87

**Tableau 17 : Angles de torsion du cycle de la Proline *N*-terminale de Fpg dans les complexes covalent ou non.**

La détermination de ces angles à été faite grâce au programme Swiss-PdbViewer version 4.0 (Guex *et al.*, 1997). La nomenclature des angles de torsion est indiquée dans la Figure 82.

### II.2.2.3.2. Bases fonctionnelles et sélectivité de la réaction suicide entre *L/Fpg* et le cHyd

L'analyse structure/activité de la réaction suicide entre *L/Fpg* et le duplexe cHyd/C suggère fortement que la formation du complexe covalent résulte d'une attaque nucléophile de l'azote de P1 en C5 du cHyd. Pour parfaire notre analyse, nous avons testé la capacité des formes de type mutant de l'enzyme à catalyser la réaction suicide avec le duplexe cHyd/C (Figure 83).



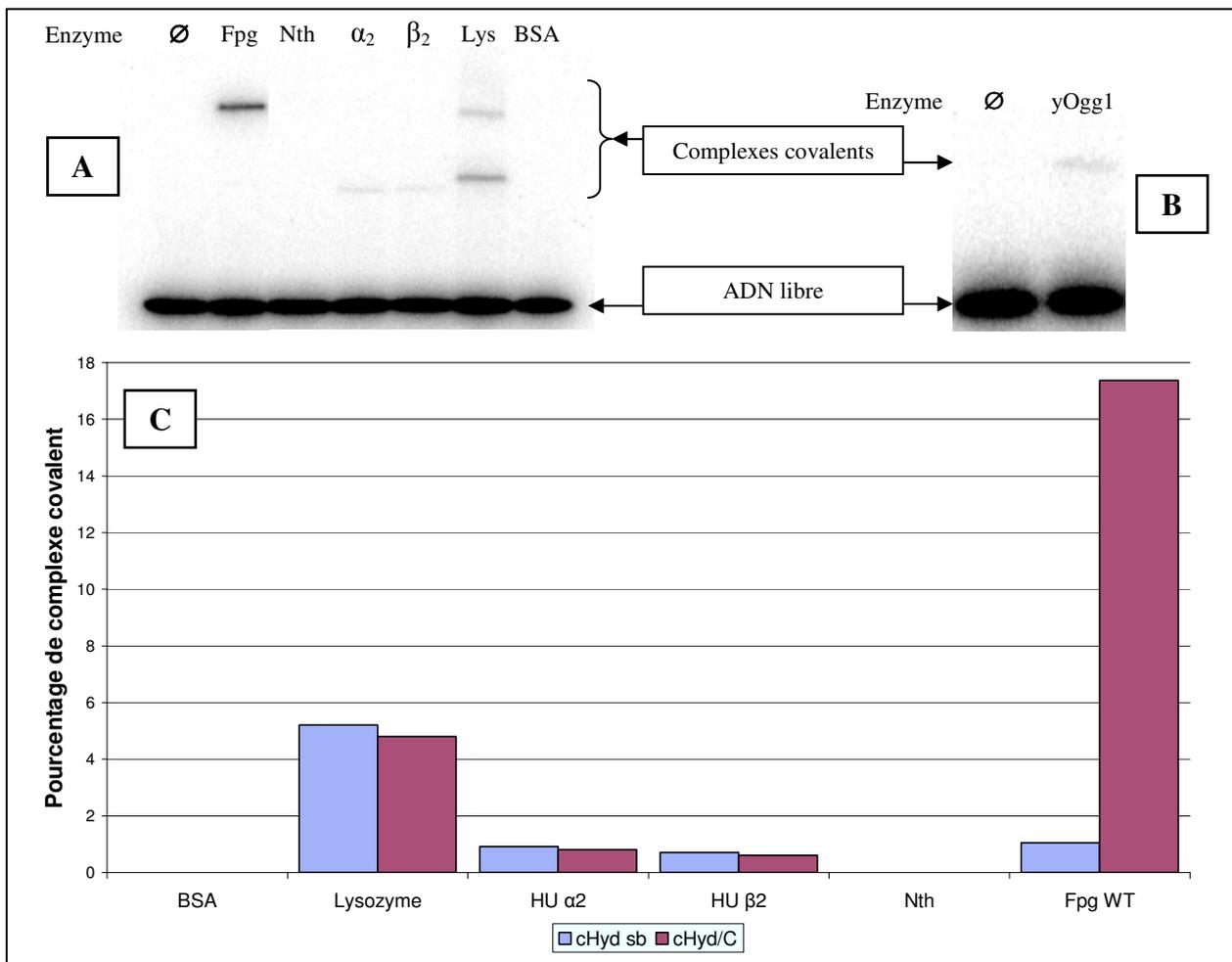
**Figure 83 : Analyse fonctionnelle par mutagenèse de l'efficacité de la réaction suicide entre *L/Fpg* et le duplexe cHyd/C.**

Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans son tampon de stockage (voir Annexe). Le temps d'incubation est de quatre jours à 37°C, et la concentration finale en Fpg est de 12,5 µM. L'analyse des résultats est faite par SDS-PAGE-TRAP (voir Matériel et méthodes). La notation Ø correspond à un essai sans Fpg.

De toute évidence, toute mutation affectant le résidu P1 de l'enzyme induit une forte diminution de l'efficacité de la réaction de pontage covalent sur le duplexe cHyd/C par rapport à l'enzyme de type sauvage (Pistes 2 à 4, Figure 83). La substitution P1G induit cependant une plus faible diminution de la quantité de pontage covalent que la délétion ΔP1, cela probablement car le résidu glycine est tout de même légèrement nucléophile (même si moins que la proline). D'autre part, les substitutions du résidu E2 de l'enzyme par un Q ou un D amènent aussi une forte diminution de la quantité de complexe covalent (Pistes 2, 5 et 6, Figure 83). Tout comme pour l'activité ADN glycosylase de *L/Fpg* (chapitre II.1.2.3), la réaction suicide implique clairement les résidus catalytiques P1 et E2, qui sont situés dans la poche du site actif de l'enzyme. Ces résultats sont en accord avec notre analyse structurale, qui indique que l'amine N-terminale de Fpg est responsable de l'attaque en C5 du cHyd, aboutissant à la création d'un pontage covalent. Toujours par analogie avec l'activité ADN-glycosylase de *L/Fpg*, le résidu E2 pourrait être responsable de l'activation de la réaction suicide par déprotonation du résidu P1 *via* une molécule d'eau (chapitre II.1.2.3 et Coste *et al.*, 2008), expliquant ainsi pourquoi les mutants E2D et E2Q forment moins de complexe

covalent avec le cHyd que l'enzyme de type sauvage. En effet, comme pour l'activité ADN glycosylase, la formation d'un complexe covalent entre l'enzyme et le cHyd nécessite que l'amine N-terminale de Fpg soit nucléophile, et donc déprotonée.

Afin d'évaluer la sélectivité de la réaction suicide sur le cHyd, nous avons examiné la possibilité de piéger différentes protéines sous formes de complexes covalents avec le duplexe cHyd/C (Figure 84).



**Figure 84 : Sélectivité de la réaction suicide entre *L/Fpg* et le cHyd.**

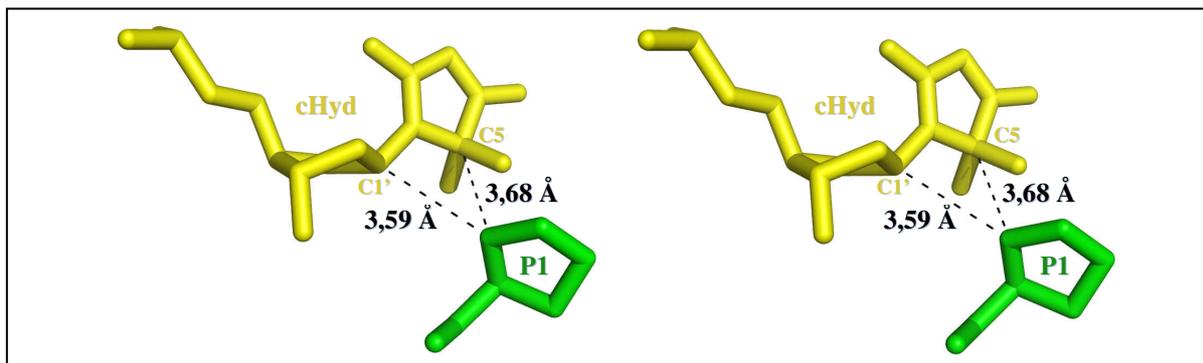
(A) : Tests SDS-PAGE-TRAP (voir Matériel et méthodes) entre le duplexe cHyd/C et différentes protéines. Le tampon d'incubation utilisé est le n°5, et les enzymes sont diluées dans le tampon de stockage de Fpg (voir Annexe). Le temps d'incubation est de quatre jours à 37°C, et la concentration finale est de 12,5  $\mu$ M pour chaque protéine. La notation  $\emptyset$  correspond à un essai sans enzyme, la notation  $\alpha_2$  correspond à la protéine HU homodimérique  $\alpha_2$ , la notation  $\beta_2$  correspond à la protéine HU homodimérique  $\beta_2$ , et la notation Lys correspond au lysozyme. (B) : Test SDS-PAGE-TRAP (voir Matériel et méthodes) entre le duplexe cHyd/C et la protéine yOgg1. Le tampon d'incubation utilisé est le n°5, et l'enzyme est diluée dans le tampon de stockage de Fpg (voir Annexe). Le temps d'incubation est de quatre jours à 37°C, et la concentration finale en yOgg1 est de 19,3  $\mu$ M. La notation  $\emptyset$  correspond à un essai sans enzyme. (C) : Quantités de complexe covalent obtenues par SDS-PAGE-TRAP entre le duplexe cHyd/C ou le cHyd simple-brin (sb) et différentes protéines. Les conditions expérimentales sont celles du (A).

L'albumine de sérum bovin (BSA) constitue un contrôle négatif pour cette expérience (protéine acide ne se fixant pas à l'ADN de pI = 5,5), et ne réalise pas de complexe covalent avec le cHyd (Figure 84C et Piste 7, Figure 84A). Le lysozyme de blanc d'oeuf, une glycosidase (protéine fortement basique ne se fixant pas à l'ADN, de pI = 9,5), semble être capable de former différents complexes covalents avec le cHyd, mais ceci de façon peu abondante et sans spécificité quant à la nature mono- ou bi-caténaire de l'ADN (Figure 84C et Piste 6, Figure 84A). Cela suggère qu'un certain nombre de résidus protéiques sont susceptibles de faire office de nucléophile pour l'attaque en C5 du cHyd. Ensuite, le caractère fortement basique (pI = 10,5) et la haute affinité pour l'ADN de la protéine de type histone HU de *E. coli* (sous ses formes homodimériques  $\alpha_2$  et  $\beta_2$ ) ne sont pas suffisants pour permettre la formation significative de complexe covalent avec le cHyd (Figure 84C et Piste 4 et 5, Figure 84A). De plus, tout comme le lysozyme, la protéine HU ne présente aucune préférence quant à la nature double-brin de l'ADN pour la formation de complexe covalent avec le cHyd (Figure 84C). Le résultat le plus intéressant de cette expérience réside dans l'observation que Nth ne peut pas être piégée covalamment avec le cHyd (Figure 84C et Piste 3, Figure 84A). Pour cette enzyme, une ADN glycosylase spécifique de l'excision des lésions pyrimidiques, le cHyd n'est pas un inhibiteur suicide, et cela alors que l'Hyd est efficacement excisée par Nth (Figure 52). On obtient donc du pontage covalent sur le cHyd avec les protéines HU, le lysozyme et Fpg, et pas avec Nth ou la BSA (Figure 84A et 84C). De plus, la protéine Fpg donne nettement plus de complexe covalent avec le duplexe cHyd/C que toutes les autres protéines (Figure 84C) sur le duplexe cHyd/C ou le cHyd simple-brin. En outre, Fpg est la seule à montrer une préférence pour l'ADN double brin parmi les protéines formant des complexes covalents avec le cHyd, les autres protéines montrant plutôt une très légère préférence pour l'ADN simple-brin (Figure 84C). On peut préciser ici que la taille des différents complexes semble correcte, étant donné que Fpg est la plus grosse des protéines donnant du pontage (MM = 31 kDa), suivie du Lysozyme (MM = 14,7 kDa), puis de HU sous sa forme monomérique (MM = 9,4 kDa). Par contre, la présence d'une seconde bande de taille élevée avec le lysozyme reste pour l'instant inexplicée. D'autre part, des expériences préliminaires semblent indiquer que l'homologue fonctionnel eucaryote yOgg1 de Fpg (issu de *S. cerevisiae*) peut être piégé de façon covalente par le cHyd (Figure 84B). Finalement, l'ensemble de ces données suggèrent que la réaction de pontage covalent entre le cHyd et Fpg est relativement sélective, et que le duplexe cHyd/C serait un piège moléculaire efficace uniquement pour les ADN glycosylases spécifiques de la réparation des lésions appariées à une pyrimidine.

#### II.2.2.4. Conclusion et discussion

Les données expérimentales présentées dans ce dernier chapitre de résultats nous ont permis d'une part de mettre en évidence la formation d'un complexe covalent et abortif entre l'analogue carbocyclique de la 5-hydroxy-5-méthyle-Hydantoïne et la protéine *LIFpg* et, d'autre part, d'en établir les bases structurales et fonctionnelles. On peut d'ailleurs indiquer ici que la structure du complexe covalent que nous avons résolue est la première qui soit disponible pour une ADN glycosylase complexée covalamment sur un ADN sans qu'un agent pontant n'ait été utilisé. D'autre part, le complexe covalent mis en évidence entre une ADN glycosylase et une lésion n'est que le troisième exemple de ce type dans la bibliographie (la 2-désoxyribonolactone : Hashimoto *et al.*, 2001; Kroeger *et al.*, 2003; et l'oxanine : Nakano *et al.*, 2003), et seulement le deuxième si on considère les lésions de bases nucléiques (l'oxanine : Nakano *et al.*, 2003). Du fait de son insensibilité au borohydrure de sodium, nous avons démontré que ce complexe est clairement différent de l'intermédiaire de type base de Schiff qui se forme de façon transitoire au cours de l'activité ADN glycosylase (Figure 75). Après avoir caractérisé les déterminants structuraux et fonctionnels de la reconnaissance et de l'excision du dommage Hyd par *LIFpg* (chapitre II.2.1), nous avons montré que le complexe covalent formé entre l'enzyme et un duplexe cHyd/C résulte d'une attaque nucléophile en C5 du dommage par l'azote de la proline *N*-terminale de l'enzyme (P1). Il s'agit d'une condensation simple de l'enzyme sur l'ADN modifié sans une réelle métabolisation du dommage. D'une façon non-ambigüe, nous avons montré que cette réaction de pontage covalent (attaque de l'azote de P1 en C5 de Hyd) et la réaction Hyd-ADN glycosylase (attaque de l'azote de P1 en C1' de Hyd) peuvent se réaliser toutes les deux sur le désoxyribonucléoside naturel Hyd (Figure 75), d'où une compétition entre ces deux réactions. La découverte de l'attaque en C5 du dommage lors de la réaction de pontage nous permet de comprendre pourquoi ce pontage covalent est observable avec le désoxyribonucléoside Hyd et son analogue carbocyclique cHyd, car ces deux molécules n'ont aucune différence au niveau de leur partie base nucléique lésée. De ce point de vue, le cHyd, qui ne peut en aucun cas être substrat pour l'activité ADN glycosylase de l'enzyme, correspond parfaitement à la définition d'un substrat suicide dans des conditions physiologiques. La possibilité pour l'enzyme d'attaquer le dommage en C1' et en C5 est parfaitement supportée par la structure cristallographique du complexe non-covalent entre *LIFpg* et le duplexe cHyd/C (Figure 85). En effet, pour les deux mécanismes catalytiques, l'amine *N*-terminale de P1 se retrouve à des distances très similaires du C1' et du C5 (0,09 Å

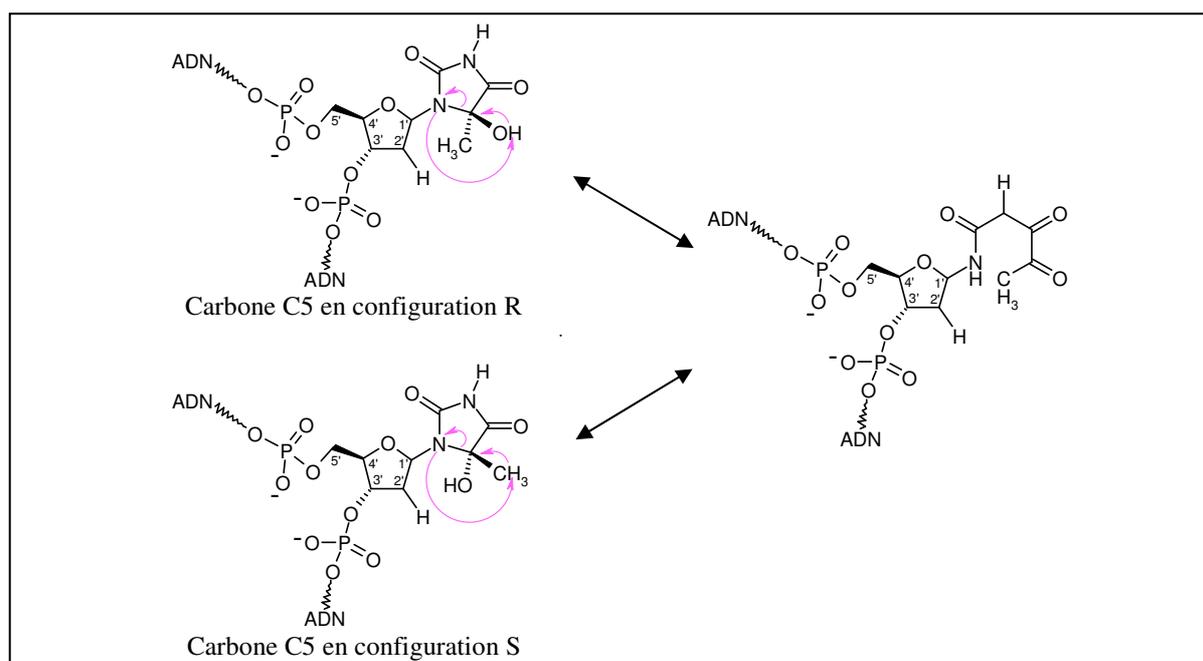
de différence, Figure 85), et toujours idéalement positionnée pour effectuer une attaque nucléophile sur ces deux carbones du dommage.



**Figure 85 : Vue stéréoscopique de la localisation du résidu P1 de Fpg par rapport au cHyd dans le complexe non-covalent.**

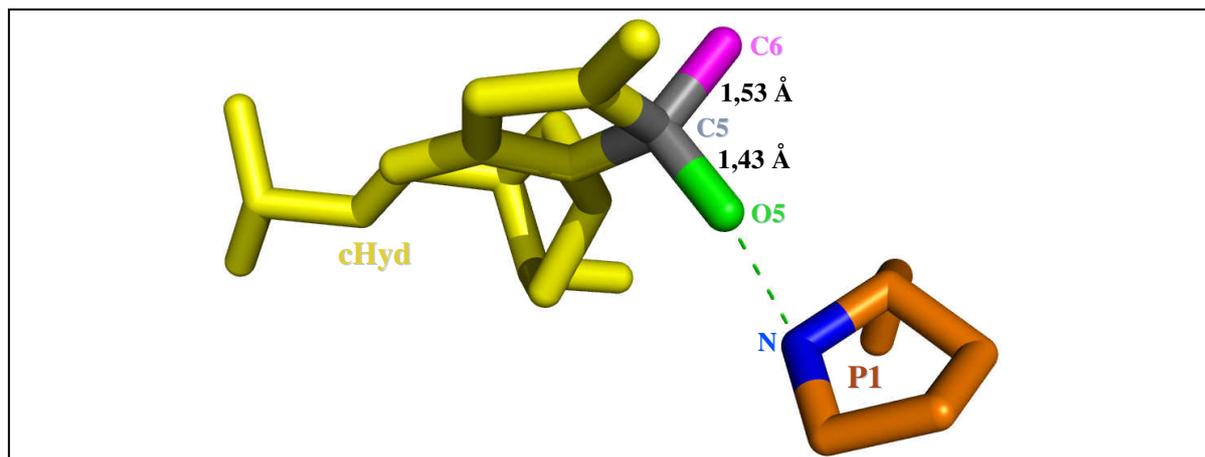
Le résidu P1 est représenté en vert, le cHyd est représenté en jaune, les distances entre l'azote de P1 et le C1' ou le C5' du cHyd sont représentées en noir, et sont localisées par des pointillés noirs.

La résolution des structures cristallographiques des complexes non-covalent et covalent formés entre *LlFpg* et le duplexe cHyd/C indique que la réaction suicide et la reconnaissance du cHyd par l'enzyme, et aussi très probablement l'activité Hyd-ADN glycosylase, sont des mécanismes moléculaires stéréospécifiques. En effet, lors de la synthèse de l'oligonucléotide contenant un cHyd, nos collègues chimistes n'ont pas réussi à séparer les molécules présentant les configurations R et S du carbone C5 à cause d'un phénomène d'épimérisation. En effet, au dessus de pH 7, les épimères sont en interconversion, donnant un équilibre avec 50% de chaque espèce (racémique) (Figure 86).



**Figure 86 : Représentation schématique du mécanisme d'épimérisation de la 5-OH-5-Me-Hydantoïne.** Cette réaction n'est possible qu'à pH > 7, ce qui est le cas dans nos essais au pH physiologique de 7,6.

Dans le modèle structural du complexe non-covalent, la grande qualité de la carte de densité électronique au niveau du dommage indique clairement que c'est l'épimère de configuration R qui est sélectionné par l'enzyme pour être stabilisé dans la poche du site actif (distance C5-CH<sub>3</sub> de 1,53 Å, contre 1,54 Å en théorie, et distance C5-OH de 1,43 Å, contre 1,50 Å en théorie, Figure 87).



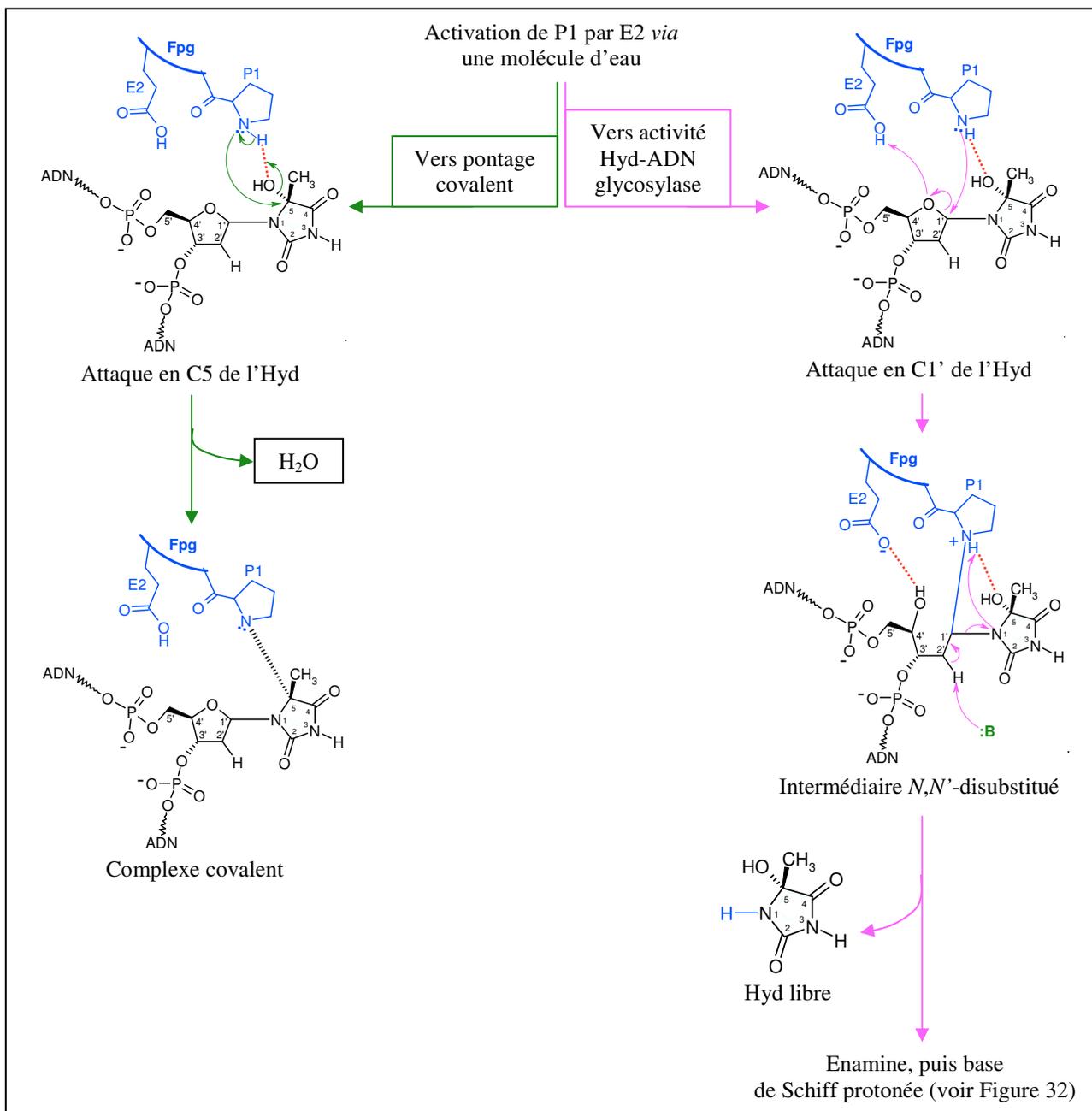
**Figure 87 : Mode de reconnaissance stéréospécifique du cHyd par L/Fpg.**

Le cHyd est en jaune, le résidu P1 de Fpg est en marron, l'azote de P1 est en bleu, le O5 du cHyd est en vert, le C5 du cHyd est en gris et le C6 du cHyd est en magenta. La longueur des liaisons covalentes C5-O5 et C5-C6 du cHyd sont indiquées. La liaison hydrogène entre l'azote de P1 et le O5 du cHyd est représentée par un pointillé vert.

Tout se passe comme si L/Fpg, en se fixant spécifiquement au duplexe cHyd/C, déplace l'équilibre en faveur de l'épimère R du C5, tout du moins dans les conditions de cristallisation que nous avons utilisées. La nature R de la configuration du C5 du cHyd dans la poche du site actif de Fpg est également supportée par le fait que l'on observe dans le complexe non-covalent une liaison hydrogène forte entre l'amine de P1 et le groupement fixé sur le C5 du cHyd qui est le plus proche d'elle (Figure 87). Or, cette interaction forte ne peut s'établir qu'avec un groupement hydroxyle, c'est-à-dire avec l'épimère R du C5. En effet, dans le cas de la stabilisation de l'épimère S du C5 dans la poche du site actif, l'hydroxyle en C5 du cHyd serait trop loin pour interagir avec P1, et le méthyle ne pourrait réaliser une liaison hydrogène forte. Cette interaction ne laisse donc aucun doute sur la nature de la configuration du C5 du dommage dans la poche du site actif. En conséquence, P1 pourrait être directement impliquée dans la sélection de l'épimère R du C5. Cela pourrait suggérer que cette sélection se réalise dans la poche du site actif de l'enzyme après extrusion de celui-ci de la double-hélice d'ADN et non pas au moment de la reconnaissance du dommage dans l'ADN. La configuration du C5 du dommage est encore confortée par la structure du complexe covalent dans lequel l'hydroxyle en C5 a été éliminé. L'amine N-terminale de

l'enzyme se retrouve covalamment reliée au C5 du dommage en lieu et place du groupement l'hydroxyle (Figure 81). Etant donnée la qualité de la carte de densité électronique observée pour le modèle structural du complexe covalent, on peut proposer que dans la réaction suicide, P1 attaque toujours du même côté le C5 de Hyd. Ainsi, seul l'épimère R du C5 serait susceptible de créer un pontage covalent avec l'enzyme. D'autre part, l'utilisation d'un analogue non métabolisable cHyd et la possibilité d'épimérisation de celui-ci au niveau de son C5, ainsi que la durée élevée de la cristallogénèse, expliquent qu'on arrive finalement à avoir 100 % de complexe covalent dans nos cristaux. Si nous considérons maintenant le complexe non-covalent formé entre *L/Fpg* et le duplexe cHyd/C comme équivalent à un complexe enzyme/substrat de Michaelis, cette analyse structurale ne nous permet pas d'expliquer pourquoi l'activité ADN glycosylase (attaque de P1 en C1' de l'Hyd) est beaucoup plus efficace que la réaction suicide, puisque P1 se retrouve à peu de chose près à égale distance de ses cibles C1' et C5.

Bien que les deux réactions (attaque en C1' ou en C5) catalysées par *L/Fpg* partagent des caractéristiques fonctionnelles communes (préférence pour la structure bicaténaire de l'ADN lésé, discrimination forte de la base opposée au dommage et implications des résidus catalytiques P1 et E2), elles manifestent une dépendance au pH très différente (Figure 76). L'activité Hyd-ADN glycosylase présente une courbe en cloche avec un optimal d'activité à pH 7,5. Ce type de comportement par rapport au pH est très similaire à celui observé pour l'activité Fapy-ADN glycosylase (B. Castaing, communication personnelle). La courbe de dépendance au pH de la réaction suicide quant à-elle présente une allure sigmoïde avec un optimal à pH 9. La dépendance au pH des activités d'une enzyme indique que l'état de protonation de un ou plusieurs résidus influe sur la catalyse enzymatique et qu'ils sont donc directement impliqués dans celle-ci. Dans le cas de l'activité ADN glycosylase, nous avons montré qu'au moins P1 et E2 sont indispensables. Conformément au mécanisme que nous avons proposé, E2 sous sa forme de base conjuguée (-COO-) pourrait-être impliquée dans l'activation de P1 *via* l'intervention d'une molécule d'eau (Coste *et al.*, 2008). Ainsi activé, P1 sous sa forme de base conjuguée (>NH) pourrait attaquer le C1' du nucléoside endommagé. Il s'ensuit une ouverture du désoxyribose associée à une attaque nucléophile de l'oxygène de l'hétérocycle en C4' sur le résidu E2 (sous sa forme acide conjugué, -COOH) qui aboutit à régénérer E2 dans son état déprotoné (ou base conjuguée) (Figure 88).



**Figure 88 :** Représentation schématique des mécanismes réactionnels d'attaque du cHyd en C1' ou en C5, aboutissant respectivement à la formation d'un pontage covalent entre *LIFpg* et l'Hyd ou à l'activité Hyd-ADN glycosylase.

L'ADN lésé est représenté en noir, Fpg en bleu, et des liaisons hydrogène sont représentées par des pointillés rouges. L'activité ADN glycosylase se déroule comme dans la Figure 32 de l'introduction, hormis pour le résidu E2 dont le rôle a été précisé. La réaction de pontage covalent se déroule via une substitution nucléophile de type  $S_N2$ , avec l'amine de P1 de Fpg comme nucléophile, et avec le groupement hydroxyle en C5 de l'Hyd comme groupement partant. Un réarrangement du produit de cette réaction permet le départ d'une molécule d'eau plutôt que d'un ion hydroxyle, et permet de conserver une charge neutre sur l'azote du résidu P1. La stéréospécificité de la réaction pour un C5 de l'Hyd en configuration R est donnée par une liaison hydrogène avec le résidu P1 de Fpg.

La courbe en cloche que l'on observe pour la dépendance au pH de l'activité Hyd-ADN glycosylase pourrait s'expliquer par le fait que E2 intervienne dans le mécanisme catalytique sous ses deux états de protonation (acide et base conjugués). Cela voudrait dire qu'au dessus d'un certain pH (ici pH 7,5), les conditions dans la poche du site actif de

l'enzyme ne sont plus réunies pour maintenir au moins de façon transitoire E2 sous sa forme protonée (acide conjugué), et cela à cause de la présence de trop d'ions  $\text{OH}^-$  dans le milieu. De même, en dessous d'un certain pH, la déprotonation du résidu P1 de l'enzyme serait rendue difficile par la présence de trop d'ions  $\text{H}_3\text{O}^+$  dans le milieu.

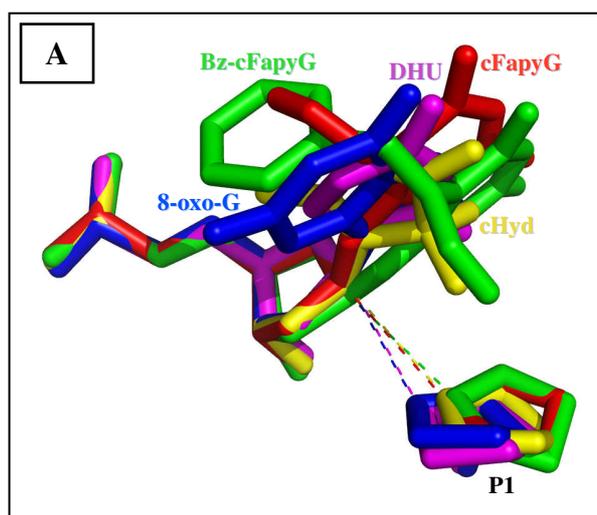
Dans le cas de la réaction suicide, nous avons également montré que P1 et E2 sont requis. Cependant au-delà de pH 7,5, l'activité continue à augmenter jusqu'au voisinage du pKa de la proline. A pH physiologique (pH 7,5), on peut penser que, comme dans le cas de l'activité ADN glycosylase, E2 doit activer P1 *via* une molécule d'eau. Ensuite, plus on se rapproche du pKa de P1, moins on a besoin de ce mécanisme pour déprotoner P1 et donc pour l'activer en vue de l'attaque en C5. Dans la réaction suicide, E2 est utile uniquement sous sa forme base conjuguée dans des conditions de pH physiologique (Figure 88). Dans des conditions de pH alcalin, P1 est activée de façon constitutive et selon ce schéma E2 n'est plus requis. L'ensemble de ces travaux permet donc de proposer un mécanisme moléculaire pour ces réactions catalysées par l'enzyme (Figure 88).

### **III. Conclusion générale et perspectives**

### III. Conclusion générale et perspectives

#### III.3. Conclusion générale

Les analyses structurales de complexes abortifs entre Fpg et des ADNs endommagés ont clairement établi que l'activité ADN glycosylase de Fpg nécessite l'extrusion du dommage hors de la double hélice d'ADN pour sa stabilisation dans la poche du site actif de l'enzyme (Fromme *et al.*, 2003a; Coste *et al.*, 2004; Banerjee *et al.*, 2006; Coste *et al.*, 2008). Cette stabilisation a pour but d'exposer la lésion aux résidus catalytiques de l'enzyme. Notamment, l'une des caractéristiques les plus importantes de l'enzyme Fpg est qu'elle utilise son amine N-terminale pour réaliser une attaque nucléophile en C1' des nucléosides endommagés, initiant ainsi son activité ADN-glycosylase (chapitre I.7.3 sur le mécanisme d'action de Fpg, Tchou *et al.*, 1995; Zharkov *et al.*, 1997; Sidorkina *et al.*, 2000). Bien que le mode de stabilisation du dommage extrahélical puisse varier de façon significative selon la nature de la base endommagée, la position relative du résidu P1 de Fpg par rapport au C1' des lésions ou analogues de lésions de bases nucléiques varie peu (Figure 89). Ainsi, quelle que soit l'origine (*L. lactis*, *B. stearothermophilus*) et la nature (type sauvage ou mutant inactif E2Q) de Fpg, l'azote du résidu P1 de Fpg se retrouve toujours bien positionné pour une attaque nucléophile en C1' du dommage extrahélical. Notre dernière structure du complexe formé entre *L*/Fpg et un duplexe cHyd/C n'échappe pas à cette observation (en jaune, Figure 89). Le duplexe Hyd/C étant un bon substrat pour l'enzyme (Figures 51 et 52), nous nous attendions en effet à ce que ce critère structural soit respecté dans notre structure avec l'analogue de lésion. Par contre, notre structure de *L*/Fpg complexée au Bz-cFapyG montre aussi un bon positionnement de P1 par rapport au C1' de l'analogue de lésion, or les FapyGs substitués en N7 par de gros adduits sont de mauvais substrats pour l'enzyme (Alekseyev *et al.*, 2004). De toute évidence, le bon positionnement du C1' du dommage par rapport à celui de l'amine de P1 de Fpg n'est pas suffisant pour expliquer une catalyse enzymatique efficace. Nous avons effectivement proposé que l'état de protonation de P1 est aussi un déterminant nécessaire à une catalyse efficace par Fpg (Coste *et al.*, 2008). En effet, même si P1 est bien positionné par rapport au C1' du dommage, il faut aussi qu'il soit déprotoné pour être nucléophile, et donc être capable d'attaquer le nucléoside endommagé en C1'.

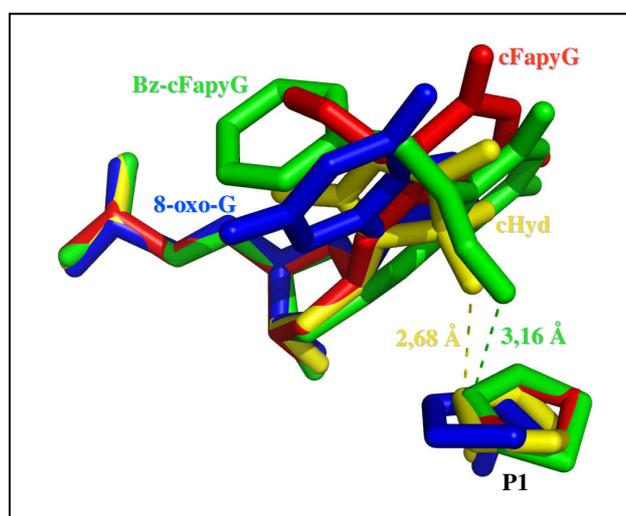


Lésion ou analogue de lésion	Origine de Fpg	Distance entre le N de P1 et le C1' de la lésion ou analogue de lésion	Publication de référence
8-oxo-G	<i>Bacillus</i>	3,46 Å	Fromme <i>et al.</i> , 2003a
DHU	<i>stearothermophilus</i>	3,83 Å	
cFapyG	<i>Lactococcus lactis</i>	3,76 Å	Coste <i>et al.</i> , 2008
Bz-cFapyG		3,57 Å	
cHyd		3,59 Å	Non publié

**Figure 89 : Positionnement du résidu P1 de Fpg par rapport aux différentes lésions ou analogues de lésions pour lesquels on a la structure d'un complexe avec l'enzyme.**

La superposition est faite sur les squelettes des nucléotides portant les lésions ou analogues de lésions, plus les atomes C $\beta$  et O des résidus 1 et 2 de la protéine. Les atomes issus de la structure du complexe *L/Fpg/cFapyG* sont représentés en rouge (code PDB 1XC8), ceux du complexe *L/Fpg/Bz-cFapyG* en vert (code PDB 3C58), ceux du complexe *BstFpgE2Q/8-oxo-G* en bleu (code PDB 1R2Y), ceux du complexe *BstFpgE2Q/DHU* en magenta (code PDB 1R2Z), et ceux du complexe *L/Fpg/cHyd* en jaune. (A) : Représentation de la superposition précédente. Les localisations des distances calculées entre le N de P1 et le C1' de chaque lésion ou analogue de lésion sont indiquées par des pointillés. (B) : Tableau indiquant les distances entre le N de P1 et le C1' de chaque lésion ou analogue de lésion.

Ainsi, nous avons émis une hypothèse s'appuyant sur l'état de protonation de P1 pour expliquer que les FapyGs substitués en N7 par des gros adduits sont de piètres substrats pour l'enzyme. En effet, dans la structure du complexe formé entre *L/Fpg* et le duplexe Bz-cFapyG, une liaison hydrogène est observée entre l'azote du résidu de P1 de l'enzyme et le O8 de l'analogue de lésion. Nous avons alors proposé que cette interaction pourrait stabiliser l'état protoné du P1 de l'enzyme (Chapitre II.2.1.4 et Coste *et al.*, 2008), diminuant ainsi l'activité de Fpg sur la lésion Bz-FapyG. En effet, cette interaction n'est présente ni dans la structure de *L/Fpg* complexée au duplexe cFapyG/C, ni dans la structure de *BstFpgE2Q* complexée au duplexe 8-oxo-G/C (Figure 90).

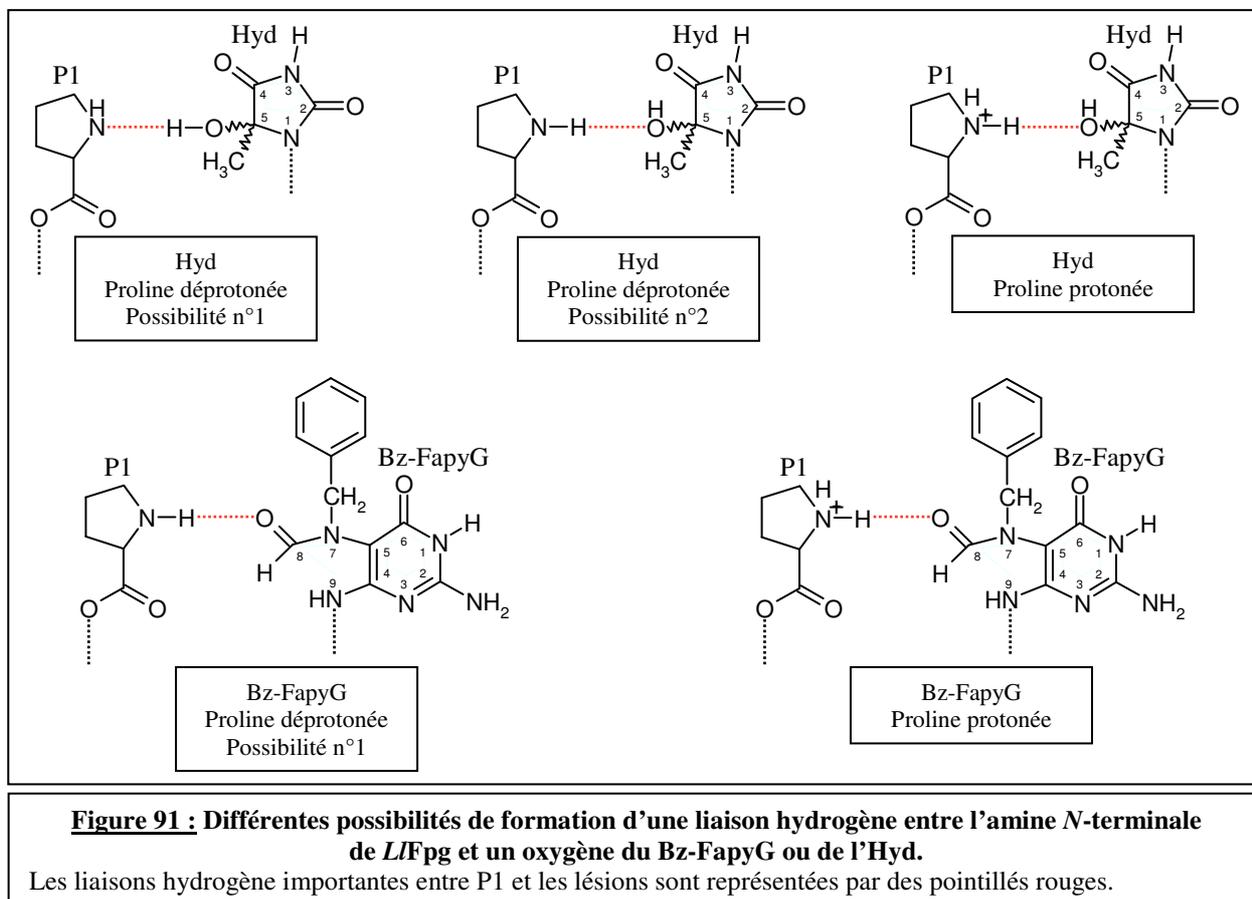


**Figure 90 : Présence d'une liaison hydrogène entre l'azote de P1 et un oxygène de certains analogues de lésions.**

La superposition est faite sur les mêmes atomes que pour la Figure 89. Les atomes issus de la structure du complexe *L/Fpg/cFapyG* sont représentés en rouge (code PDB 1XC8, Coste *et al.*, 2008), ceux du complexe *L/Fpg/Bz-cFapyG* sont représentés en vert (code PDB 3C58, Coste *et al.*, 2008), ceux du complexe *BstFpgE2Q/8-oxo-G* sont représentés en bleu (code PDB 1R2Y, Fromme *et al.*, 2003a), et ceux du complexe *L/Fpg/cHyd* sont représentés en jaune (ce travail). Les liaisons hydrogène sont représentées des pointillés, et leur longueur est indiquée.

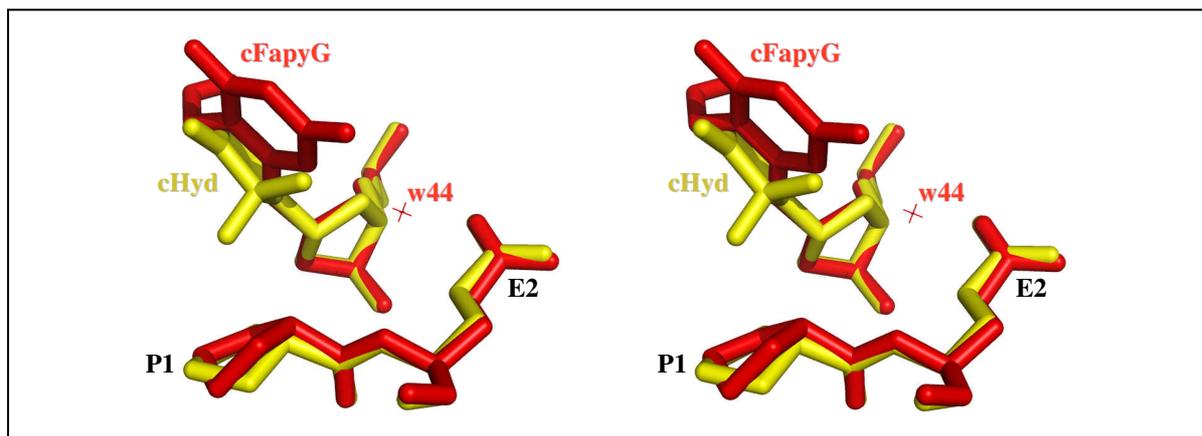
Cependant, une liaison hydrogène similaire est observée dans la structure du complexe *L/Fpg/cHyd* entre l'azote du résidu P1 de l'enzyme et le O5 de l'analogue de lésion. On peut même indiquer que cette dernière est plus forte que celle observée entre Fpg et le Bz-cFapyG, étant donné que leur longueur respective est de 2,68 Å (interaction forte) et 3,16 Å (interaction plus faible) (Figure 90). Si cette interaction était défavorable à l'activité de Fpg, alors on devrait avoir un effet plus fort de celle-ci avec l'Hyd qu'avec le Bz-FapyG. Sachant que le duplexe Hyd/C est un bon substrat pour Fpg (Figures 51 et 52), la présence d'une telle liaison hydrogène ne semble donc pas si défavorable à l'activité de Fpg. Cependant, l'analyse structurale par cristallographie ne peut pas nous renseigner sur l'état de protonation de P1 (les hydrogènes ne sont pas visibles à cause de leur faible densité électronique, et de la résolution insuffisante de notre structure). Ainsi, notre hypothèse concernant l'effet défavorable d'une liaison hydrogène avec le résidu P1 protoné de l'enzyme pourrait être erronée, car cette interaction est aussi possible avec une proline déprotonée (Figure 91). En fait, il est tout à fait possible que l'interaction observée dans notre structure entre P1 et le cHyd fasse intervenir une proline déprotonée (Figure 91), et donc compétente pour l'attaque nucléophile. Cette interprétation serait ainsi en accord avec nos résultats qui indiquent que le duplexe Hyd/C est un bon substrat de *L/Fpg* (Figures 51 et 52). D'autre part, il est tout à fait possible que la liaison hydrogène présente dans notre autre structure entre P1 et le Bz-cFapyG fasse intervenir une proline protonée (Figure 91), et donc incompétente pour l'attaque nucléophile.

Cette hypothèse irait d'ailleurs dans le sens des conclusions de notre publication (Coste et al., 2008), car elle permettrait d'expliquer pourquoi les FapyGs substitués en N7 par des adduits encombrants sont de mauvais substrats de Fpg (Alekseyev *et al.*, 2004).



En outre, la confrontation entre les études fonctionnelles et l'analyse structurale nous a conduits à proposer que l'activation de P1 par déprotonation serait assurée par le résidu E2 *via* une molécule d'eau (w44 dans le complexe formé entre LIFpg et le duplexe cFapyG/C, Figure 92 et Coste *et al.*, 2008). Cette hypothèse est supportée par des données de la bibliographie, qui indiquent que la Fpg de type mutant E2Q est complètement déficiente en activité, et qu'aucun équivalent de la w44 n'est retrouvé dans la structure de BstFpg de type mutant E2Q complexée à un duplexe 8-oxo-G/C (Lavrukhin *et al.*, 2000; Fromme *et al.*, 2003a). En fait, dans la structure du complexe avec la 8-oxo-G, le bon positionnement d'une molécule d'eau en lieu et place de la w44 est impossible, car le groupement 8-oxo de la lésion occupe cette position. Dans cette situation particulière, il est probable que le résidu Q2 contraigne la 8-oxo-G à adopter une conformation *syn* en contactant le groupement 8-oxo du dommage. D'autre part, notre hypothèse sur l'intervention d'une molécule d'eau est aussi étayée par nos résultats de mutagenèse dirigée sur le résidu E2, qui indiquent qu'un résidu acide à cette position est

nécessaire pour l'activité de *LIFpg* (Figures 44 et 45, et Tableau 9). En effet, le remplacement d'un résidu E2 (chaîne latérale acide) par un résidu dont la chaîne latérale est neutre ou basique (versions mutantes E2N, E2R, E2T, E2L, E2M et E2Q) abolit complètement l'activité de l'enzyme. Par contre, le mutant E2D (chaîne latérale acide, mais plus courte que pour le résidu initial) de *LIFpg* possède toujours une légère activité. Ainsi, on peut penser que l'acidité d'un aspartate permettrait de déprotoner le P1 de Fpg *via* la molécule d'eau évoquée précédemment, mais de façon moins efficace qu'un glutamate à cause de la différence de taille entre les chaînes latérales d'un E et d'un D. L'intervention d'une molécule d'eau dans le mécanisme catalytique que nous proposons est tout à fait en accord avec le mécanisme réactionnel de l'activité ADN glycosylase proposé à la fin de l'introduction du manuscrit (Figure 32). Après avoir arraché un proton à la molécule d'eau, le résidu E2 se trouve sous sa forme d'acide conjugué (protoné), et peut donc être l'acide responsable de la protonation de l'oxygène de l'hétérocycle du site AP qui permet sa stabilisation sous forme ouverte (Figure 32 et Figure 88). Sachant qu'aucun équivalent de la w44 n'est présent dans la structure de *LIFpg* complexée à un duplexe Bz-cFapyG/C, nous avons ainsi proposé que cette observation supportait le fait que les FapyGs substitués en N7 par des adduits encombrants sont de mauvais substrats pour l'enzyme. Il nous faut cependant nuancer notre hypothèse. En effet, aucun équivalent de cette molécule d'eau n'est présent dans la structure du complexe formé entre *LIFpg* et un duplexe cHyd/C (Figure 92), alors que le duplexe Hyd/C est un bon substrat pour Fpg (Figures 51 et 52). L'observation de cette molécule d'eau stabilisée dans le site actif par le résidu E2 n'est donc pas un critère structural suffisant pour conclure qu'une lésion est un bon substrat. En conséquence, l'absence d'une telle molécule d'eau avec le Bz-cFapyG n'est pas non plus un argument suffisant pour dire que le Bz-FapyG est un substrat médiocre pour Fpg. On peut cependant proposer que l'absence de cette molécule d'eau dans la structure avec le cHyd n'implique pas pour autant que cette molécule d'eau n'existe pas (l'emplacement de la w44 reste disponible dans la structure, Figure 92), mais que pour des raisons qui nous échappent, elle n'a pas été stabilisée lors de la cristallisation du complexe. De plus, dans l'hypothèse où cette molécule d'eau aurait déjà été utilisée pour activer P1 dans le complexe entre *LIFpg* et le duplexe cHyd/C (P1 serait déprotonée), on peut proposer qu'elle serait de ce fait moins bien ligandée, d'où son absence à la position attendue dans la structure.



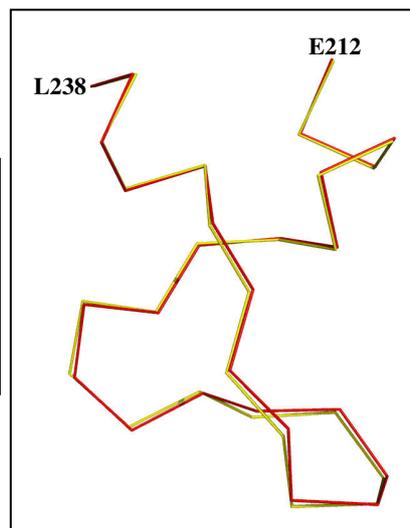
**Figure 92 :** Vue stéréoscopique montrant la présence d'une molécule d'eau à proximité du résidu E2 de Fpg dans la structure avec le cFapyG, et pas avec le cHyd.

La superposition est faite sur les mêmes atomes que pour la Figure 89. Les résidus P1 et E2 de Fpg et les analogues de lésions sont représentés pour chaque structure. Le complexe Fpg/cFapyG est représenté en rouge, et le complexe Fpg/cHyd en jaune. La molécule d'eau importante dans la structure du complexe Fpg/cFapyG est représentée par une croix rouge, et est notée **w44**.

Au cours de nos travaux, nous avons proposé qu'un autre élément structural pouvait expliquer les différences d'activité de Fpg sur ses substrats, qui concerne la conformation de la boucle  $\alpha$ F- $\beta$ 9 de Fpg (Coste *et al.*, 2008). Nos études comparatives de la conformation de cette boucle dans nos structures et dans celles déjà disponibles pour Fpg nous ont conduit à proposer qu'il existait deux conformations majeures de celle-ci, une conformation « ouverte » et une conformation « fermée » (Figure 5 de Coste *et al.*, 2008). Nous avons alors proposé que la conformation « fermée » était nécessaire pour la reconnaissance correcte des guanines oxydées et pour l'excision efficace de ces lésions. Cependant, cette conformation « fermée » n'est pas uniquement présente pour les structures de Fpg libre ou en complexe avec des ADNs contenant une purine oxydée (FapyG et 8-oxo-G, codes PDB 1XC8 et 1R2Y, respectivement, Figure 5 de Coste *et al.*, 2008), mais aussi dans la structure de la Fpg de *B. stearothermophilus* de type mutant E2Q complexée à un ADN contenant une DHU (pyrimidine oxydée) (code PDB 1R2Z, Fromme *et al.*, 2003a) (Figure 93). Dans ce cadre, on pourrait être tenté de proposer que la conformation « fermée » de la boucle  $\alpha$ F- $\beta$ 9 serait en fait une caractéristique nécessaire à la reconnaissance correcte des lésions oxydatives et pour l'excision efficace de ces lésions par Fpg (et non réservée uniquement aux purines oxydées). Cependant, la boucle est en conformation « ouverte » dans notre structure du complexe Fpg/cHyd (Figure 58C), alors que nous avons montré qu'un duplexe Hyd/C est un bon substrat pour Fpg (Figures 51 et 52). Ainsi, la structuration de la boucle en conformation « fermée » n'est pas une caractéristique structurale nécessaire et suffisante pour expliquer une bonne excision des lésions par l'enzyme.

**Figure 93 : Boucle  $\alpha$ F- $\beta$ 9 de *BstFpgE2Q* en conformation « fermée » dans les complexes avec des ADN contenant une 8-oxo-G ou une DHU.**

La superposition est faite sur les C $\alpha$  des résidus 212 à 238 et la vue est similaire à celle de la Figure 5 de Coste *et al.*, 2008.



De plus, la boucle  $\alpha$ F- $\beta$ 9 de *L/Fpg* adopte une conformation « fermée » dans le complexe formé avec le cFapyG, et une conformation « ouverte » dans le complexe formé avec le Bz-cFapyG pour des raisons d'accommodation du groupement benzyle (Figure 5A de Coste *et al.*, 2008). Nous avons alors émis l'hypothèse que cela constituait un bon argument pour dire que le Bz-FapyG était stabilisé dans la poche du site actif de Fpg dans une conformation catalytiquement non-productive. Cependant, cette boucle adoptant aussi une conformation « ouverte » dans la structure du complexe formé entre *L/Fpg* et le duplexe cHyd/C (Figure 58C), la structuration de la boucle en conformation « ouverte » avec le Bz-cFapyG n'est pas un élément suffisant pour dire que la lésion Bz-FapyG est stabilisée dans une conformation non productive par l'enzyme. L'implication de la boucle  $\alpha$ F- $\beta$ 9 de Fpg dans la catalyse enzymatique proprement dite doit donc être démontrée par d'autres études.

Par ailleurs, nous avons montré que *L/Fpg* est capable d'exciser efficacement l'Hyd et la 8-oxo-G quand celles-ci sont contenues dans un oligonucléotide double-brin, et appariées à une cytosine (Figure 51). Par contre, l'enzyme excise inefficacement ces deux lésions quand celles-ci sont contenues dans un oligonucléotide simple-brin (Figure 51). Ce résultat est en accord avec les données antérieures, qui indiquaient que Fpg est une enzyme spécifique de l'élimination des lésions contenues dans l'ADN double-brin (Tchou *et al.*, 1991). De plus, cette caractéristique peut assez bien s'expliquer structurellement parlant, étant donné que l'enzyme fait quelques interactions avec le brin opposé au dommage (Figure 28). En outre, le fait que cette spécificité ait été sélectionnée durant l'évolution n'est en fait pas étonnant, étant donné que l'enzyme Fpg nécessite d'avoir une base opposée à la lésion pour permettre la resynthèse de l'information génétique correcte par une ADN polymérase (chapitre I.5). D'autre part, nous avons montré que l'Hyd est inefficacement excisée par Fpg quand celle-ci

est appariée à une adénine dans un oligonucléotide double-brin (Figures 51 et 52). Ce résultat est aussi en accord avec les données de la littérature, qui indiquent que Fpg est capable d'exciser efficacement les lésions opposées à toute base nucléique hormis une adénine (Tableau 7 de l'introduction, et références correspondantes). De plus, la spécificité de Fpg pour la base opposée au dommage est aussi tout à fait en accord avec le rôle principal qui a été émis pour cette enzyme, en l'occurrence de lutter contre les effets mutagènes de la 8-oxo-G (système GO, chapitre I.6 et Michaels *et al.*, 1992b). Structuralement, cette spécificité est partiellement expliquée par la reconnaissance de la base opposée au dommage par une arginine conservée de l'enzyme (chapitre II.2.1.4). Par contre, bien qu'aucune des publications sur cet aspect de la spécificité de substrat de Fpg ne déroge réellement à la règle du « tout sauf un A opposé au dommage », il semble que quelques lésions amènent une moins forte spécificité de l'enzyme pour ce paramètre (Tableau 18).

**Tableau 18 : Différences d'activité de *EcFpg* en fonction de la base opposée à la lésion pour différentes lésions.**

Les noms des lésions sont les mêmes que dans la légende des Figures 29 et 30.

Lésion	Ordre de préférence pour la base opposée à la lésion	Publication(s) de référence
<b>Règle du « tout sauf un A opposé au dommage » strictement respectée</b>		
8-oxo-G	dT ≥ dG ≥ dC >> dA	Tchou <i>et al.</i> , 1991; Tchou <i>et al.</i> , 1994; Asagoshi <i>et al.</i> , 2000
	dG ≥ dT > dC >> dA	Zaika <i>et al.</i> , 2004
	dC ≥ dT > dG >> dA	Castaing <i>et al.</i> , 1993
N <sup>7</sup> -Me-FapyG	dC > dT ≥ dG > dA	Asagoshi <i>et al.</i> , 2000
FapyG	dC >> dA	Wiederholt <i>et al.</i> , 2003
Gh	dC ≈ dT ≈ dG >> dA	Leipold <i>et al.</i> , 2000
Sp	dC ≈ dT ≈ dG >> dA	Leipold <i>et al.</i> , 2000
FapyA	dG > dT > dA	Wiederholt <i>et al.</i> , 2002a
εA2 et εA3	dT > dC > dG >> dA	Speina <i>et al.</i> , 2001
5-OH-U	dG >> dA	Hatahet <i>et al.</i> , 1994
Xanthine	Excision avec dC et pas avec dG, dT et dA	Wuenschell <i>et al.</i> , 2003
Ac.HI-urée	dC ≥ dG ≥ dT > dA	Jurado <i>et al.</i> , 1998
Site AP	dC >> dA	Tchou <i>et al.</i> , 1994
Hyd	dC >> dA	Notre travail (Figures 51 et 52)
<b>Règle du « tout sauf un A opposé au dommage » moins strictement respectée</b>		
DHU	Equivalente pour les quatre bases	Zaika <i>et al.</i> , 2004
5-OH-C	dG ≈ dA	Hatahet <i>et al.</i> , 1994
Ug	dG ≥ dA	Purmal <i>et al.</i> , 1998

Cette discrimination négative d'un A opposé au dommage peut être également associée à d'autres paramètres, tels que des effets de l'environnement séquentiel du dommage et/ou des effets dus à la nature du dommage lui-même (déformabilité/stabilité de la paire de base au niveau du dommage, ou présence/absence de déterminants structuraux du dommage dans l'un des sillons de l'ADN, etc). Finalement, les deux spécificités de Fpg évoquées précédemment (contexte double brin de la lésion, et appariement à une base autre qu'une adénine) apparaissent comme des déterminants essentiels pour une bonne activité de l'enzyme.

Enfin, nos données indiquent que le duplexe Hyd/A, qui correspond au seul appariement de l'Hyd qui soit physiologiquement relevant (Figure 50; Halliwell *et al.*, 1991; Gasparutto *et al.*, 2000), est nettement moins bien réparé par *LIFpg* que par *EcNth* (Figure 52), ce qui n'est pas en accord avec une comparaison des Fpg et Nth de *E. coli* (Gasparutto *et al.*, 2009). Ces différences impliquent donc que l'on ne peut pas réellement conclure à partir de nos données quant à l'enzyme majoritairement responsable de l'excision de l'Hyd *in vivo* chez *L. lactis*. Pour ce faire, il aurait fallu que l'on compare les Fpg et Nth provenant de cette bactérie. Nos résultats pourraient donc indiquer une différence de spécificité de substrat entre *EcFpg* et *LIFpg*.

Un dernier point important ressortant de nos études structure/fonction sur *LIFpg* concerne la capacité de Fpg à former un pontage covalent avec la 5-Hydroxy-5-Méthyle-Hydantoïne (chapitre I.2.2). En effet, nous avons montré que cette réaction a lieu entre l'azote du résidu P1 de l'enzyme et le C5 en configuration R de la lésion Hyd ou de son analogue cHyd. De plus, cette réaction est plus efficace avec le cHyd qu'avec l'Hyd, ce qui s'explique par une compétition entre cette réaction et l'activité ADN glycosylase de Fpg pour l'Hyd (compétition qui n'est pas possible pour l'analogue non-métabolisable). Ainsi, le cHyd correspond bien à la définition d'un substrat suicide pour *LIFpg*. L'effet du pH sur la réaction suicide avec le cHyd et sur l'activité Hyd-ADN glycosylase, nous ont permis de proposer un mécanisme réactionnel pour ces réactions (Figure 88). Les mécanismes que nous proposons tiennent compte de l'état de protonation des résidus P1 et E2 de l'enzyme, ainsi que de la stéréospécificité des réactions par rapport à la configuration de l'atome C5 de la lésion. On peut noter ici que notre structure d'un tel complexe covalent entre une ADN-glycosylase et un substrat suicide est le premier exemple de ce type de complexe dans la littérature. Deux autres cas de substrats suicides avec des ADN glycosylases sont connus (la 2-désoxyribonolactone : Hashimoto *et al.*, 2001; Kroeger *et al.*, 2003; et l'oxanine : Nakano *et al.*, 2003), dont un seul avec la partie base nucléique d'une lésion (l'oxanine). De plus, la réaction que nous avons

mise en évidence avec le cHyd semble être plutôt plus spécifique que ne le sont les deux exemples précédents. Cependant notre étude de pontage avec différentes protéines n'est que préliminaire, étant donné qu'il aurait fallu tester de nombreuses autres protéines se fixant à l'ADN afin de préciser nos résultats.

### **III.3. Perspectives**

#### **- Reconnaissance des lésions encombrantes par Fpg et Ogg1**

Nous avons proposé que le mode particulier de reconnaissance du  $N^7$ -Benzyle-FapyG par *L*Fpg conduisait à l'apparition d'un complexe catalytiquement non- ou peu-productif (Coste *et al.*, 2008). Sur la base des données structurales, nous suggérons que ce mode de reconnaissance du Bz-FapyG pourrait être une stratégie moléculaire générale de Fpg pour accommoder dans son site actif toute purine à cycle ouvert substitué en  $N^7$  par un adduit plus encombrant qu'un groupe « méthyle » (Figure 29). Cette proposition permet de fournir une explication moléculaire raisonnable qui soit en accord avec l'observation que ces lésions sont de médiocres substrats pour Fpg. Afin d'étayer cette hypothèse, une étude structurale par cristallographie des complexes formés entre *L*Fpg et des ADNs contenant les analogues carbocycliques  $N^7$ -Me-cFapydG,  $N^7$ -Et-cFapydG et  $N^7$ -AFB1-cFapydG doit être entreprise. Si notre proposition est correcte, le Me-FapyG devrait être stabilisé dans la poche du site actif de Fpg de la même manière que le FapyG, tandis que l'Et-cFapyG et l'AFB1-cFapyG devraient l'être de manière similaire au Bz-cFapyG. En parallèle, le mode de reconnaissance des purines à cycle imidazole substitué en  $C^8$  par de gros adduits tels que l'aminofluorène pourrait être aussi étudié par cristallographie (Figure 29) (Boiteux *et al.*, 1989). On peut d'ailleurs dire qu'à cet effet, l'analogue carbocyclique du  $C^8$ -AF-irodG est en cours de synthèse chez nos collègues chimistes de Munich.

Au cours de cette étude, nous avons montré également que la protéine Ogg1 (homologue fonctionnel de Fpg chez les eucaryotes) reconnaît indifféremment un ADN contenant un FapyG et un Bz-FapyG (Coste *et al.*, 2008). Ceci suggère que comme Fpg, Ogg1 est capable d'accueillir des lésions encombrantes malgré la petite taille apparente de la poche de son site actif (Bruner *et al.*, 2000). Pour être complète, cette étude structurale devra donc être étendue à la reconnaissance des lésions encombrantes par la protéine Ogg1.

### - Sélection de la conformation *syn/anti* de la liaison *N*-glycosydique des purines oxydées dans la poche du site actif de Fpg

Concernant la reconnaissance des purines oxydées, nous disposons de 2 modèles structuraux : un complexe entre une protéine Fpg de type « mutant inactif » E2Q provenant d'une bactérie thermophile (*B. stearothermophilus*) et un oligonucléotide contenant une 8-oxoG (Fromme *et al.*, 2003a) et un complexe entre une protéine de type sauvage provenant d'une bactérie mésophile (*L. lactis*) et un oligonucléotide contenant un analogue de FapyG, le cFapyG (Coste *et al.*, 2004; Coste *et al.*, 2008). La comparaison de ces modèles suggère que Fpg peut stabiliser la base endommagée en position extrahélicale en conformation *syn* pour la 8-oxoG ou *anti* pour le FapyG. Ces données structurales peuvent s'interpréter de plusieurs façons : **(1)** la protéine Fpg a sélectionné plusieurs modes de stabilisation des purines oxydées dans son site actif (conformation *anti* pour les résidus FapyG et *syn* pour la 8-oxoG); **(2)** en fonction de l'origine de l'enzyme (bactéries mésophiles ou thermophiles), les protéines Fpg n'ont pas retenu le même mode de stabilisation du substrat dans leur site actif; **(3)** il n'y a qu'un seul mode de reconnaissance (en *anti*) et le remplacement de E2 par Q (Fromme *et al.*, 2003a), étant associé à une interaction entre l'amine de Q2 et l'oxygène du groupement C8-oxo, contraint la protéine *BstFpg* mutante E2Q à sélectionner de la conformation *syn* du lien *N*-glycosidique de la 8-oxoG; **(4)** l'analogue de FapyG non-métabolisable par l'enzyme, le cFapyG, dans lequel l'oxygène hétérocyclique du sucre a été remplacé par un carbone (Coste *et al.*, 2004; Coste *et al.*, 2008) serait responsable de la stabilisation de l'analogue de base en conformation *anti*; **(5)** du fait de la thermostabilité de la protéine Fpg de *B. stearothermophilus* (*BstFpg*), la cristallisation à température ambiante du complexe avec l'ADN contenant une 8-oxoG pourrait avoir piégé l'enzyme thermostable dans un complexe abortif non physiologique. Dans ce cas, la 8-oxoG adopterait une conformation *syn* qui ne serait pas productive pour une catalyse qui s'effectue normalement à des températures supérieures à 50°C. Chez *L. lactis* comme chez *E. coli* (Lavrukhin *et al.*, 2000), ce mutant n'est pas complètement déficient en activité ADN glycosylase (2000 fois moins actif que la protéine de type sauvage). Dans les essais de cristallisation à 20°C de *LIFpg* mutante E2Q complexé avec un duplexe d'ADN contenant une 8-oxoG, l'ADN est progressivement clivé par la protéine et le mélange hétérogène obtenu ne cristallise pas (Castaing, communication personnelle). Dans le cas de *B. stearothermophilus*, la protéine *BstFpg* de type mutant E2Q doit être complètement inactive à 20°C, ce qui doit permettre de la cristalliser sans problème avec un substrat vrai.

Même si nous ne pouvons conclure en l'état actuel des connaissances, nous avons la conviction que la sélection en *anti* de la purine oxydée est le mode de reconnaissance du dommage efficace pour sa métabolisation par Fpg, d'autant plus que la protéine hOgg1 stabilise elle aussi la 8-oxoG en *anti* dans la poche de son site actif (Bruner *et al.*, 2000). Pour trier parmi ces hypothèses, nous sommes en mesure de résoudre les structures de la protéine Fpg de *L. lactis* de type sauvage ou de type mutant inactif (E2Q et autres mutants de E2) engagée dans un complexe avec un duplexe d'ADN contenant un cFapyG ou une c8-oxoG. Durant mon travail de thèse, j'ai obtenu des cristaux de bonne qualité de LIFpg de type sauvage complexée à un duplexe c8-oxoG/C. Malheureusement, seul le cyclopentane du c8-oxoG est visible dans le modèle structural et aucune densité électronique n'est associée à la partie base nucléique du dommage, suggérant que la 8-oxoG reste très dynamique dans la poche du site actif (rotation libre autour de la liaison *N*-glycosydique). Cette dynamique de la 8-oxoG est en accord avec des simulations de dynamique moléculaire proposant que la 8-oxoG peut être stabilisée indifféremment en *syn* ou *anti* contrairement au cFapyG qui lui serait préférentiellement stabilisé en *anti* (Perlow-Poehnelt *et al.*, 2004; Zaika *et al.*, 2004), la conformation *anti* étant légèrement favorisée. L'utilisation des protéines LIFpg de type mutant sur E2 devrait nous permettre d'éliminer certaines des hypothèses que nous avons formulées plus haut. Afin d'écartier tout biais lié à l'utilisation de analogues carbocycliques de lésions, il serait aussi souhaitable de réaliser au moins la structure d'un complexe entre l'enzyme thermostable BstFpg de type sauvage avec un ADN contenant une c8-oxoG et/ou un cFapyG.

#### **- Rôle fonctionnel de la boucle $\alpha$ F- $\beta$ 9 dans l'activité des enzymes de la superfamille de Fpg**

L'ensemble des données structurales sur les protéines de la sous-famille structurale de Fpg indique que la boucle dynamique  $\alpha$ F- $\beta$ 9 peut adopter une conformation « fermée » qui se replie en contactant certains dommages (8-oxoG, FapyG, DHU) et une (ou des) conformation(s) « ouverte(s) » et dynamique(s) clairement observable(s) dans des complexes entre Fpg et des ADNs contenant certaines lésion ou des sites abasiques. Le caractère dynamique de cette boucle repose sur 3 glycines strictement conservées dans les séquences primaires de Fpg (G216, G226 et G229 chez LIFpg, Figure 22; Figure 5 de Coste *et al.*, 2008) qui peuvent jouer le rôle de « charnières » ou « rotules ». Malgré des tailles et des séquences très variables de la boucle, quand elles se replient sur le dommage, elles adoptent des conformations « fermées » quasi identiques (Fromme *et al.*, 2003; Coste *et al.*, 2008). Nous et

d'autres avions pensé que la conformation « fermée » était requise pour la stabilisation de la base endommagée dans la poche du site actif dans une conformation clivable par l'enzyme. Ainsi, la boucle  $\alpha$ F- $\beta$ 9 adoptant une conformation de type ouverte dans le complexe formé avec le Bz-cFapyG, nous avons proposé que cette conformation « ouverte » pouvait être associée au fait que les lésions encombrantes étaient des substrats médiocres. Hors nos derniers travaux structuraux sur les complexes entre *L*/Fpg et la 5-hydroxy-5-Méthyle-Hydantoïne (Hyd, lésion très peu encombrante) révèlent que la boucle reste « ouverte » alors que le duplexe d'ADN Hyd/C est un bon substrat pour l'enzyme.

En l'état actuel de nos connaissances, il est donc difficile d'accorder un rôle fonctionnel essentiel à la conformation « fermée » de la boucle  $\alpha$ F- $\beta$ 9. La seule donnée qui permette encore de soutenir un rôle fonctionnelle indispensable de la boucle réside dans l'observation que la version de type mutant Y222A de *L*/Fpg de l'enzyme est complètement défectueuse en activité 8-oxoG-ADN glycosylase (Tableau 9).

L'alignement des séquences primaires de Fpg indique que dans tous les cas, on retrouve une Tyrosine ou une Phénylalanine à cette position dans la boucle. Mises à part les 3 glycines strictement conservées (G216, G226 et G229), c'est le seul résidu contenu dans la boucle qui soit bien conservé. Cette dernière observation nous a conduits à formuler l'hypothèse que ce résidu aromatique (Y ou F) pourrait interagir avec la base endommagée (*via* des interactions de types  $e^-/e^-\pi$  ou  $e^-\pi /e^-\pi$ ) après le clivage du lien *N*-glycosidique. Les interactions de la boucle avec la base lésée clivée (*via* les groupements « amino » de la chaîne principale des résidus S217, I219, R220, T221 et Y222 et *via* la chaîne latérale de Y222) associées à la dynamique intrinsèque de la boucle (passage d'une conformation « fermée » à « ouverte ») permettrait à l'enzyme de rompre les interactions fortes établies avec la base lésée dans la poche du site actif et ainsi d'évacuer la base excisée hors de la poche (à l'image d'une catapulte). Suite à la libération de la base, la boucle peut alors se relâcher et adopter une conformation « ouverte » comme on l'observe dans la structure de l'enzyme complexée à un ADN contenant un site abasique (Serre *et al.*, 2002 ; Pereira de Jésus *et al.*, 2005). Sous sa forme « ouverte ou relâchée », la boucle n'autoriserait plus le retour de la base libre dans le site actif de Fpg. Dans le cadre de cette hypothèse, l'évacuation de la base endommagée clivée serait indispensable pour le clivage du site abasique (AP lyase). Cela revient à dire que les activités ADN glycosylase et AP lyase de Fpg sont découplées. Ce découplage entre les 2 activités a été clairement démontré pour la protéine Ogg1 et il a même été proposé que Ogg1 se dissocie de l'ADN contenant le site AP après l'activité ADN glycosylase avant de se réassocier à l'ADN pour enclencher l'activité AP lyase (Vidal *et al.*, 2001). En présence d'un

excès de 8-oxoG libre dans le milieu réactionnel, Ogg1 est même capable de catalyser la réaction inverse de l'AP lyase (religation d'un site AP clivé, Chung *et al.*, 2004). Toutes ces données semblent donc indiquer que la base lésée clivée qui serait encore capable de rester dans le site actif pourrait inhiber l'activité AP lyase de hOgg1. Afin d'éviter cette situation, Fpg utiliserait le caractère dynamique de la boucle  $\alpha$ F- $\beta$ 9, ce qui lui permettrait du même coup d'éviter de se dissocier de l'ADN contenant le site abasique comme cela est proposé pour Ogg1. Cette stratégie moléculaire permet à Fpg d'avoir un rendement ADN glycosylase/AP lyase quasi quantitative entre les 2 activités ce qui est loin d'être le cas de Ogg1 (certains auteurs proposent même que l'activité AP lyase de Ogg1 n'est pas opérationnelle *in vivo*). Ce mécanisme actif d'extraction de la base oxydée clivée par Fpg serait d'autant plus nécessaire que celle-ci est fortement ligandée dans la poche du site actif. Dans le cas contraire (c'est le cas de l'Hyd), la base lésée pourrait diffuser facilement en dehors du site actif sans la nécessité de faire intervenir un mécanisme actif utilisant la boucle  $\alpha$ F- $\beta$ 9 (la boucle adopte une conformation « ouverte » lorsque l'enzyme est fixé à un ADN contenant un cHyd).

Pour éprouver l'hypothèse d'un lien entre la dynamique de la boucle  $\alpha$ F- $\beta$ 9 et un découplage des activités ADN glycosylase et AP lyase de Fpg, une étude structure-activité de *L*/Fpg après le ciblage par mutagenèse dirigée de la boucle serait nécessaire. Dans un premier lieu, on pourrait examiner la question en faisant une délétion complète de la boucle  $\alpha$ F- $\beta$ 9. Il serait possible également d'atténuer fortement le caractère dynamique de la boucle en remplaçant par exemple les glycines conservées (G216, G226 et G229) par des prolines (de façon ponctuelle, 2 à la fois, ou les 3 en même temps). Un ciblage systématique de Y222 pour des substitutions devrait permettre d'éclaircir le rôle de ce résidu très conservé. D'autre part, l'étude fonctionnelle de ces protéines de type mutant (évaluation des activités ADN glycosylase, ADN glycosylase/AP lyase et AP lyase sur des ADN contenant une 8-oxoG, une Hyd ou un site AP, et effet de la 8-oxoG ou du FapyG libre sur l'activité AP lyase) devra être associée à l'étude structurale des complexes correspondant avec l'ADN lésé (cFapyG et cHyd). Il serait également intéressant d'étendre cette étude structure-activité aux protéines de la sous-famille structurale de l'endonucléase VIII (Nei, ADN glycosylase/AP lyase spécifique des pyrimidines oxydées) qui possèdent également une boucle de plus petite taille que celles des Fpg (Figure 5 de Coste *et al.*, 2004; Zharkov *et al.*, 2002). Jusqu'à aujourd'hui, il n'existe aucune donnée structurale sur le mode de reconnaissance des pyrimidines oxydées par Nei. Dans ce cadre, le cHyd constitue un outil de choix pour initier ce type d'étude. Il n'est pas exclu non plus d'utiliser le cFapyG pour le même type d'approches car il a été montré que les

protéines de la sous-famille de Nei (Nei et NEIL1) sont capables d'exciser les résidus Fapy (Wiederholt *et al.*, 2005).

### **- Conception d'inhibiteurs suicides de Fpg et/ou Ogg1**

Les dernières perspectives qui découlent de mon travail de thèse concernent la possibilité d'inhiber sélectivement et irréversiblement la protéine Fpg (et/ou Ogg1) : le cHyd dans l'ADN étant le premier prototype de tels inhibiteurs « suicides ». En effet, l'inhibition suicide et sélective de hOgg1 pourrait permettre de diminuer significativement la radiorésistance de certains cancers (Weiss *et al.*, 2005) et, dans le cas de maladies neurodégénératives, de stopper l'évolution de la maladie (maladie de Huntington et d'Alzheimer, McMurray, 2008). L'idée ici est de considérer la protéine hOgg1 comme une cible pharmacologique (ce concept reste à valider). Nous avons ouvert la voie pour l'exploration de la faisabilité de cette stratégie en montrant pour la première fois que le dérivé carbocyclique de la 5-hydroxy-5-méthyl-hydanthoine (cHyd) est un substrat suicide pour les protéines Fpg et Ogg1. Nous avons établi les bases biochimiques et structurales de la réaction suicide avec Fpg en résolvant la structure cristallographique du complexe covalent enzyme-ADN. Le même type d'études avec la protéine hOgg1 pourrait donc être conduit. Sur la base des données structurales, il sera possible d'examiner ensuite avec nos collègues chimistes les possibilités de modifier les substituants de cHyd dans des stratégies nucléosides, nucléotides et/ou oligonucléotides, l'ensemble de ces modifications du cHyd ayant pour but de rendre la réaction suicide plus efficace et sélective. L'approche structurale permet ici de rationaliser le « *design* » des dérivés du cHyd. L'action de ces nouvelles molécules sur les autres ADN glycosylases du système BER procaryotes et eucaryotes devra être analysée afin d'établir les bases moléculaires et structurales de la sélectivité de ces substrats suicides. L'effet de ces nouvelles molécules devra être ensuite évalué *in vivo* sur la radiorésistance de cellules tumorales (nous pensons en effet qu'en présence de ce type de molécules, un stress oxydant modéré sera suffisant pour induire une mort cellulaire) et sur des souris modèles pour la maladie d'Huntington pour laquelle il a été montré que Ogg1 et les stress oxydants contribuent significativement à l'évolution de la maladie (McMurray, 2008).

Toujours sur la base des travaux obtenus dans le cadre de ma thèse, nous pourrions tirer partie de la conformation du Bz-cFapyG stabilisée dans la poche du site actif de Fpg pour concevoir des inhibiteurs « suicides » de Fpg (et/ou hOgg1). L'adduit encombrant en  $N^7$  du dommage contraint l'enzyme à stabiliser un rotamère du groupe formamido particulier qui

conduit à une interaction entre l'oxygène du formyle en  $C^8$  du dommage et l'amine du nucléophile P1 de Fpg. En exploitant cette sélection rotamérique, il devrait être possible de positionner idéalement un groupe électrophile en  $C^8$  du Bz-cFapyG susceptible d'être attaqué par P1 et compatible avec les contraintes spatiales de la poche du site actif. Par exemple, la substitution dans le Bz-cFapyG de l'hydrogène du formyle par des groupements électrophiles peu encombrants tels qu'un thiocarbamate, un chloroacétamide ou un acrylamide devrait être possibles. L'attaque de ce type de groupements réactifs par P1 conduirait à la formation d'un complexe covalent irréversible entre l'ADN lésé et l'enzyme. Pour étendre cette étude à la protéine hOgg1, il faudrait établir d'abord les bases structurales de la reconnaissance du Bz-cFapyG par l'enzyme comme nous l'avons proposé au début de ce paragraphe.

## **IV. Matériel et méthodes**

## IV. Matériel et méthodes

### IV.1. Matériel biologique utilisé

#### IV.1.1. Mutagenèse dirigée sur LIFpg

Le plasmide utilisé est le plasmide pFlag-CTC, contenant le gène de résistance à l'ampicilline, et dans lequel a été sous-cloné au niveau du site multiple de clonage la phase ouverte du gène *Llfpg* codant pour la protéine Fpg de type sauvage de *Lactococcus lactis* (LIFpg), donnant ainsi le plasmide pFlag-*Ll-fpg*. Le gène *Llfpg* est placé sous dépendance d'un promoteur *tac* inductible en utilisant l'IPTG comme dé-répresseur, et la séquence du plasmide pFlag-*Ll-fpg* a été légèrement modifiée de telle sorte que l'on produise une protéine à laquelle on n'ait ajouté aucun acide aminé en N-terminal, et cela car le résidu N-terminal de Fpg est crucial pour son activité.

A partir de ce plasmide, nous avons généré six versions mutantes de la protéine par mutagenèse dirigée au niveau du deuxième codon du gène. Les mutations induites au niveau de la séquence de LIFpg sont les suivantes : E2N, E2R, E2T, E2L, E2M, E2D et E2Q. La mutagenèse dirigée a été faite grâce au système commercial « QuickChange Site-Directed Mutagenesis Kit » (Stratagene), qui permet de modifier le plasmide initial pour obtenir la séquence codant pour la version mutante de la protéine *via* une PCR un peu particulière. La technique est similaire à une PCR classique mais différente dans le choix des amorces. Pour insérer la mutation, il suffit de choisir des amorces s'hybridant toutes les deux au niveau de la zone comprenant la future mutation et de substituer volontairement le codon ciblé par la séquence du codon voulu. L'ensemble des autres étapes est similaire à une PCR classique, où chaque cycle d'élongation permet l'amplification de tout le plasmide, hormis la zone de mutation qui est apportée par les amorces. En fin d'expérience, le mélange PCR résultant est composé de deux populations : la matrice qui correspond au plasmide contenant la séquence de la protéine non mutée et les plasmides nouvellement synthétisés qui contiennent la mutation désirée. Une différence issue de leur origine nous permet alors de discriminer ces deux populations plasmidiques. Les plasmides matrices sont issus d'une préparation bactérienne alors que les plasmides contenant la mutation ont été synthétisés *in vitro*. De ce fait, nous savons que les plasmides matrices sont nécessairement méthylés, alors que les

plasmides néo-synthétisés ne le sont pas. Pour éliminer la matrice, nous avons donc réalisé une digestion par l'enzyme Dpn1 qui ne clive que les molécules d'ADN méthylées.

Le produit de la digestion est ensuite utilisé pour transformer des bactéries ultra compétentes (fournies dans le système commercial). Cette étape a deux objectifs : premièrement celui de réaliser les ligations entre les amorces et les brins néo-synthétisés et deuxièmement celui d'obtenir, après extraction, une quantité importante de plasmide muté. Toutes les étapes d'ici à ce que l'on ait extrait les plasmides des bactéries sont réalisées en milieu stérile afin d'éviter les contaminations bactériennes.

Pour la transformation bactérienne, on réalise le mélange suivant pour chaque plasmide:

- 50  $\mu$ L de solution de bactéries ultra-compétentes (XL1-Blue Supercompetent cells, Stratagene).
- 1  $\mu$ L de mélange de digestion par Dpn1.

La transformation des cellules est alors réalisée dans les conditions suivantes :

- 30 minutes à 4°C (permet l'entrée du vecteur dans les cellules).
- 45 secondes à 42°C (lyse des cellules de trop grande fragilité).
- 2 minutes à 4°C (arrête la lyse des cellules).

Ensuite, 0,5 mL de milieu NZY+ (voir annexe) est ajouté et l'ensemble est incubé à 37°C sous agitation (225-250 rpm) pendant une heure. 15  $\mu$ L et 150  $\mu$ L de la suspension bactérienne obtenue sont respectivement étalés sur deux boîtes de Pétri contenant du milieu LB/agar/ampicilline (voir annexe), suivi d'une incubation sur la nuit à 37°C. Ce protocole est appliqué avec les huit plasmides recombinants.

Une colonie isolée est alors reprise dans 25 mL de milieu LB/ampicilline. Arrivée à saturation ( $DO_{600nm} > 2,1$ ), la culture est centrifugée (4500g, 30 minutes, 4°C) et le surnageant est éliminé. Afin d'obtenir le plasmide, nous avons eu recours à un système commercial de midi-préparation d'ADN plasmidique (Nucleobond Xtra Midi, Macherey-Nagel), permettant l'extraction et la purification rapide de l'ADN plasmidique.

Enfin, les plasmides ont subi une concentration par précipitation éthanolique, et une petite quantité de chaque plasmide a alors été expédiée pour séquençage (société MWG) afin de vérifier que l'on avait bien la mutation voulue pour chaque plasmide.

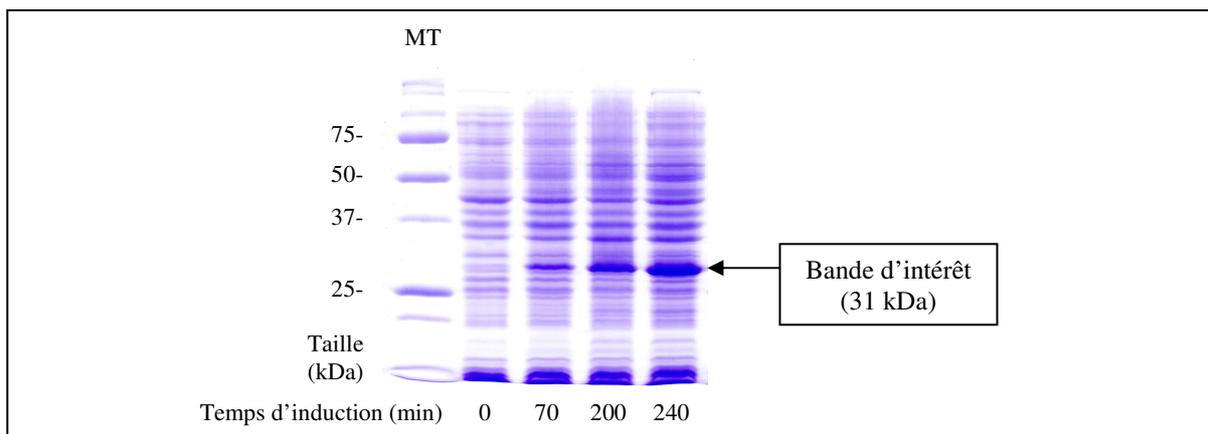
#### IV.1.2. Production de protéines recombinantes *LIFpg* WT et mutantes

La souche bactérienne utilisée est une souche BH540 (*fpg* :: *Kn*) de *Escherichia coli* dont le gène codant pour la protéine Fpg a été inactivé par insertion (recombinaison homologue) dans sa séquence du gène de résistance à la kanamycine (*Kn*).

Les huit plasmides recombinants ont été utilisés pour transformer cette souche bactérienne selon le protocole exposé dans le sous-chapitre précédent. De plus, une neuvième transformation a été faite en utilisant le plasmide pFlag-CTC intacte, ne contenant donc pas le gène *fpg*.

##### IV.1.2.1. Cultures de 250 mL pour les extraits cellulaires

Un clone isolé sur boîte LB/agar/ampicilline est prélevé afin d'ensemencer une pré-culture de 10 mL de milieu LB/ampicilline. Après une nuit à 37°C sous agitation, la culture saturée obtenue est centrifugée (4500g, 30 minutes, 4°C) et reprise dans 1 mL de milieu et utilisée pour ensemencer une culture de 250 mL de LB/ampicilline, ce qui permet de commencer à une  $DO_{600nm}$  de environ 0,2. La culture est placée dans un bain marie à 37°C sous agitation. Lorsque que la  $DO_{600nm}$  est comprise entre 0,6 et 0,8, la production de *LIFpg* (WT ou mutantes) est induite par ajout d'IPTG dans le milieu (1 mM final). Au cours de la production, des prélèvements sont effectués permettant de suivre par SDS-PAGE (voir annexe) la quantité de Fpg surproduite (Figure 94).



**Figure 94 : Gel SDS-PAGE représentant l'évolution de la quantité de *LIFpg* WT produite en fonction du temps d'induction.**

Le gel est un SDS-PAGE à 12 % (voir annexe). Des marqueurs de tailles protéiques (MT) sont présents dans la première piste (Precision Plus Protein Unstained Standards, Biorad). Les temps d'induction sont indiqués sous le gel.

On voit aussi que toutes les formes de Fpg que l'on voulait obtenir sont surproduites. Par contre, la quantité de plusieurs des mutants de Fpg n'évoluant plus après 2 heures d'induction, nous avons choisi cette durée pour nos essais. Après deux heures d'induction, la culture est arrêtée par centrifugation (4500g, 30 min, 4°C) et les culots sont repris deux fois dans une solution de lavage (voir annexe) et transférés dans un Falcon pour être ensuite pesés et conservés à -20°C.

#### **IV.1.2.2. Cultures de 1L pour les purifications protéiques**

La procédure est rigoureusement la même que pour les cultures de 250 mL (voir le sous-chapitre précédent), hormis le fait que les 10 mL de pré-culture servent ici àensemencer 1L.

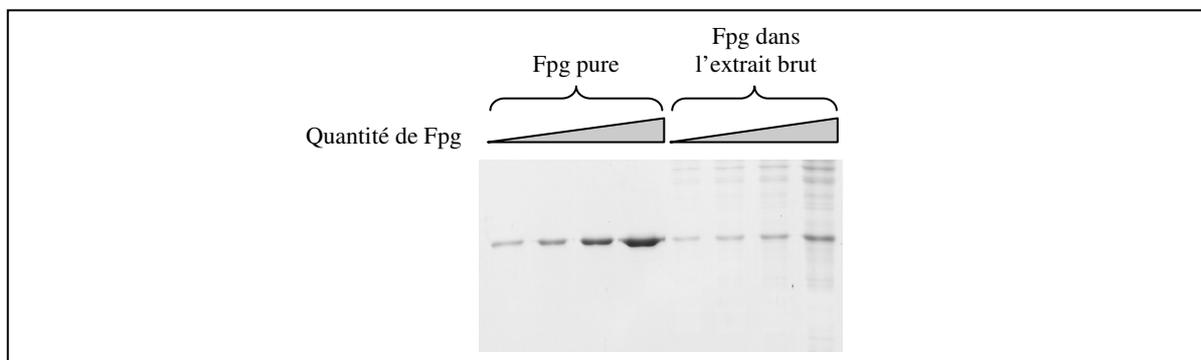
#### **IV.1.3. Lyse des cellules et obtention d'extraits cellulaires**

Après décongélation, chaque culot bactérien est repris dans du tampon de lyse des cellules (voir annexe) à raison de 10 mL par gramme de bactérie. La suspension cellulaire obtenue est complétée par du lysozyme (0,5 mg.mL<sup>-1</sup> final), puis incubée à 30°C pendant 15 min. Après quatre cycles de congélation-décongélation dans l'azote liquide, le lysat cellulaire résultant est centrifugé (1h30, 96 000g, 4°C). Le surnageant résultant constitue la fraction protéique soluble, ou extrait cellulaire brut (EB). Ce protocole a été suivi pour chaque bactérie recombinante surproduisant la protéine *L/Fpg* sous sa forme sauvage ou ses formes mutantes. Après centrifugation, les extraits cellulaires bruts sont fractionnés par 500 µL, congelés rapidement dans l'azote liquide et stockés à -80°C.

#### **IV.1.4. Détermination de la quantité de protéine dans les extraits cellulaires**

Afin de pouvoir comparer plus précisément les propriétés des différentes formes de Fpg obtenues, il est nécessaire de déterminer la concentration en protéines totale dans nos extraits cellulaires, ainsi que la concentration en Fpg. Pour ce faire, deux techniques sont utilisées. La concentration en protéines totale des différents extraits cellulaires est déterminée en utilisant la méthode de dosage colorimétrique de Bradford (Bradford, 1976). Par contre, la détermination de la concentration en protéine Fpg dans l'extrait nécessite une quantification directe sur SDS-PAGE. En l'occurrence, on fait migrer sur un même gel SDS-PAGE (voir

annexe) une gamme de dilution sous-saturante en Fpg de l'extrait et une gamme étalon de protéine *LIFpg* purifiée de concentration connue (un exemple de gel obtenu avec *LIFpg* de type sauvage est donné dans la Figure 95).



**Figure 95 : Gel SDS-PAGE pour la quantification de *LIFpg* WT dans les extraits bruts.**

Le gel est un SDS-PAGE à 12 % (voir annexe). Les quatre premières pistes correspondent respectivement à des quantités de 0,25 ; 0,5 ; 1 et 2 µg de *LIFpg* purifiée. Les pistes 5 à 8 correspondent respectivement à des volumes de 1,875 ; 2,5 ; 3,75 et 7,5 µl d'extrait brut contenant Fpg dilué au 1/25<sup>ème</sup>.

Après avoir coloré au bleu de Coomassie puis scanné le gel, les intensités des bandes sur chaque piste sont mesurées à l'aide du logiciel « Image Quant TL » (Amersham Biosciences). Les concentrations en Fpg sont alors déterminées par comparaison des intensités des bandes correspondant à Fpg dans les extraits à celles obtenues avec la gamme étalon de Fpg pure. Cette méthode étant relativement imprécise, plusieurs quantifications indépendantes ont été réalisées pour chaque forme de Fpg, et les concentrations ont ensuite été moyennées. Les concentrations obtenues sont présentées dans le Tableau 19.

Plasmide utilisé	[Fpg] ( $\mu\text{g}\cdot\mu\text{L}^{-1}$ )	[Protéines totales] ( $\mu\text{g}\cdot\mu\text{L}^{-1}$ )
pFlag-LIFpg E2N	2,53	8,56
pFlag-LIFpg E2R	2,37	9,82
pFlag-LIFpg E2T	3,58	13,33
pFlag-LIFpg E2L	1,84	8,91
pFlag-LIFpg E2N	2,9	9,24
pFlag-LIFpg E2D	1,86	11,73
pFlag-LIFpg E2Q	2,73	12,04
pFlag-LIFpg WT	1,71	12,45
pFlag-CTC	0	13,87

**Tableau 19 : Tableau récapitulatif des quantifications de Fpg et des protéines totales dans les extraits bruts.**

Les versions mutantes de Fpg sont notées E2X avec X le résidu substituant du glutamate 2. Le plasmide pFlag-CTC correspond au plasmide ne contenant pas le gène *fpg*.

Connaissant les concentrations en protéine dans nos extraits, nous avons ensuite complété les extraits contenant le moins de protéines annexes avec l'extrait ne contenant pas Fpg (issu du plasmide pFlag-CTC), ceci dans le but d'avoir le même rapport de concentrations entre Fpg et les protéines totales pour chaque extrait. Enfin, on a légèrement dilué certains extraits afin d'obtenir des concentrations finales en Fpg WT ou mutantes identiques dans chaque extrait.

#### **IV.1.5. Purification de *Ll*Fpg WT et mutantes**

Toutes les étapes de purification sont identiques pour les trois formes de Fpg que l'on a purifiées (WT, E2Q et E2D), et ont été réalisées à 4°C. Au cours de ces étapes, la présence de la protéine *Ll*Fpg est détectée par l'analyse des fractions de purification par électrophorèse en conditions dénaturantes (SDS-PAGE). La purification commence en utilisant un extrait brut réalisé dans des conditions identiques à précédemment, mais en partant d'une culture de 1 L.

##### **IV.1.5.1. Chromatographie Accell QMA (Waters)**

La colonne QMA est un échangeur fort d'anions dont les groupements fonctionnels (ammonium quaternaire), dans des conditions de haute force ionique (250 mM NaCl), vont fixer uniquement les anions fortement chargés. Cette étape de chromatographie permet donc d'éliminer principalement les molécules d'acides nucléiques encore présentes dans l'extrait brut qui pourraient interférer avec les étapes suivantes de purification.

La colonne est utilisée au débit de 2 mL.min<sup>-1</sup>, elle est équilibrée pendant 30 minutes dans du tampon QMA (voir annexe), puis l'extrait brut (~40 mL) est injecté dans la colonne. Dans ces conditions, la plupart des protéines contenues dans l'échantillon, dont l'enzyme *Ll*Fpg, ne se fixent pas à la résine et sont directement éluées. On fait ensuite un léger lavage de la colonne pour finir d'élué les molécules non fixées sur celle-ci. La fraction non fixée sur la colonne constitue la fraction « QMA ».

#### **IV.1.5.2. Chromatographie Sulphopropyl Sepharose Fast Flow (SPFF, Pharmacia)**

La colonne SPFF est un échangeur fort de cations. Cette étape chromatographique permet de fixer les protéines chargées positivement grâce aux groupements sulphopropyls greffés sur la résine, ce qui est le cas de *LIFpg* ( $pI = 9,6$ ) dans le tampon SP (voir annexe) à  $pH = 7,6$ . La colonne est utilisée au débit de  $2 \text{ mL}\cdot\text{min}^{-1}$  et elle est équilibrée pendant 30 minutes dans du tampon SP à 50 mM NaCl. Ensuite, la protéine *LIFpg* est fixée sur la colonne, puis on fait un lavage pendant 30 minutes à 100 mM NaCl pour éliminer les molécules faiblement cationiques. On élue ensuite tout ce qui reste fixé sur la colonne en une seule fois à 500 mM NaCl, ce qui constitue la fraction « SPFF ».

#### **IV.1.5.3. Chromatographie Sulphopropyl Sepharose High Performance (SPHP, Pharmacia)**

Les propriétés de cette colonne sont les mêmes que la colonne utilisées pour l'étape précédente. La seule différence réside dans son plus grand pouvoir de séparation lié à des billes de résine plus petites. Après une équilibration identique à celle de l'étape précédente, la fraction « SPFF » est diluée que si pour 50 mM NaCl avec du tampon SP sans sel (tampon SP de type A, voir annexe) et chargée à  $2 \text{ mL}/\text{min}$  sur la colonne, suivi d'un lavage comme précédemment. Les protéines fixées sur la colonne sont ensuite éluées par un gradient linéaire de concentration en NaCl allant de 100 à 600 mM sur 400 mL, l'élution étant fractionnée par échantillons de 3 mL au moyen d'un collecteur de fractions. Les fractions sont alors analysées par SDS-PAGE (données non présentées) afin de déterminer celles qui contiennent la protéine *Fpg*. En l'occurrence, la protéine *LIFpg* est éluee entre 250-300 mM NaCl, et les fractions correspondantes sont rassemblées, constituant la fraction « SPHP ».

#### **IV.1.5.4. Chromatographie Ultrogel AcA54 (LKB)**

Par cette colonne ( $L = 1\text{m}40$  et  $\varnothing = 2 \text{ cm}$ ), nous séparons les différentes protéines de l'échantillon selon le principe de gel filtration et donc selon leur taille et/ou leur forme. En effet, les molécules sont séparées à l'image d'un tamis moléculaire par leur capacité à diffuser au travers d'un réseau de billes poreuses que constitue la résine AcA54. Pratiquement, seules les molécules de moins de 80 KDa sont séparées sur cette résine, étant éluées d'autant plus

lentement qu'elles sont petites, et les molécules de plus de 80 KDa sont directement éluées dans le volume mort de la colonne.

La colonne est utilisée au débit de  $0,5 \text{ mL}\cdot\text{min}^{-1}$  et elle subit une équilibration pendant 24 heures dans le tampon SP à 1M NaCl (tampon SP de type B, voir annexe). Ensuite, la fraction « SPHP », concentrée par concentrateur (Amicon Ultra 15, Millipore) jusqu'à un volume final de 3 mL, est injectée sur le gel. L'éluion est fractionnée par tubes de 5 mL, et les fractions sont analysées par SDS-PAGE (données non présentées). Les fractions contenant la protéine d'intérêt sont rassemblées et constituent la fraction « AcA54 ».

#### **IV.1.5.5. Chromatographie HS20 (POROS)**

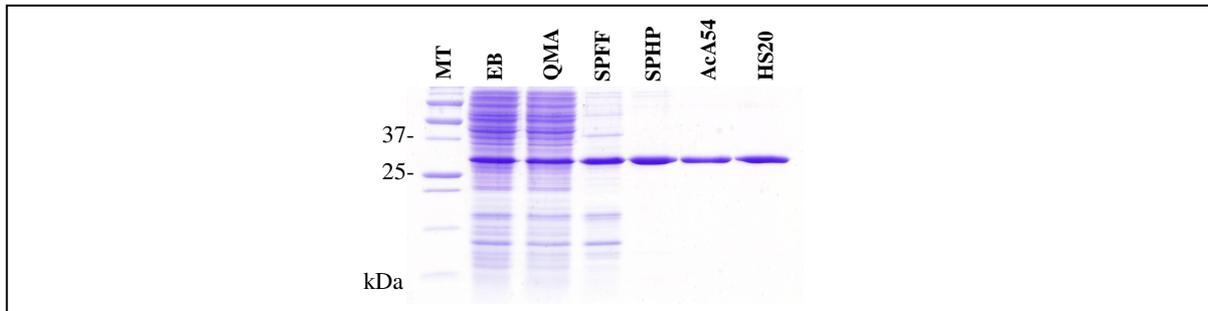
Cette colonne contient un échangeur fort de cations. La structure poreuse de la résine sur laquelle sont greffés des groupements sulphopropyls, ainsi que la forte homogénéité de taille des billes de résine, lui apportent un fort pouvoir séparateur. Dans ce cas précis, elle permet d'éliminer le seul contaminant restant, en l'occurrence un résultat de dégradation de la protéine Fpg retrouvé en sortie de AcA54, dont la taille et la charge sont très proches de celles de Fpg active (ce qui fait qu'on ne les différencie pas sur SDS-PAGE, Figure 96). L'échantillon « AcA54 » est d'abord concentré, permettant aussi de changer de tampon pour se retrouver dans du tampon SP (voir annexe) à 100 mM NaCl. La colonne est utilisée à  $3 \text{ mL}\cdot\text{min}^{-1}$ , et est équilibrée 30 min dans du tampon SP à 100 mM NaCl. Ensuite, l'équivalent de 5 mg de protéine *L/Fpg* est injecté à  $0,5 \text{ mL}\cdot\text{min}^{-1}$ , puis la colonne subit un lavage à  $3 \text{ mL}\cdot\text{min}^{-1}$  pendant 15 minutes avec du tampon SP à 100 mM NaCl. La protéine est enfin éluée à  $3 \text{ mL}\cdot\text{min}^{-1}$  avec un gradient linéaire de concentration en NaCl allant de 100 à 650 mM sur 136 mL. Les fractions éluées entre 300 et 435 mM NaCl contiennent la protéine *L/Fpg* pure, et sont rassemblées pour constituer la fraction « HS20 ».

#### **IV.1.5.6. Récapitulatif du suivi de purification**

La Figure 96 qui suit montre un exemple de gel SDS-PAGE (voir annexe) représentatif du suivi des principales étapes de notre purification protéique, dans ce cas pour la protéine *L/Fpg*-E2D.

On voit sur cette figure que l'on a bien purifié jusqu'à une apparente homogénéité la protéine *L/Fpg*. Par contre, on ne voit plus de contaminant sur le gel à partir de la fraction AcA54. Cependant, quand on passe cette fraction sur la colonne HS20, on voit à l'éluion un

léger pic correspondant à un contaminant, ce qui prouve que l'on avait une inhomogénéité de Fpg qui n'était pas visible sur SDS-PAGE. De plus, on a remarqué que ce pic de contaminant avait tendance à augmenter lentement avec le temps, ce qui prouve qu'il correspond à une dégradation de Fpg n'influant que peu sur les propriétés de la protéine en SDS-PAGE (la perte du zinc du doigt à zinc par exemple).



**Figure 96 : Bilan de la purification de E2D-L/Fpg.**

Le gel est un SDS-PAGE à 15 % (voir annexe). Le nom de la fraction déposée sur chaque piste est indiqué au dessus du gel. Des marqueurs de tailles (MT) ont été déposés sur la première piste (Precision Plus Unstained Standards, Biorad).

#### **IV.1.5.7. Dosage et stockage de L/Fpg WT et mutantes**

Comme indiqué précédemment, la protéine Fpg a légèrement tendance à se dégrader au cours du temps. Sachant que la seule étape chromatographique permettant d'éliminer le contaminant généré est le passage sur colonne HS20, nous avons décidé d'aliqoter la fraction « Aca54 » par tubes d'environ 5 mg de protéine Fpg (déterminé rapidement en utilisant un dosage de Bradford) pour stockage à  $-80^{\circ}\text{C}$ . Ensuite, il ne reste donc plus qu'à décongeler un de ces aliqots et à le passer sur colonne HS20 à chaque fois que l'on débute une série de manipulations nécessitant d'avoir un lot de Fpg bien pure.

Notamment, nous avons remarqué que les études cristallographiques de Fpg nécessitaient d'avoir un lot de Fpg fraîchement purifié sur HS20, l'idéal étant de réaliser cette étape de purification et les essais de cristallogenèse au cours d'une même journée. Dans ce cadre, les fractions d'intérêt en sortie de colonne HS20 sont regroupées et concentrées à l'aide d'un concentrateur (Amicon Ultra 15, Millipore). La protéine est alors dosée par la technique de Bradford puis diluée dans du tampon de stockage de Fpg à 5 % glycérol (voir annexe) pour obtenir la concentration désirée (généralement environ  $160\ \mu\text{M}$ , soit  $5\ \text{mg.mL}^{-1}$ ), et cette solution est directement utilisée pour réaliser nos essais de cristallogenèse.

Par contre, dans le cas des études biochimiques des propriétés de Fpg, il est apparu que la dégradation de celle-ci était suffisamment lente pour permettre de stocker la protéine

en sortie de colonne HS20 durant plusieurs semaines à plusieurs mois à -20°C sans que cela n'influe réellement sur les propriétés de l'enzyme vis-à-vis de ses substrats ou analogues de substrats (les constantes de dissociations apparentes par exemple). Dans ce cadre, dès la sortie de colonne HS20, les fractions contenant la protéine sont regroupées et concentrées à l'aide d'un concentrateur (Amicon Ultra 15, Millipore). La protéine est alors dosée par la technique de Bradford puis diluée dans du tampon de stockage de Fpg à 50 % glycérol (voir annexe) pour obtenir la concentration désirée (généralement 100 µM). L'ensemble est alors conservé à -20°C.

#### **IV.1.6. Source des autres protéines utilisées**

Tout d'abord, les protéines yOgg1 et Nth qui ont été utilisées dans certaines manipulations sont issues du laboratoire du Dr Serge BOITEUX (Laboratoire de radiobiologie de l'ADN, CEA Fontenay-aux-Roses, 18 route du panorama, 92265 Fontenay-aux-roses cedex, France), qui a eu l'amabilité de nous fournir les quantités de protéines nécessaires à nos manipulations.

Ensuite, la protéine HU ( $\alpha_2$  et  $\beta_2$ ) a été produite et purifiée dans notre équipe au laboratoire par le Dr Bertrand CASTAING.

Enfin, les protéines BSA et lysozyme sont des protéines commerciales.

#### **IV.1.7. Production des oligonucléotides utilisés**

Tout d'abord, les séquences des oligonucléotides utilisés pour nos études sont données dans le Tableau 20.

<b><u>Tableau 20 : Oligonucléotides utilisés pour nos études.</u></b>	
X représente la guanine, la 5-OH-5-Me-Hyd, la 8-oxo-G, le FapyG, le Bz-FapyG, ou leur analogue carbocyclique, Y représente la base nucléique opposée à la guanine ou aux lésions (A, G, C ou T).	
14 mere, brin contenant la lésion	5'- CTCTTT <b>X</b> TTTCTCG -3'
14 mere, brin opposé à la lésion	5'- GCGAGAAA <b>Y</b> AAA GA -3'
24 mere, brin contenant la lésion	5'- GCGCCTCTTT <b>X</b> TTTCTCGCCGGG -3'
24 mere, brin opposé à la lésion	5'- CCCGGCGAGAAA <b>Y</b> AAAGAGGCCGC -3'

D'autre part, les oligonucléotides 14 meres contenant une 5-OH-5-Me-Hydantoïne naturelle ou greffée sur un cyclopentane, ainsi que ceux contenant un FapyG ou un Benzyl-

FapyG greffés sur un cyclopentane, nous ont été fournis par le laboratoire d'un de nos collaborateurs chimiste, le Dr Thomas CARELL (Center for Integrated Protein Science, Department of Chemistry and Biochemistry, Ludwig-Maximilians Universität München, Feodor-Lynen straÙe 25, D-81377 Munich, Germany).

De même, l'oligonucléotide 14 mere contenant une 8-oxo-G greffée sur un cyclopentane nous ont été fournis par le laboratoire d'un de nos collaborateurs chimiste, le Dr Didier GASPARUTTO (Laboratoire des Lésions des Acides Nucléiques, CEA de Grenoble, 38054 Grenoble Cedex 9, France).

Enfin, tous les autres oligonucléotides utilisés ont été commandés chez la société Eurogentec.

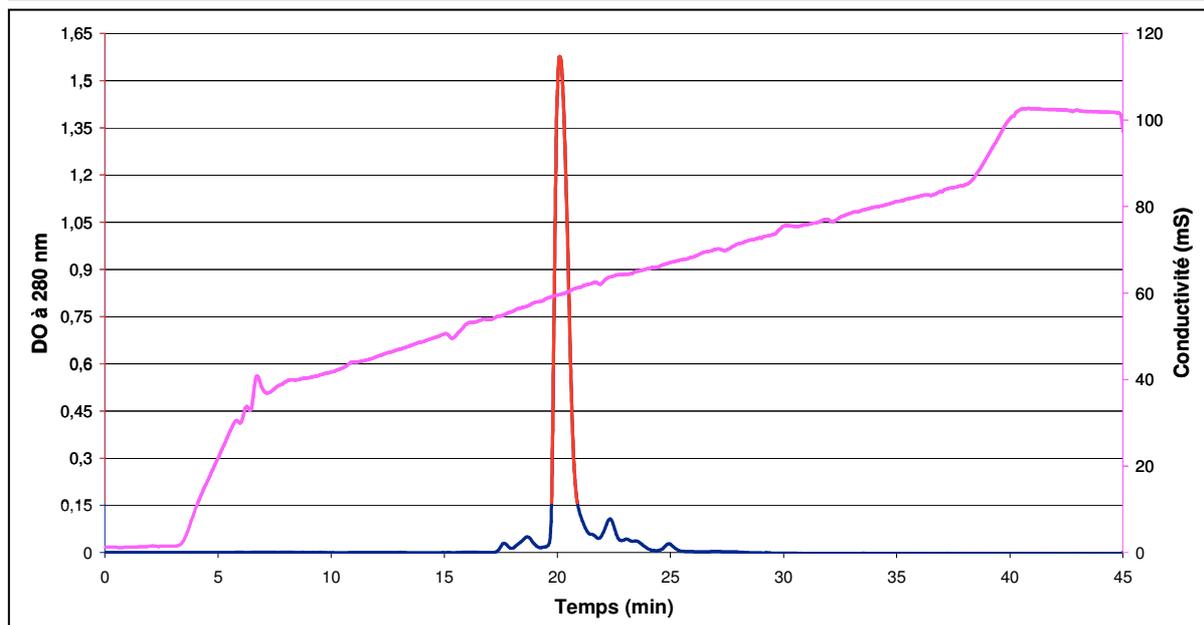
#### **IV.1.8. Purification des oligonucléotides utilisés**

Les oligonucléotides utilisés lors de nos études n'étant généralement pas assez purs pour nos études cristallographiques (souvent une déprotection incomplète), nous les avons purifiés sur colonne échangeuse d'anions de type MonoQ 10/100 GL (Amersham Biosciences) en conditions dénaturantes. Pour ce faire, la colonne est utilisée à 3 mL.min<sup>-1</sup>, et est équilibrée 30 min dans du tampon MonoQ de type A (voir annexe). Un oligonucléotide est ensuite chargé sur la colonne dans du tampon MonoQ de type A, puis celle-ci subit un lavage pendant 15 minutes avec du tampon MonoQ de type A. Enfin, l'oligonucléotide est élué par un gradient allant de 0 à 30 % de tampon MonoQ de type B en 5 minutes, puis de 30 à 80 % de tampon MonoQ de type B en 30 minutes. L'oligonucléotide pur est alors collecté manuellement dès que l'on observe une forte augmentation de DO à 280 nm (après environ 25 min d'éluion pour les oligonucléotides 14 meres contenant une lésion, et après environ 20 min d'éluion pour les oligonucléotides 14 meres complémentaires sans lésion).

Un exemple de chromatogramme obtenu lors d'une telle purification est donné dans la Figure 97, correspondant à la purification d'un oligonucléotide complémentaire sans lésion 14-mere-C. Afin de s'assurer de ne pas collecter de molécules contaminantes, on ne collecte que la partie du pic étant au dessus de 0,15 DO<sub>280 nm</sub> (Figure 97, en rouge). On peut noter ici que l'on suit l'éluion grâce à la DO à 280 nm et non pas à 254 nm, car cela permet de ne pas saturer le détecteur, et donc de pouvoir purifier une plus grande quantité d'oligonucléotide à chaque passage sur la colonne. Les tubes collectés (soit environ 3,5 mL) sont alors stockés à -20°C, et on lave la colonne avec 100 % de tampon MonoQ de type B pendant 10 min pour éliminer toute molécule fixée.

**Figure 97 : Chromatogramme représentatif de la purification d'un oligonucléotide 14-merc-C.**

La conductivité représentant le suivi du gradient de concentration en sel est représenté en rose. La DO à 280 nm est représentée en bleu et en rouge, la partie bleue correspondant à ce que l'on ne garde pas, et la partie rouge correspondant à la fraction purifiée que l'on récupère.



A ce niveau, on peut indiquer que les oligonucléotides contenant une 8-oxo-G greffée sur un désoxyribose n'ont pas pu être purifiés par cette chromatographie, car cette lésion est apparue comme étant dégradée par la soude contenue dans le tampon MonoQ. De plus, nos tests pour purifier cette lésion sur MonoQ en conditions natives ou par HPLC n'ont pas été concluants à cause d'une résolution trop faible. En fait, pour ce cas particulier, nous n'avons fait qu'une purification sur gel non dénaturant à 20 % en acrylamide/bisacrylamide 19/1, ce qui nous a tout de même permis d'éliminer le contaminant majeur contenu dans notre lot d'oligonucléotide issu de Eurogentec.

Ensuite, il a été nécessaire de rétablir un pH neutre, puis de dessaler nos oligonucléotides, ce qui a été réalisé grâce à des petites colonnes hydrophobes de type « Sep-Pak Cartridge C18 » (Waters). Cette étape nous a permis d'obtenir des oligonucléotides en solution dans de l'eau ultrapure, avec des quantités de sels et de tampon négligeables. Les oligonucléotides ont alors été stockés à -20°C après dosage.

#### **IV.1.9. Marquage au $^{32}\text{P}$ d'oligonucléotides**

Ce type de marquage est systématiquement réalisé pour nos manipulations de biochimie, que ce soit avec les oligonucléotides 14 ou 24 meres. Afin de faciliter la quantification de l'intensité des bandes après migration sur gel, seul le brin contenant la lésion est marqué radioactivement.

Pour chaque type de duplexe d'ADN désiré, on réalise le mélange suivant :

- 10 pmol de brin contenant la lésion (sous forme sèche).
- 1  $\mu\text{L}$  de PolyNucléotide Kinase à  $10 \text{ u.}\mu\text{L}^{-1}$  (PNK, Fermentas).
- 1  $\mu\text{L}$  de tampon de PNK 10X.
- 8  $\mu\text{L}$  de  $[\gamma\text{-}^{32}\text{P}]\text{-ATP}$  à  $3,33 \text{ pmol.}\mu\text{L}^{-1}$  et  $10 \text{ }\mu\text{Ci.}\mu\text{L}^{-1}$  (Amersham Biosciences).

On incube ce mélange 1 h à  $37^\circ\text{C}$ , on ajoute à nouveau 1  $\mu\text{L}$  de PNK, puis on incube 1 h à  $37^\circ\text{C}$ . Le mélange réactionnel est alors séché à l'évaporateur rotatif, puis dilué dans 6  $\mu\text{L}$  de tampon de dénaturation Formamide 1X (voir annexe), et incubé 2 à 3 minutes à  $70^\circ\text{C}$ .

On peut noter ici que le tampon de dénaturation Formamide utilisé ici ne contient pas de soude (car aucune activité AP/lyase n'est souhaitée) et ne contient pas de Bleu de Bromophénol (car celui-ci migrerait sur gel d'acrylamide/bisacrylamide 19/1 à 20 % comme l'oligonucléotide de 14 bases, compliquant ainsi la purification).

L'échantillon obtenu est alors déposé sur minigel (dimensions  $83\times 73\times 0,75 \text{ mm}$ , Mini-protean 2, Biorad) d'acrylamide/ bisacrylamide 19/1 à 20 %, 1X TBE, puis on fait une migration à  $200 \text{ V}^{\text{cst}}$  pendant une heure afin d'éliminer l'ATP en excès.

Le gel obtenu est révélé par autoradiographie, la bande d'intérêt est découpée puis déposée dans 1 mL d'eau ultrapure, et l'élution est réalisée par agitation pendant deux heures à  $50^\circ\text{C}$ , puis une nuit à température ambiante.

La solution obtenue est enfin filtrée à  $0,2 \text{ }\mu\text{m}$  sur Eppendorf, puis séchée à l'évaporateur rotatif.

#### **IV.1.10. Hybridation d'oligonucléotides**

On hybride le brin marqué avec un excès en brin complémentaire, et ce afin de s'assurer que la totalité du brin marqué soit hybridée à son complémentaire.

A ce niveau, les duplexes de 24 paires de bases étant relativement stables à température ambiante, ceux-ci sont hybridés avec un excès de 10 fois en brin complémentaire par rapport au brin lésé.

D'autre part, dans les conditions diluées de biochimie, le faible  $T_m$  des duplexes de 14 paires de bases fait que l'on hybride le brin contenant la lésion avec un excès de 80 fois en brin complémentaire, ce qui permet de favoriser l'hybridation. A l'opposé, dans les conditions concentrées de cristallographie, l'hybridation est plus efficace, et ne nécessite donc qu'un excès de 1,1 fois en brin complémentaire.

A ce niveau, les oligonucléotides complémentaires dissouts dans de l'eau ultrapure sont mis en présence, puis le mélange est séché à l'évaporateur rotatif. Le culot obtenu est alors repris dans 50  $\mu\text{L}$  de tampon d'hybridation de l'ADN n°1 (voir annexe) pour les expériences de biochimie. Par contre, pour les expériences de cristallographie, le culot est repris dans du tampon d'hybridation de l'ADN n°2 (voir annexe) afin d'avoir une concentration finale en duplexe de 2,64 mM.

Enfin, on procède à l'hybridation des duplexes par dépôt de l'échantillon dans un bain marie à 100°C, dont on fait ensuite descendre la température le plus lentement possible jusqu'à 4°C. Les duplexes sont alors stockés à -20°C.

On peut noter ici que l'on a au maximum 50 % de rendement lors de la purification sur gel du brin marqué, soit environ 5 pmol après cette étape. On considèrera donc que l'on a 5 pmol de brin contenant une lésion à hybrider. De plus, on a en moyenne (selon la « fraîcheur » de notre stock d'ATP radioactif) 500 cps. $\mu\text{L}^{-1}$  dans les 50  $\mu\text{L}$  obtenus après hybridation. On pourra donc considérer que nos duplexes marqués présentent une radioactivité d'environ 5000 cps.pmol<sup>-1</sup>.

## **IV.2. Etudes structurales de complexes Fpg/ADN contenant une lésion**

### **IV.2.1. Cristallogenèse sur les complexes Fpg/ADN contenant une lésion**

Premièrement, on peut dire que la cristallogenèse a été réalisée grâce au protocole définis en 2002 dans l'équipe (Pereira de Jesus *et al.*, 2002).

Les complexes sont formés dans les conditions suivantes :

- 20  $\mu\text{l}$  de protéine Fpg stock à 161,3  $\mu\text{M}$ , soit environ 147,8  $\mu\text{M}$  final en Fpg.
- 1,83  $\mu\text{L}$  de stock de duplexe d'ADN à 2,64 mM, soit environ 221,3  $\mu\text{M}$  final en duplexe d'ADN (soit un excès de 1,5 fois en ADN par rapport à la protéine).

En fait, les concentrations indiquées ci-dessus correspondent aux conditions de base utilisées, et ont été légèrement ajustées selon les essais de cristallisation. Notamment, pour les essais de cristallisation à 20°C, les solutions précédentes ont été généralement un peu diluées pour des raisons de solubilité inférieure des complexes à 20°C par rapport à 37°C.

A ce niveau, on réalise nos gouttes de cristallisation dans les conditions suivantes :

- 1,5  $\mu\text{L}$  de solution de cristallisation.
- 1,5  $\mu\text{L}$  de solution de complexe.

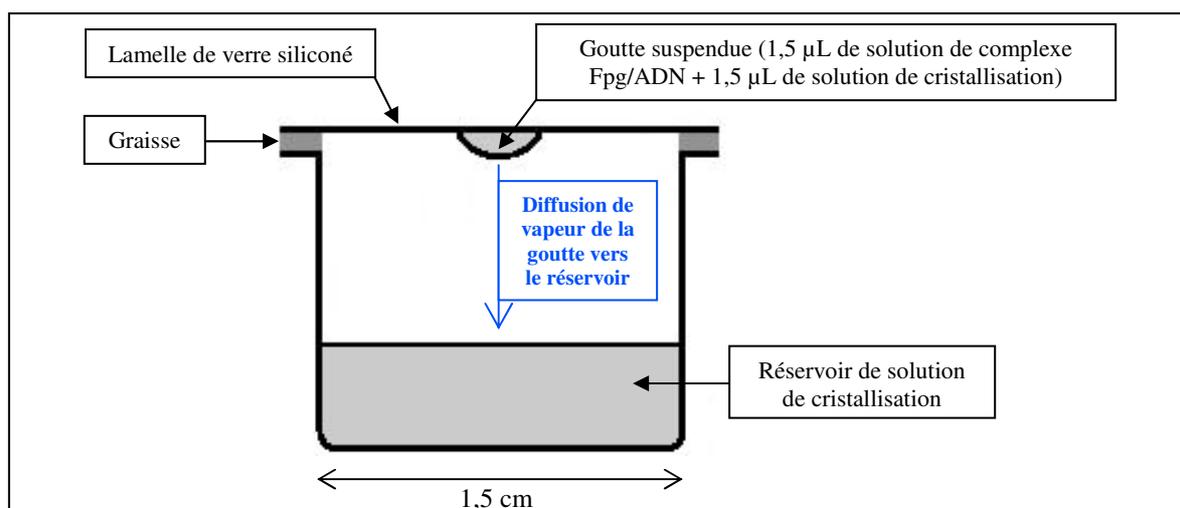
On peut noter que la solution de cristallisation n'est pas exactement identique pour tous nos complexes, et cela à nouveau à cause de la différence de solubilité de nos complexes à 20 et 37°C. En effet, la solubilité supérieure à 37°C nous a obligés à utiliser une solution de cristallisation plus concentrée en agent cristallisant qu'à 20°C. Finalement, pour les complexes non-covalents formés entre Fpg et les duplexes cFapyG/C, Bz-cFapyG/C et cHyd/C, la solution de cristallisation est la suivante :

- 1,4 M citrate, 0,1M Hepes/NaOH, pH 7,6.

Par contre, pour le complexe covalent formé entre Fpg et le duplexe cHyd/C, la solution de cristallisation est la suivante :

- 1,6 M citrate, 0,1M Hepes/NaOH, pH 7,6.

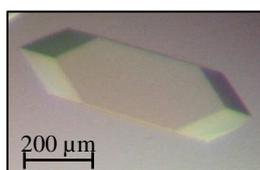
La cristallisation est alors faite par la méthode de diffusion de vapeur en goutte suspendue contre un réservoir de 1 mL de solution de cristallisation (Figure 98), et est réalisée à T° ambiante pour les complexes non-covalents Fpg/(cFapyG/C), Fpg/(Bz-cFapyG/C) et Fpg/(cHyd/C), et à 37°C pour le complexe covalent Fpg/(cHyd/C).



**Figure 98 : Cristallogénèse par la méthode de diffusion de vapeur en goutte suspendue.**

La concentration en agent cristallisant est deux fois moindre dans la goutte que dans le réservoir, ce qui implique une diffusion de vapeur d'eau de la goutte vers le réservoir.

Les conditions de cristallisation précédentes ont permis d'obtenir des cristaux de taille suffisante pour la cristallographie en quelques jours, soit environ 200 à 600 µm sur la plus grande longueur. Un exemple du type de cristal obtenu est donné dans la Figure 99.



**Figure 99 : Cristal obtenu avec le complexe Fpg/(cHyd/C) à 37°C.**

#### **IV.2.2. Vérification de la qualité des cristaux sur générateur de rayons X à anode tournante**

Dans le but d'amener au synchrotron seulement les cristaux ayant le meilleur pouvoir diffractant, une étude préliminaire de diffraction des rayons X est réalisée au Centre de Biophysique Moléculaire sur un générateur de rayons X à anode tournante ( $\lambda=1,54 \text{ \AA}$ , raie  $K\alpha$  du cuivre) de type « Rigaku RU-200 » fonctionnant à 40 kV et 100 mA.

Les cristaux sont prélevés sous loupe binoculaire dans leur solution de cristallisation grâce à une boucle de nylon de taille adaptée (Hampton Research), puis ils sont rapidement montés sur une tête goniométrique et figés à 100 K grâce au système « Oxford Cryosystems 600 Series ». La forte concentration en citrate dans les cristaux (1,4 ou 1,6 M) nous a permis de nous passer de l'ajout d'un cryoprotectant supplémentaire.

Les clichés de diffraction sont collectés grâce à un détecteur 2D à plaque photosensible « MAR 300 » (MarResearch). Ces clichés sont tout d'abord analysés visuellement puis traités par le programme « MOSFLM » pour vérifier que les cristaux que nous souhaitons emmener au synchrotron sont bien monocristallins, et qu'ils peuvent diffracter à haute résolution. Les cristaux ayant un pouvoir diffractant suffisant ( $< 3 \text{ \AA}$  de résolution) sont retirés rapidement de la tête du goniomètre et stockés dans un dewar de transport rempli d'azote liquide.

#### **IV.2.3. Collection et traitement des données de diffraction, affinement des modèles moléculaires et vérification de leur qualité**

##### **IV.2.3.1. Caractéristiques des lignes de lumière utilisées à l'ESRF**

Nous avons utilisé les lignes de lumière « ID23-1 » et « BM30A » de l'Installation Européenne de Rayonnement Synchrotron (ERSF, Grenoble, France), qui sont dédiées aux expériences de cristallographie sur les macromolécules. Ces lignes de lumière possèdent toutes les deux un faisceau de rayons X monochromatiques de longueurs d'ondes respectivement fixées à 0,97935 et 0,92004  $\text{\AA}$  (soit respectivement des énergies de 12,66 et 13,48 keV), et permettent de collecter des données de très haute résolution (allant idéalement jusqu'à environ 0,9  $\text{\AA}$ ). La taille du faisceau a été réglée à 100  $\mu\text{m} \times 100 \mu\text{m}$  lors de nos expériences. En plus du faisceau de rayons X, plusieurs dispositifs sont nécessaires à la collection des données de diffraction à l'ESRF :

- Un Goniomètre, sur lequel nous avons monté nos échantillons. On peut noter qu'afin d'éviter les problèmes d'évaporation de la solution de cristallisation, les échantillons cristallins sont maintenus sous réfrigération à l'azote liquide pendant le transfert depuis le dewar de transport jusqu'à la tête du goniomètre.
- Un système de refroidissement de type « Oxford Cryostream 700 series » fournissant un jet d'azote réfrigéré à 100 K sur l'échantillon, couplé à un système superdissécatrice permettant d'éviter autant que possible la formation de glace sur les cristaux.
- Un détecteur CCD de type « ADSC Q4 ».
- Un système de centrage automatique du cristal dans le faisceau de rayons X.

#### **IV.2.3.2. Collection et traitement des données de diffraction**

##### **- Collection des données et indexage des images**

La collection des données est dirigée dans la cabine de contrôle en utilisant l'interface « ProDC », qui contrôle l'axe phi, les obturateurs, le positionnement du détecteur, ainsi que les autres paramètres expérimentaux. Le traitement des images de diffraction est réalisé grâce au programme « MOSFLM » (Leslie, 1992). La première étape d'un traitement de données est l'autoindexage des images, soit la détermination de la position de chaque réflexion dans le réseau réciproque. Le programme enregistre alors la matrice d'orientation, la distance cristal-détecteur et la longueur d'onde de radiation, et estime la taille maximale de la maille grâce à la séparation inter-réflexions. L'utilisateur est alors amené à sélectionner une maille et un groupe d'espace parmi plusieurs choix. « MOSFLM » affine ensuite la solution sélectionnée et applique la bonne symétrie, puis réalise un post-affinement sur deux cadrans de données adjacentes afin de déterminer les paramètres de maille de façon plus précise. « MOSFLM » est alors en mesure d'intégrer l'intensité des réflexions sur chaque image, et fournit un fichier de sortie (.MTZ) contenant la position dans le réseau réciproque et l'intensité de chacune des réflexions.

##### **- Mise à l'échelle, fusionnement des données et génération des normes des facteurs de structure**

Cette étape est réalisée à partir du fichier de sortie (.MTZ) de « MOSFLM » grâce au programme « SCALA » (Diederichs *et al.*, 1997) de la suite « CCP4 ». Ce programme permet

de mettre toutes les réflexions à l'échelle par comparaison des intensités de réflexions localisées sur des images adjacentes. De plus, ce programme permet de fusionner les données après la mise à l'échelle, ce qui correspond au calcul de la moyenne des intensités pour chaque jeu de réflexions liées par symétrie.

L'analyse des résultats de la mise à l'échelle permet de déterminer :

- Le facteur R de corrélation entre les réflexions liées par symétrie ( $R_{\text{sym}}$ ), qui est un indicateur de la qualité des données.
- La résolution utile du jeu de données par analyse de la variation du rapport signal sur bruit en fonction de la résolution (on estime que  $I/\sigma I$  doit être  $\geq$  à trois).
- Le groupe d'espace effectif du cristal (par fusionnement selon différents groupes d'espaces et analyse de l'évolution de  $R_{\text{sym}}$ , ainsi que par l'analyse des absences systématiques).

La génération des normes des facteurs de structure à partir des intensités de diffraction est alors réalisée grâce au programme « TRUNCATE » (French *et al.*, 1978) de « CCP4 ». Le programme sélectionne alors 5 % des données au hasard, afin de les utiliser pour le calcul du facteur  $R_{\text{free}}$ , qui permettra de suivre la qualité de l'affinement de notre structure.

### **- Résolution du problème des phases et obtention des facteurs de structure**

La méthode de remplacement moléculaire a été choisie pour déterminer les phases car nous possédons plusieurs modèles moléculaires de *LIFpg* complexée à des ADNs lésés. Cette étape est réalisée grâce au programme « AMoRe » (Navaza, 1994) de « CCP4 », en utilisant la structure de *LIFpg* de type sauvage complexée à un ADN contenant un FapyG greffé sur un cyclopentane (code PDB 1XC8).

Le remplacement moléculaire correspond à l'utilisation des phases calculées à partir d'un modèle moléculaire (fichier PDB) proche de celui que l'on étudie pour les appliquer aux normes des facteurs de structure obtenus lors de notre expérience de diffraction des rayons X, et ainsi résoudre le problème des phases. En pratique, cette méthode consiste à déterminer la matrice d'orientation [R] (*via* la corrélation des fonctions de Patterson) puis le vecteur de translation T qui permettent le passage du jeu de coordonnées initiales du modèle ( $\mathbf{X}_1$ ) au jeu de coordonnées de la molécule étudiée ( $\mathbf{X}_2$ ) grâce à la formule suivante :

$$\mathbf{X}_2 = [\mathbf{R}].\mathbf{X}_1 + \mathbf{T}$$

Enfin, le programme réalise un affinement en corps rigide de la position du modèle. Plusieurs solutions sont déterminées par « AMoRe, » et classées suivant la corrélation sur les facteurs de structure et suivant le facteur R. Le jeu de coordonnées choisi permettra alors de pouvoir associer les phases du modèle initial aux modules des facteurs de structure de la molécule étudiée, permettant ainsi le calcul de cartes de densité électronique pour la molécule étudiée. La génération des facteurs de structure calculés est réalisée grâce au programme « SFALL » (Agarwal, 1978) de « CCP4 » par transformée de Fourier inverse.

### **- Génération des cartes de densité électronique et positionnement du modèle atomique**

Cette étape est réalisée grâce au programme « FFT » (Read *et al.*, 1988) de « CCP4 ». Dans notre cas, deux cartes de type SIGMAA ont été générées, soit une carte de type ( $mF_o-DF_c$ ) et une carte de type ( $2mF_o-DF_c$ ). Les cartes de densité électronique sont ensuite visualisées grâce au programme « TURBO-FRODO » (Roussel *et al.*, 1991).

### **- Affinement des modèles moléculaires**

L'affinement de la plus grande partie de chaque modèle est réalisé comme suit. Après avoir ajusté manuellement le modèle atomique dans la carte de densité électronique, on réalise une étape d'affinement du modèle grâce au programme « REFMAC » (Murshudov *et al.*, 1997) de « CCP4 ». Nous avons utilisé un affinement restreint, qui correspond à un ajustement à la fois sur les facteurs de structure observés et calculés, ainsi que sur la géométrie de la molécule. Cet affinement a été paramétré afin de réaliser dix cycles d'optimisation du modèle atomique envers les données expérimentales, et afin d'affiner les facteurs de température de façon mixte (un affinement anisotrope de ce paramètre est préférable pour l'atome de zinc). Le programme génère ensuite de nouvelles cartes de densité électronique, que l'on visualise avec « TURBO-FRODO », et qu'on utilise pour continuer à ajuster manuellement le modèle. On réalise ainsi des cycles d'affinement suivis de modifications du modèle atomique, ceci jusqu'à ce que ce modèle s'inscrive au mieux dans les cartes de densité électronique. Pour suivre ces cycles d'affinement, le programme « REFMAC » estime la corrélation entre le modèle atomique et les données expérimentales *via* les facteurs de structure observés et calculés, donnant un facteur R calculé sur toutes les données utilisées pour l'affinement ( $R_{work}$ ), ainsi qu'un facteur R libre calculé sur une partie

des données exclues de l'affinement ( $R_{\text{free}}$ ), et permettant de vérifier la qualité de l'affinement.

Ensuite les molécules de solvant sont ajoutées aux modèles grâce au programme « wARP » (Asselt Van *et al.*, 1998) de « CCP4 ». Celui-ci permet le positionnement automatique des atomes d'oxygène de l'eau en fonction de la carte  $mF_o-DF_c$  contournée à  $3\sigma$  et de la géométrie des liaisons hydrogène qui s'établissent entre les molécules d'eau et le complexe protéine-ADN.

Enfin, pour la reconstruction des parties de chaque modèle pour lesquelles la carte de densité électronique est trop imprécise, on réalise un affinement par carte d'omission. Cette étape est réalisée afin d'éliminer les biais potentiels provenant de la différence entre la structure à étudier et la structure utilisée pour le remplacement moléculaire. Tout d'abord, on supprime manuellement les atomes du fichier PDB correspondant à la partie du modèle que l'on n'arrive pas à reconstruire correctement. Ensuite, on applique une perturbation de 0,3 Å aux coordonnées x, y et z des atomes restants grâce au programme « Moleman2 » (Kleywegt *et al.*, 2004). Enfin, on réalise quelques cycles d'affinement avec « REFMAC ». Les nouvelles phases obtenues se trouvent ainsi débarrassées des informations fournies par la partie du modèle posant des problèmes d'interprétation.

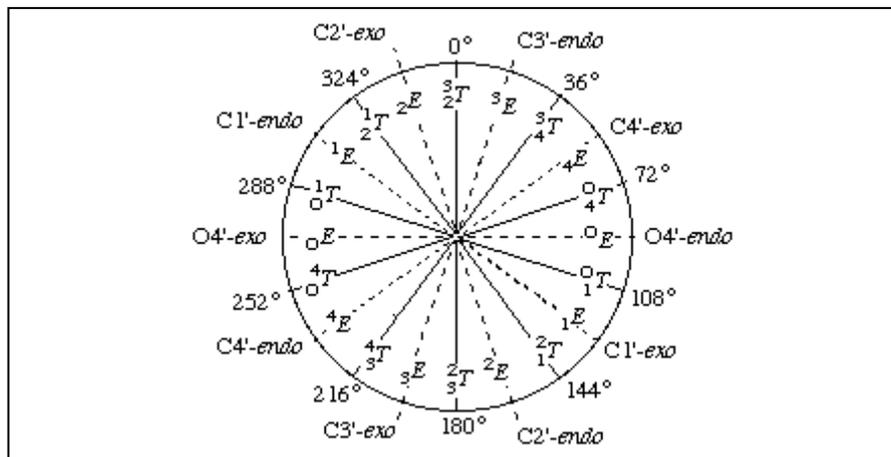
#### **IV.2.4. Vérification de la qualité et analyse des modèles moléculaires**

##### **IV.2.4.1. Vérification de la qualité stéréochimique des modèles protéiques**

Cette étape est réalisée grâce au programme « PROCHECK » (Laskowski *et al.*, 1993) de « CCP4 », qui permet d'analyser la qualité stéréochimique d'un modèle protéique à partir d'un fichier PDB, ceci par comparaison avec des banques de données de paramètres stéréochimiques issus de structures bien définies à haute résolution (angles de torsion, longueur des liaisons...). Les diagrammes de Ramachandran présentés ont été réalisés grâce au programme RP version 2.0 (Gopalakrishnan *et al.*, 2007).



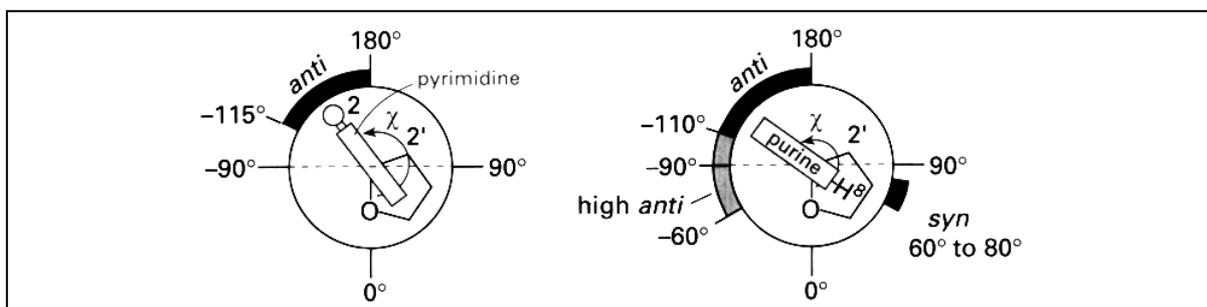
La valeur de P en degrés nous donne alors le type de plissement du sucre grâce au cercle de pseudorotation (Figure 101).



**Figure 101 :** Cercle de pseudorotation du désoxyribose dans l'ADN. (Saenger, 1984).

Dans nos études structurales, le formalisme précédent sera en fait utilisé pour définir le plissement du cyclopentane qui remplace le désoxyribose dans nos analogues carbocycliques de substrats de Fpg.

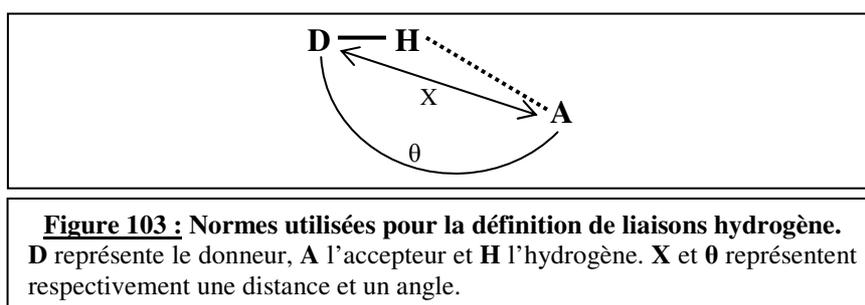
D'autre part, la conformation de la base nucléique par rapport au désoxyribose est définie comme indiqué dans la Figure 102 grâce à la valeur de l'angle de torsion  $\chi$  (Figure 100) (Blackburn *et al.*, 1996) :



**Figure 102 :** Formalisme utilisé pour la définition de la conformation d'une base nucléique par rapport au désoxyribose dans l'ADN.

## - Détermination des réseaux de liaisons hydrogène

Les réseaux de liaisons hydrogène dans nos modèles structuraux ont été déterminés grâce au programme « HBAT » (Tiwari *et al.*, 2007). Pour ce faire, le programme a été utilisé avec un seuil de distance maximum entre les atomes donneur et accepteur de liaison hydrogène de 3,2 Å ( $X$  dans la Figure 103), ce qui permet de ne pas considérer les liaisons hydrogène de faible intensité, ainsi qu'un angle minimum de 90° entre le donneur, son hydrogène, et l'accepteur ( $\theta$  dans la Figure 103), ce qui permet d'éliminer les liaisons impossibles géométriquement.



Par contre, on peut noter que le programme « HBAT » nécessite d'avoir les hydrogènes dans les structures, cet ajout a donc été fait grâce au programme « PyMol ». Or « PyMol » ajoute les hydrogènes à une seule orientation même quand plusieurs positions sont possibles (ex : les molécules d'eau). Dans ce cas, certaines liaisons hydrogène ne sont pas données par le programme « HBAT ». Une vérification manuelle est donc nécessaire pour ne pas passer à côté des certaines liaisons. De plus, le seuil de distance utilisé est relativement restreint. En effet, les liaisons hydrogène sont généralement considérées comme possibles jusqu'à 4 Å, cependant leur force diminuant très rapidement avec la distance, nous avons pris un seuil de 3,2 Å qui permet de ne pas considérer que les liaisons hydrogène de faible intensité.

### **IV.3. Etudes biochimiques de complexes Fpg/ADN contenant une lésion**

#### **IV.3.1. Etudes biochimiques dans des extraits bruts**

##### **IV.3.1.1. Etude de la capacité de fixation de Fpg sur l'ADN lésé dans des extraits bruts**

Lors de ce type d'expériences de gel-retard, la taille du complexe protéine/ADN étant plus élevée que celle de l'ADN seul, la migration du complexe sur gel est retardée par rapport à celle de l'ADN libre, et on voit apparaître deux bandes sur les gels. De plus, le signal détecté au phospho-imageur est proportionnel à la radioactivité présente dans le gel, qui est elle-même proportionnelle à la quantité de duplexe radiomarqué. Sachant cela, la quantification par le logiciel « ImageQuant TL » permet de déterminer la quantité relative de complexe par rapport à l'ADN libre. Dans le cas des ces manipulations sur des extraits cellulaires, la détermination précise des quantités relatives de complexe et d'ADN libre n'est pas primordiale étant donné que la quantification de la quantité de protéine dans les extraits n'est pas très précise, empêchant donc la détermination de constantes de dissociations. Par contre, cette détermination sera très utile par la suite dans nos manipulations de gel-retard sur les protéines purifiées, car dans ce cas on pourra déterminer des constantes de dissociations à partir des quantifications de l'intensité des bandes sur les gels.

Les conditions présentées ici correspondent aux manipulations de gel-retards réalisées avec les Fpgs de type sauvage et mutantes (E2N, E2R, E2T, E2L, E2M, E2D et E2Q) dans des extraits bruts sur les duplexes de 14 paires de bases G/C et c8-oxo-G/C. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Tout d'abord, on réalise une dilution de chaque extrait brut à partir du stock dans du tampon de dilution de Fpg (voir annexe), afin d'obtenir une concentration finale en protéine de 128 nM. De même, on réalise une dilution des différents duplexes dans du tampon TE à 0,4 M NaCl afin d'obtenir des solutions à 10 cps.μL<sup>-1</sup> (soit environ 2 nM en duplexe).

Pour chaque type de duplexe, on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- N×2 μL de duplexe à 10 cps.μL<sup>-1</sup>.

- N×2 µL de tampon d'incubation 4X n°3 (voir annexe).
- N×2 µL d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 6 µL de prémélange + 2 µL d'extrait contenant Fpg WT ou mutantes ou 2 µL d'extrait ne contenant pas de Fpg.

La composition finale de nos essais est donc :

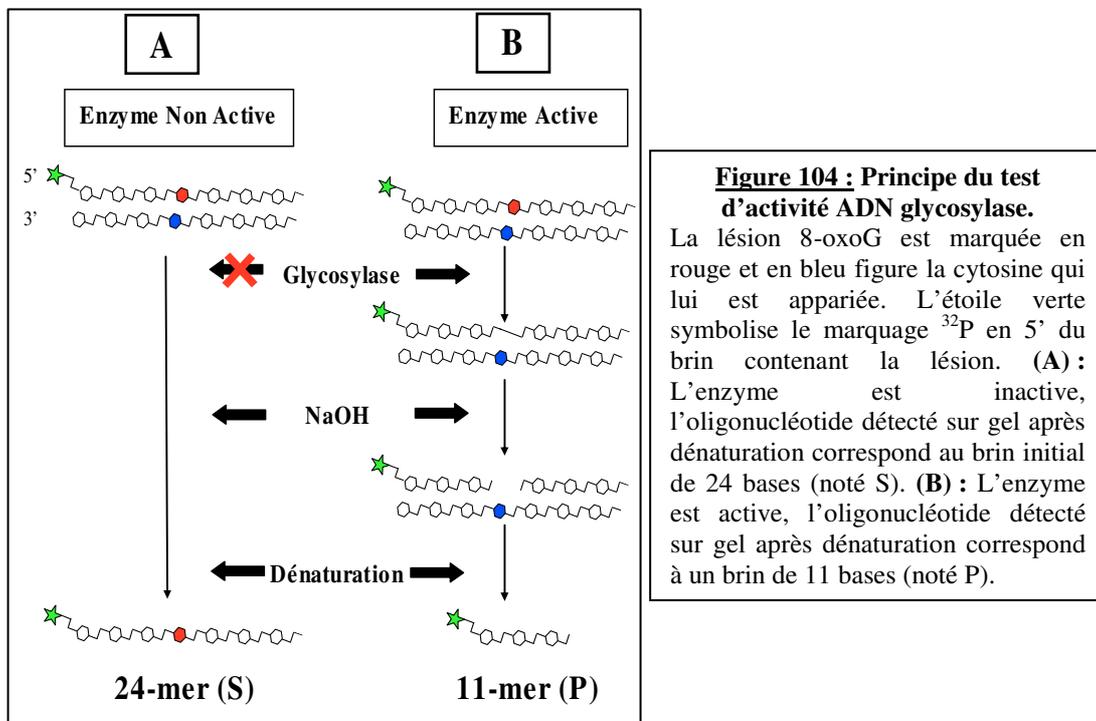
- Environ 32 nM de Fpg WT ou mutantes (sauf pour l'échantillon ne contenant pas de Fpg).
- 0,5 nM en duplexe d'ADN radiomarqué.
- 25 mM HEPES/NaOH (pH 7,6).
- 2,5 mM Tris/HCl (pH 7,6).
- 125 mM NaCl.
- 5,25 mM EDTA.
- 5 mM β-mercaptoéthanol.
- 0,1 mM PMSF.
- 8,75 % glycérol.
- 1X BSA (100 µg.mL<sup>-1</sup>).

On laisse alors la formation des complexes arriver jusqu'à l'équilibre pendant 30 minutes, on charge 5 µL de chaque essai sur minigel (gel-retard de type n°1, voir annexe) (dimensions 200×200×1,5 mm), et on fait migrer à 200 V constant pendant 210 minutes. On procède alors à la fixation des bandes sur les gels dans une solution à 10 % d'acide acétique et 10 % d'éthanol, on sèche ceux-ci, et on met à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL ».

#### **IV.3.1.2. Etude de l'activité ADN glycosylase de Fpg sur l'ADN lésé dans des extraits bruts**

Tout d'abord, on peut rappeler ici que seul le brin contenant la lésion est marqué en 5' par du <sup>32</sup>P. Dans ce cas, après migration sur gel dénaturant l'ADN et révélation, seul ce brin sera détecté. De plus, dans le cas où la protéine Fpg a réalisé son activité ADN glycosylase sur la 8-oxo-G, alors on aura apparition d'un site AP sur ce même brin. Fpg possédant aussi une activité AP/lyase, le produit final de réaction de Fpg sera en fait un brin contenant une

lacune d'un nucléoside. Cependant, afin de s'assurer que tous les sites AP générés par Fpg soient clivés (dans le cas où l'activité AP lyase de Fpg ne serait pas complète, ce qui n'est normalement pas le cas), nous avons ajouté de la soude à notre tampon de dénaturation Formamide (voir annexe), ce qui permet de cliver les sites AP en générant une lacune d'un nucléoside. Dans ces conditions, nous pouvons donc suivre uniquement l'activité ADN glycosylase de Fpg sans avoir à se soucier de l'efficacité de l'activité AP lyase de celle-ci. En effet, après dénaturation de l'ADN, le résultat de la migration sur gel dénaturant montrera deux bandes, l'une correspondant au simple-brin non clivé portant la lésion (bande ayant le moins migré), et l'autre correspondant au produit final clivé (bande ayant le plus migré) (Figure 104).



Les conditions présentées ici correspondent aux tests d'activité ADN glycosylase des Fpgs de type sauvage et mutantes (E2N, E2R, E2T, E2L, E2M, E2D et E2Q) dans des extraits bruts sur le duplexe de 24 paires de bases 8-oxo-G/C.

Tout d'abord, on réalise une dilution de chaque extrait brut à partir du stock dans du tampon de dilution d'enzyme (voir annexe), afin d'obtenir les concentrations finales en Fpgs voulues. De même, on réalise une dilution des différents duplexes dans du tampon TE à 0,4 M NaCl afin d'obtenir des solutions à 80 cps.μL<sup>-1</sup> (soit environ 16 nM en duplexe).

Pour chaque type de duplexe, on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- $N \times 2 \mu\text{L}$  de duplexe à  $80 \text{ cps} \cdot \mu\text{L}^{-1}$ .
- $N \times 2 \mu\text{L}$  de tampon d'incubation 4X n°3 (voir annexe).
- $N \times 2 \mu\text{L}$  d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- $6 \mu\text{L}$  de prémélange +  $2 \mu\text{L}$  d'extrait contenant Fpg WT ou mutantes ou  $2 \mu\text{L}$  d'extrait ne contenant pas de Fpg.

La composition finale de nos essais est donc :

- $4 \text{ nM}$  en duplexe d'ADN radiomarqué.
- $25 \text{ mM}$  HEPES/NaOH (pH 7,6).
- $2,5 \text{ mM}$  Tris/HCl (pH 7,6).
- $125 \text{ mM}$  NaCl.
- $5,25 \text{ mM}$  EDTA.
- $5 \text{ mM}$   $\beta$ -mercaptoéthanol.
- $0,1 \text{ mM}$  PMSF.
- $8,75 \%$  glycérol.
- $1\text{X}$  BSA ( $100 \mu\text{g} \cdot \text{mL}^{-1}$ ).

On peut noter ici que pour avoir des temps d'incubation identiques pour toutes les formes de Fpg testées, et donc des résultats plus comparables, l'enzyme est déposée en haut des tubes, et la réaction est démarrée en même temps pour tous les essais par centrifugation. Ensuite, les essais sont incubés 20 min à  $37^\circ\text{C}$ . Enfin, toujours dans un souci d'avoir un temps de réaction identique pour tous les échantillons, la réaction est arrêtée dans tous les essais en même temps par congélation rapide dans l'azote liquide.

On ajoute alors du tampon de dénaturation Formamide (avec NaOH, voir annexe) que si pour 1X, puis on fait une dénaturation 2 min à  $95^\circ\text{C}$ . Ensuite, on dépose l'équivalent de 15 cps de chaque échantillon sur minigel dénaturant l'ADN (voir annexe), on fait migrer à  $300 \text{ V}$  constant pendant 50 minutes, et on met le gel à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL », ce qui permet de déterminer le pourcentage de coupure de l'oligonucléotide.

## IV.3.2. Etudes biochimiques avec des protéines purifiées

### IV.3.2.1. Etude de la capacité de fixation de Fpg pure sur l'ADN lésé

#### - Détermination de constantes de dissociations apparentes par gel retard

On considère l'équilibre entre une protéine P, son ligand L et le complexe qu'ils forment PL :



La constante de dissociation est donnée par :

$$K_D = \frac{[P] \times [L]}{[PL]}$$

Où  $K_D$  est la constante d'équilibre de dissociation, et  $[P]$ ,  $[L]$  et  $[PL]$  sont respectivement les concentrations à l'équilibre en protéine, en ligand et en complexe protéine/ligand.

De plus, si on considère  $[P]_0$  et  $[L]_0$  comme les concentrations initiales en protéine et en ligand, alors on a :

$$[P]_0 = [P] + [PL] \quad \text{et} \quad [L]_0 = [L] + [PL]$$

Si on se place dans le cas où  $[P]_0 \gg [L]_0$ , alors  $[P]$  est très peu différent de  $[P]_0$  à l'équilibre, et on peut dire que :

$$K_D = \frac{[P]_0 \times [L]}{[PL]}$$

Si on pose :

$$f = \text{fraction de ligand libre à l'équilibre} = [L] / [L]_0$$

On a alors :

Soit :

$$K_D = \frac{[P]_0 \cdot [L] / [L]_0}{([L]_0 - [L]) / [L]_0} \quad K_D = \frac{[P]_0 \cdot f}{1 - f}$$

D'où :

$$f = \frac{K_D}{[P]_0 + K_D}$$

On voit donc que quand la fraction de ligand libre est égale à 0,5, alors la constante d'équilibre de dissociation est égale à la concentration protéique initiale. Or lors de manipulations de type gel-retard, on peut déterminer la concentration protéique nécessaire pour avoir 50% de complexe, soit une fraction de ligand libre de 0,5, ce qui nous permet de déterminer la constante d'équilibre de dissociation correspondante.

### - Conditions standard utilisées avec Fpg WT et yOgg1 sur les cFapyG, Bz-cFapyG et cHyd

Les conditions présentées ici correspondent aux manipulations de gel-retards réalisées avec Fpg de type sauvage pure sur les duplexes de 14 paires de bases du cFapyG, du Bz-cFapyG et du cHyd appariés aux quatre bases possibles, ainsi qu'aux manipulations réalisées avec yOgg1 pure sur les duplexes de 14 paires de bases G/C, cFapyG/C et Bz-cFapyG/C. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Tout d'abord, on réalise une série de dilutions en cascade de la protéine Fpg ou yOgg1 à partir du stock à 100 µM dans du tampon de dilution de Fpg (voir annexe). De même, on réalise une dilution des différents duplexes dans du tampon TE à 0,5 M NaCl afin d'obtenir des solutions à 10 cps.µL<sup>-1</sup> (soit environ 2 nM en duplexe).

Pour chaque type de duplexe, on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- N×1 µL de duplexe à 10 cps.µL<sup>-1</sup>.
- N×2,5 µL de tampon d'incubation 4X n°1 (voir annexe).
- N×4,5 µL d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 8 µL de prémélange + 2 µL de tampon de dilution d'enzyme n°1 ou d'enzyme à une concentration donnée diluée dans ce tampon.

La composition finale de nos essais est donc :

- Quantité variable de protéine selon l'essai (autour du nM pour Fpg et autour de 100 nM pour yOgg1).
- 200 pM en duplexe d'ADN radiomarqué.
- 30 mM HEPES/NaOH (pH 7,6).

- 1 mM Tris/HCl (pH 7,6).
- 140 mM NaCl.
- 2,5 mM EDTA.
- 2,4 mM  $\beta$ -mercaptoéthanol.
- 0,12 mM PMSF.
- 11 % glycérol.
- 3,6X BSA (360  $\mu\text{g}\cdot\text{mL}^{-1}$ ).

On laisse alors la formation des complexes arriver jusqu'à l'équilibre pendant 20 minutes, on charge 5  $\mu\text{L}$  de chaque essai sur minigel (gel-retard de type n°1, voir annexe), et on fait migrer à 110 V constant pendant une heure dix. On procède alors à la fixation des bandes sur les gels dans une solution à 10 % d'acide acétique et 10 % d'éthanol, on sèche ceux-ci, et on met à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL ».

On peut donc noter ici que l'on n'est pas vraiment dans des conditions où l'on a une concentration initiale en ADN négligeable devant la concentration initiale en protéine. En effet, on a des concentrations en protéines finales de l'ordre du nanomolaire pour Fpg, donc seulement 5 fois plus de protéine que d'ADN. C'est un peu mieux pour yOgg1, avec des concentrations autour de 100 nM en protéine, soit 500 fois plus de protéine que de duplexe. Tout ceci fait que les  $K_D$  que l'on détermine ne sont pas tout à fait corrects, c'est pourquoi on parlera de  $K_D$  apparents. Par contre, les mêmes conditions ayant été utilisées pour les séries de données avec nos différents duplexes, ces  $K_D$  apparents seront tout de même valables dans un but de comparaison de nos différentes sondes oligonucléotidiques.

### **- Conditions alternatives n°1, utilisées avec les Fpg WT et mutantes sur le duplexe 8-oxo-G/C**

Les conditions utilisées ici sont identiques à celles présentées dans le sous-chapitre III.3.1.1 traitant de la fixation de Fpg contenue dans des extraits cellulaires. Par contre, ici ne sont utilisées que les protéines Fpg WT et mutantes E2Q et E2D purifiées (pour lesquelles on réalise des gammes de dilutions en cascade dans le but de déterminer des constantes de dissociations), et l'oligonucléotide 14 mere c8oxoG/C. Les conditions expérimentales sont

identiques à celles utilisées avec les extraits bactériens hormis le fait que les électrophorèses sont réalisées sur mini-gels à 110V pendant 1h10.

### - Conditions alternatives n°2, utilisées avec Fpg WT et Nth sur le duplexe cHyd/A

C'eut été intéressant de reprendre les mêmes conditions que pour Fpg WT et yOgg1 sur les cFapyG, Bz-cFapyG et cHyd, car dans ce cas on aurait pu comparer tous les résultats obtenus avec Fpg WT, Ogg1 et Nth. Cependant dans ces conditions aucun complexe n'était observable avec la protéine Nth. Afin de pouvoir observer les complexes avec les protéines Fpg et Nth, nous avons donc du changer les conditions expérimentales. De plus, la 5-OH-5-Me-Hydantoïne n'étant normalement présente qu'une face d'une adénine, seul le duplexe cHyd/A a été utilisé pour ces manipulations. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Tout d'abord, on réalise une série de dilutions en cascade de la protéine Fpg ou Nth à partir du stock dans du tampon de dilution d'enzyme n°2 (voir annexe). De même, on réalise une dilution des différents duplexes dans du tampon TE à 0,5 M NaCl afin d'obtenir des solutions à 10 cps.μL<sup>-1</sup> (soit environ 2 nM en duplexe).

Pour chaque type de duplexe, on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- N×1 μL de duplexe à 10 cps.μL<sup>-1</sup>.
- N×2,5 μL de tampon d'incubation 4X n°2 (voir annexe).
- N×4,5 μL d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 8 μL de prémélange + 2 μL de tampon de dilution d'enzyme n°2 ou d'enzyme à une concentration donnée diluée dans ce tampon.

La composition finale de nos essais est donc :

- Quantité variable de protéine selon l'essai (autour de 70 nM pour Fpg et autour de 1,4 μM pour Nth).
- 200 pM en duplexe d'ADN radiomarqué.
- 24 mM HEPES/NaOH (pH 7,6).

- 1 mM Tris/HCl (pH 7,6).
- 170 mM NaCl.
- 0,7 mM EDTA.
- 0,6 mM DTT.
- 0,12 mM PMSF.
- 6 % glycérol.

On laisse alors la formation des complexes arriver jusqu'à l'équilibre pendant 20 minutes, on charge 5  $\mu$ L de chaque essai sur minigel (gel-retard de type n°2, voir annexe), et on fait migrer à 110 V constant pendant une heure dix. On procède alors à la fixation des bandes sur les gels dans une solution à 10 % d'acide acétique et 10 % d'éthanol, on sèche ceux-ci, et on met à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL ».

A nouveau, on peut noter ici que l'on n'est pas vraiment dans des conditions où l'on a une concentration initiale en ADN négligeable devant la concentration initiale en protéine Fpg. En effet, on a des concentrations en protéines finales de l'ordre de 70 nanomolaires pour Fpg, donc seulement 350 fois plus de protéine que d'ADN. C'est par contre nettement mieux pour Nth, avec des concentrations autour de 1,4  $\mu$ M en protéine, soit 7000 fois plus de protéine que de duplexe. Tout ceci fait que l'on parlera ici aussi de  $K_D$  apparents. Par contre, les mêmes conditions ayant été utilisées pour les séries de données avec nos deux protéines, ces  $K_D$  apparents seront tout de même valables dans un but de comparaison de nos différentes protéines.

#### **IV.3.2.2. Etude de l'activité d'ADN glycosylases pures sur l'ADN lésé**

##### **- Conditions utilisées pour les cinétiques d'activité de Fpg WT sur la 8-oxo-G et la 5-OH-5-Me-Hydantoïne**

Les conditions présentées ici correspondent aux tests de cinétiques d'activité réalisées avec Fpg de type sauvage pure sur les duplexes de 14 paires de bases 8-oxoG/C, nHyd/C et nHyd/A, et sur les oligonucléotides simple-brins 8-oxo-G et nHyd. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Tout d'abord, on réalise des dilutions de la protéine Fpg à partir du stock dans du tampon de dilution d'enzyme n°1 (voir annexe) pour avoir les concentrations voulues en protéine. De même, on réalise une dilution des différents duplexes dans du tampon TE à 0,5 M NaCl afin d'obtenir des solutions à 100 cps. $\mu\text{L}^{-1}$  (soit environ 20 nM en duplexe).

Pour chaque type de duplexe, on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- $N \times 1 \mu\text{L}$  de duplexe à 100 cps. $\mu\text{L}^{-1}$ .
- $N \times 2,5 \mu\text{L}$  de tampon d'incubation 4X n°1 (voir annexe).
- $N \times 4,5 \mu\text{L}$  d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 8  $\mu\text{L}$  de prémélange + 2  $\mu\text{L}$  de tampon de dilution d'enzyme n°1 ou d'enzyme à une concentration donnée diluée dans ce tampon.

La composition finale de nos essais est donc :

- Quantité variable de Fpg selon l'essai (5 nM pour les duplexes 8-oxoG/C et nHyd/C, et 1  $\mu\text{M}$  pour les autres sondes oligonucléotidiques).
- 2 nM en duplexe d'ADN radiomarqué.
- 30 mM HEPES/NaOH (pH 7,6).
- 1 mM Tris/HCl (pH 7,6).
- 140 mM NaCl.
- 2,5 mM EDTA.
- 2,4 mM  $\beta$ -mercaptoéthanol.
- 0,12 mM PMSF.
- 11 % glycérol.
- 3,6X BSA (360  $\mu\text{g}.\text{mL}^{-1}$ ).

Des prélèvements sont alors réalisés à différents temps de réaction, et on y ajoute immédiatement du tampon de dénaturation Formamide (avec NaOH, voir annexe) que si pour 1X, puis on fait une dénaturation 2 min à 95°C. Ensuite, on dépose l'équivalent de 15 cps de chaque échantillon sur minigel dénaturant l'ADN (voir annexe), on fait migrer à 250 V constant pendant 40 minutes, et on met le gel à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL », ce qui permet de déterminer le pourcentage de coupure de l'oligonucléotide.

### **- Conditions utilisées pour les tests d'activité de Fpg WT et Nth sur la 5-OH-5-Me-Hydantoïne**

Les conditions présentées ici correspondent aux tests de d'activité réalisés avec Fpg de type sauvage pure et Nth pure sur les duplexes de 14 paires de bases nHyd/A et nHyd/C, et sur l'oligonucléotide nHyd simple-brin. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C. De plus, les conditions présentées ici sont identiques à celles du paragraphe précédent, hormis le fait que c'est ici la concentration en enzyme qui variable avec un temps de réaction identique de 30 minutes à 4°C pour tous les essais.

Comme dans les essais avec les extraits bruts, pour avoir des temps d'incubation identiques pour toutes les formes de Fpg testées, et donc des résultats plus comparables, l'enzyme est déposée en haut des tubes, et la réaction est démarrée en même temps pour tous les essais par centrifugation. Ensuite, les essais sont incubés 30 min à 4°C. Enfin, toujours dans un souci d'avoir un temps de réaction identique pour tous les échantillons, la réaction est arrêtée dans tous les essais en même temps par congélation rapide dans l'azote liquide.

### **- Conditions utilisées pour les tests d'activité de Fpg WT et mutantes E2Q et E2D sur le duplexe 8-oxo-G/C**

Les conditions présentées ici correspondent aux tests d'activité réalisées avec Fpg de type et mutantes E2D et E2Q pures sur le duplexe de 24 paires de bases 8-oxo-G/C. Ces conditions sont exactement les mêmes que pour l'étude correspondante sur les extraits bruts (chapitre III.3.1.2), hormis le fait que l'on travaille ici avec des protéines purifiées.

### **- Conditions utilisées pour tester l'effet du pH sur l'activité de Fpg WT sur le duplexe nHyd/C**

Les conditions présentées ici correspondent aux tests de l'effet du pH d'incubation sur l'activité de Fpg sur le duplexe nHyd/C. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Afin de réaliser ces manipulations, il est nécessaire de réaliser des solutions tampons de dilution d'enzyme et d'incubation des échantillons ayant différents pH, en l'occurrence des

pH de 6 , 6,5 , 7 , 7,5 , 8 , 8,5 et 9. Cependant, afin de garder un pouvoir tampon suffisant, nous avons du prendre deux tampons différents pour réaliser notre gamme de pH. Les tampons utilisés sont le tampon Phosphate pour les pH allant de 6 à 7,5, et le tampon Tris pour les pH allant de 7,5 à 9. On peut noter que le pH de 7,5 a été fait à la fois pour le tampon Phosphate et pour le tampon Tris, afin de pouvoir normaliser les valeurs de pourcentage de coupure par Fpg à ce pH, ce qui nous permet de nous affranchir d'éventuels effets du type de tampon sur l'activité de Fpg. Cependant, la réalisation de tampons de différents pH implique l'ajout de HCl ou de NaOH dans nos solutions, ce qui apporte plus ou moins de sels dans nos solutions suivant le pH voulu. Afin de s'affranchir d'éventuels effets dus aux forces ioniques différentes dans nos solutions tampons, nous avons mesuré la conductivité de nos solutions, puis normalisé celle-ci par ajout de plus ou moins de NaCl dans nos solutions. Cela nous a donc permis d'obtenir des solutions tampons à différents pH, mais avec une force ionique identique pour chaque solution. De plus, on peut noter que les forces ioniques finales dans nos solutions tampons de dilution d'enzyme et d'incubation ont été corrigées pour avoir des valeurs identiques à celles que l'on avait dans les tampons de dilution d'enzyme et d'incubation n°1 (voir annexe).

Une fois que ces solutions tampons de dilution d'enzyme et d'incubation (n°4 voir annexe) ont été réalisées, nous avons fait des tests d'activité suivant les mêmes conditions que précédemment (voir le paragraphe sur les conditions utilisées pour les cinétiques d'activité de Fpg WT sur la 8-oxo-G et la 5-OH-5-Me-Hydantoïne).

La composition finale de nos essais est donc :

- 5 nM en protéine Fpg.
- 2 nM en duplexe d'ADN radiomarké.
- 25 mM de tampon Phosphate ou Tris au pH voulu.
- 1 mM Tris/HCl (pH 7,6).
- Quantité de NaCl nécessaire pour avoir une conductivité identique dans tous nos essais.
- 2,5 mM EDTA.
- 2,4 mM  $\beta$ -mercaptoéthanol.
- 0,12 mM PMSF.
- 11 % glycérol.
- 3,6X BSA ( $360 \mu\text{g.mL}^{-1}$ ).

Enfin, les résultats de quantification des expériences ont été normalisés, en fixant une valeur de 100 unités arbitraires (UA) pour les pourcentages de coupure obtenus au pH physiologique de 7,5 (avec les deux types de Tampon, phosphate et Tris).

#### IV.3.2.3. Etude de l'inhibition de l'activité ADN glycosylase de Fpg pure sur le nHyd par son analogue cHyd

Les conditions présentées ici correspondent aux tests d'inhibition de l'activité ADN glycosylase de Fpg sur le duplexe nHyd/C par un duplexe sans lésion G/C ou un duplexe cHyd/C. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, toutes les étapes précédant la fixation du gel sont réalisées à 4°C.

Les conditions utilisées ici sont quasi-identiques à celles que l'on a utilisées pour les manipulations précédentes pour l'activité de Fpg sur les 14 meres (voir le paragraphe sur les conditions utilisées pour les cinétiques d'activité de Fpg WT sur la 8-oxo-G et la 5-OH-5-Me-Hydantoïne, et les annexes correspondantes). La seule différence vient du fait que l'on doit apporter aussi de l'ADN compétiteur cHyd/C ou G/C non radiomarké en plus du duplexe nHyd/C radiomarké. Cet ajout implique donc un apport de sel supplémentaire par rapport à ce que l'on faisait précédemment, ce qui fait que nous avons utilisé ici un tampon d'incubation 4X n°1 sans NaCl, contre 200 mM NaCl pour les manipulations précédentes.

Tout d'abord, on réalise une dilution de la protéine Fpg à partir du stock dans du tampon de dilution d'enzyme n°1 (voir annexe) pour avoir une concentration finale de 60 nM. De même, on réalise une dilution du duplexe nHyd/C radiomarké dans du tampon TE à 0,5 M NaCl afin d'obtenir une solution à 100 cps.µL<sup>-1</sup> (soit environ 20 nM en duplexe). D'autre part, on réalise une série de dilutions en cascade des duplexes G/C et cHyd/C à partir du stock à 4 µM dans du tampon TE à 0,5 M NaCl pour obtenir des concentrations de 30, 60, 120, 240, 480 et 960 nM.

Ensuite on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- N×1 µL de duplexe à 100 cps.µL<sup>-1</sup>.
- N×2,5 µL de tampon d'incubation 4X n°1 (voir annexe).
- N×3,5 µL d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 7 µL de prémélange + 2 µL de Fpg à 60 nM (dans le tampon de dilution d'enzyme n°1) + 1 µL d'ADN inhibiteur G/C ou cHyd/C.

La composition finale de nos essais est donc :

- 12 nM en Fpg.
- 2 nM en duplexe d'ADN radiomarqué.
- 3 à 96 nM en duplexe compétiteur G/C ou cHyd/C.
- 30 mM HEPES/NaOH (pH 7,6).
- 2 mM Tris/HCl (pH 7,6).
- 140 mM NaCl.
- 2,6 mM EDTA.
- 2,4 mM  $\beta$ -mercaptoéthanol.
- 0,12 mM PMSF.
- 11 % glycérol.
- 3,6X BSA ( $360 \mu\text{g}\cdot\text{mL}^{-1}$ ).

Afin d'avoir un même temps d'incubation pour tous les échantillons, l'enzyme est ajoutée en même temps dans tous les échantillons, l'incubation dure 30 min à 4°C, puis la réaction est arrêtée en même temps dans tous les échantillons. On ajoute alors du tampon de dénaturation Formamide (avec NaOH, voir annexe) que si pour 1X, puis on fait une dénaturation 2 min à 95°C. Ensuite, on dépose l'équivalent de 15 cps de chaque échantillon sur minigel dénaturant l'ADN (voir annexe), on fait migrer à 300 V constant pendant 50 minutes, et on met le gel à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL », ce qui permet de déterminer le pourcentage de coupure de l'oligonucléotide en fonction de la quantité d'ADN compétiteur dans chaque essai.

#### **IV.3.2.4. Etude de la formation d'un complexe covalent entre Fpg et la 5-OH-5-Me-Hydantoïne**

##### **IV.3.2.4.1. Test SDS-PAGE-TRAP pour la mise en évidence de pontages covalents entre LIFpg et l'Hyd ou le cHyd**

Le principe de base de cette méthode est identique à celui d'un SDS-PAGE classique, reposant sur la présence de SDS dans le milieu réactionnel pour permettre la dénaturation protéique. Par contre, on utilise ici cette méthode pour mettre en évidence des pontages covalents ADN-protéine. Ainsi, une fois la protéine dénaturée par le SDS, seuls les complexes ADN-protéine covalents seront conservés car la dénaturation protéique sépare les partenaires des complexes non-covalents. Enfin, le brin d'ADN contenant l'espèce pontable covalamment

par des enzymes (l'Hyd ou le cHyd) est marqué radioactivement au  $^{32}\text{P}$ , ce qui permet de révéler les résultats de ces expériences par autoradiographie, contrairement aux SDS-PAGEs classiques qui sont révélés au bleu de Coomassie.

#### **IV.3.2.4.2. Mise en évidence et caractérisation préliminaire du complexe covalent**

##### **- Vérification du caractère nucléoprotéique du complexe covalent (correspondant à la Figure 70)**

Les conditions présentées ici correspondent aux tests de pontage covalent par SDS-PAGE-TRAP de *L/Fpg* sur le duplexe cHyd/C, avec et sans traitement subséquent à la Protéinase K, et dont le but est de vérifier la nature nucléoprotéique du complexe formé. Les duplexes de 14 paires de bases ayant une forte tendance à se déshybrider, les échantillons sont d'abord incubés 10 min à 4°C pour laisser les complexes non-covalents se faire, puis ils sont incubés à 37°C. De plus, la solution de tampon d'incubation utilisée ici ne contient pas de BSA (tampon d'incubation n°5, voir annexe), car nous avons remarqué que la présence de cette protéine dans le milieu réactionnel avait tendance à être défavorable à la création de pontage covalent entre Fpg et la 5-OH-5-Me-Hydantoïne (données non présentées).

Tout d'abord, on utilise de la Fpg à 50  $\mu\text{M}$  en solution dans son tampon de stockage de Fpg à 5% de glycérol (voir annexe). De plus, on réalise une dilution du duplexe nHyd/C radiomarqué dans du tampon TE à 0,4 M NaCl afin d'obtenir une solution à 100 cps. $\mu\text{L}^{-1}$  (soit environ 20 nM en duplexe).

Ensuite on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- N×2,5  $\mu\text{L}$  de duplexe à 100 cps. $\mu\text{L}^{-1}$ .
- N×1,875  $\mu\text{L}$  de tampon d'incubation 4X n°5 (voir annexe).
- N×3,125  $\mu\text{L}$  d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- 7,5  $\mu\text{L}$  de prémélange + 2,5  $\mu\text{L}$  de tampon de stockage de Fpg ou de Fpg à 50  $\mu\text{M}$  diluée dans ce tampon.

La composition finale de nos essais est donc :

- 12,5  $\mu$ M en Fpg (sauf pour le contrôle négatif, qui n'en contient pas).
- 5 nM en duplexe cHyd/C radiomarqué.
- 21,25 mM HEPES/NaOH (pH 7,6).
- 2,5 mM Tris/HCl (pH 7,6).
- 125 mM NaCl.
- 1 mM EDTA.
- 1,5 mM  $\beta$ -mercaptoéthanol.
- 0,25 mM TCEP.
- 0,1 mM PMSF.
- 6,5 % glycérol.

L'enzyme est ajoutée en même temps dans tous les échantillons, les essais sont incubés 10 min à 4°C puis 18 heures à 37°C, et la réaction est arrêtée en même temps dans tous les échantillons. Dans l'un des essais contenant de la Fpg, on ajoute alors 2  $\mu$ L de Protéinase K (Fermentas), et on incube 60 min supplémentaires à 37°C.

On ajoute alors du tampon de dénaturation Laemmli (voir annexe) que si pour 1X, puis on fait une dénaturation 2 min à 90°C. Ensuite, on dépose l'équivalent de 15 cps de chaque échantillon sur minigel de type SDS-PAGE (voir annexe), et on fait migrer à 200 V constant pendant 40 minutes. On sèche alors les gels obtenus, et on met le gel à révéler pendant une nuit contre un écran radiosensible. Enfin, on scanne cet écran grâce à un phospho-imageur (Storm 860, Amersham Pharmacia Biotech), et on quantifie l'intensité des bandes sur les images générées grâce au programme « ImageQuant TL ».

**- Effet de la température d'incubation sur l'efficacité des pontages covalents formés entre Fpg et la 5-OH-5-Me-Hydantoïne (correspondant à la Figure 71)**

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de l'effet de la température d'incubation sur l'efficacité de pontage covalent entre Fpg de type sauvage et le duplexe cHyd/C ou le cHyd simple-brin. Ces manipulations auront donc deux buts, celui de voir l'effet de la température d'incubation, et celui de voir l'effet du contexte simple- ou double-brin du cHyd. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que des échantillons identiques ont été incubés quatre jours à 4, 25 et 37°C.

**- Cinétiques de pontage covalent entre Fpg WT et le duplexe cHyd/C**  
(correspondant à la Figure 72)

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de cinétiques de pontage covalent entre Fpg de type sauvage et le duplexe cHyd/C, ceci afin de tester l'effet du temps d'incubation à 37°C sur l'apparition du pontage covalent. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que l'on réalise deux cinétiques de pontage avec des concentrations finales en Fpg de 12,5 à 25 µM.

**- Tests de l'effet de l'étape de dénaturation thermique sur la quantité de pontage covalent détectée**  
(correspondant à la Figure 73)

Ce sous-chapitre a pour but de tester l'effet de l'étape de dénaturation thermique des échantillons durant le SDS-PAGE-TRAP sur la quantité de pontage covalent que l'on détecte entre Fpg et le duplexe cHyd/C.

Les conditions sont identiques à celles du sous-chapitre sur la « vérification du caractère nucléoprotéique du complexe covalent », hormis pour la concentration finale en protéine Fpg qui est ici de 12,5 µM, et pour l'incubation qui dure ici quatre jours à 37°C. Par contre, après incubation chaque essai est divisé en deux, et chaque partie est traitée de deux façons différentes. La première partie est analysée comme précédemment, en faisant une étape de dénaturation thermique des échantillons 2 min à 90°C, alors que la seconde partie des échantillons ne subit pas cette étape. Le reste de l'analyse se déroule de façon identique pour tous les échantillons.

**- Tests de l'effet du pH sur la stabilité du complexe covalent formé entre L/Fpg et le duplexe cHyd/C**  
(correspondant à la Figure 74)

Ce sous-chapitre a pour but de tester l'effet de l'ajout d'acide ou de base après formation du pontage covalent entre Fpg et le duplexe cHyd/C, et avant l'analyse par SDS-PAGE-TRAP. On teste donc la stabilité acido-basique du complexe covalent créé.

Les conditions sont identiques à celles du sous-chapitre sur la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que l'incubation est réalisée pendant 4 jours à 37°C, et que la concentration finale en protéine Fpg est ici de 25 µM. Après cette incubation, on ajoute un même volume correspondant à des quantités variables de soude ou d'acide chlorhydrique dans chaque essai, afin d'obtenir des concentrations finales en NaOH ou HCl de 10 et 100 mM. Une quantité variable d'acide ou de base est alors ajoutée à chaque essai, puis on fait une incubation de 30 min à 50°C. La quantité de complexe covalent est normalisée par rapport à celle que l'on a dans l'essai sans ajout d'acide ou de base, pour lequel on fixe une valeur de 100 unités arbitraires (UA). De plus, on réalise un dernier essai dans lequel on ajoute uniquement de l'eau, ceci afin d'avoir un contrôle positif pour notre manipulation. Enfin, les essais sont analysés comme précédemment par SDS-PAGE-TRAP.

#### **IV.3.2.4.3. Analyse comparative de la spécificité de la réaction de pontage et de l'activité Hyd-ADN glycosylase de *LI*Fpg**

##### **- Tests de pontages covalents entre Fpg et le cHyd/C ou le nHyd/C, en présence ou en absence de borohydrure de sodium** **(correspondant à la Figure 75)**

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de pontage covalent de Fpg sur les duplexes cHyd/C et nHyd/C, avec et sans traitement subséquent au borohydrure de sodium (NaBH<sub>4</sub>). Ces manipulations ont en fait pour but de vérifier si le pontage covalent entre Fpg et le cHyd est aussi possible avec le nHyd, ainsi que de vérifier que ce complexe covalent n'est pas de type base de Schiff réduite. Les conditions sont quasiment les mêmes que pour le sous-chapitre précédent, hormis le fait que l'on a du adapter la préparation de nos échantillons pour pouvoir ajouter à certains du borohydrure de sodium sans rien changer à la composition globale des échantillons.

Tout d'abord, on utilise de la Fpg à 100 µM en solution dans son tampon de stockage de Fpg à 5% de glycérol (voir annexe). De plus, on réalise une dilution du duplexe nHyd/C radiomarké dans du tampon TE à 0,4 M NaCl afin d'obtenir une solution à 100 cps.µL<sup>-1</sup> (soit environ 20 nM en duplexe). Enfin, on a fait une solution de NaBH<sub>4</sub> à 0,9 M juste avant de réaliser nos échantillons (car l'ion borohydrure en solution donne rapidement du dihydrogène et l'ion boroxyde, qui n'est plus réducteur).

A ce niveau, nous avons voulu vérifier si le complexe covalent que l'on observe était ou non le même complexe covalent que l'intermédiaire de type base de Schiff réduite que l'on peut observer en conditions réductrices avec n'importe quel substrat de Fpg. De plus, les manipulations précédentes ayant été faites avec le cHyd, nous avons aussi testé la capacité du nHyd à réaliser ce type de complexe covalent avec Fpg. Des incubations de quatre jours à 37°C de Fpg avec un duplexe d'ADN contenant un cHyd ou un nHyd placé en face d'un C ont donc été réalisées, en testant dans chaque cas l'effet de l'ajout de borohydrure de sodium ( $\text{NaBH}_4$ , réducteur fort). On peut préciser que les essais en présence de borohydrure de sodium n'ont été incubés que 10 min à 37°C, afin de permettre à la réaction borohydrure dépendante de se faire, mais sans avoir trop de réaction borohydrure indépendante.

Ensuite on réalise le prémélange suivant (N étant le nombre d'échantillons à réaliser) :

- $N \times 2,5 \mu\text{L}$  de duplexe à  $100 \text{ cps} \cdot \mu\text{L}^{-1}$ .
- $N \times 1,875 \mu\text{L}$  de tampon d'incubation 4X n°5 (voir annexe).
- $N \times 1,125 \mu\text{L}$  d'eau ultrapure.

Chaque échantillon est alors réalisé dans les conditions suivantes :

- $5,5 \mu\text{L}$  de prémélange +  $2,5 \mu\text{L}$  de tampon de stockage de Fpg ou de Fpg à  $100 \mu\text{M}$  diluée dans ce tampon +  $2 \mu\text{L}$  d'eau ultrapure ou de  $\text{NaBH}_4$  fraîchement dissout.

La composition finale de nos essais est donc :

- $25 \mu\text{M}$  en Fpg (sauf pour le contrôle négatif, qui n'en contient pas).
- $5 \text{ nM}$  en duplexe cHyd/C ou nHyd/C radiomarqués.
- $180 \text{ mM}$  de  $\text{NaBH}_4$  (seulement pour certains échantillons).
- $21,25 \text{ mM}$  HEPES/NaOH (pH 7,6).
- $2,5 \text{ mM}$  Tris/HCl (pH 7,6).
- $125 \text{ mM}$  NaCl.
- $1 \text{ mM}$  EDTA.
- $1,5 \text{ mM}$   $\beta$ -mercaptoéthanol.
- $0,25 \text{ mM}$  TCEP.
- $0,1 \text{ mM}$  PMSF.
- $6,5 \%$  glycérol.

Enfin, les réactions sont arrêtées et analysées comme précédemment par SDS-PAGE-TRAP.

**- Effet du pH sur la formation du complexe covalent entre Fpg et le duplexe cHyd/C**

**(correspondant à la Figure 76)**

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de l'effet du pH d'incubation sur l'efficacité de pontage covalent entre Fpg de type sauvage et le duplexe cHyd/C. Les tampons utilisés sont les tampons de dilution d'enzyme et d'incubation n°4 (voir annexe), la concentration protéique finale est de 20  $\mu\text{M}$ , et l'incubation dure ici quatre jours à 37°C.

On peut noter que le stock d'enzyme utilisé ici est à 396,25  $\mu\text{M}$  dans du tampon HEPES pH 7,6 à 10 mM. On fait une dilution de l'enzyme que si pour 100  $\mu\text{M}$  avec du tampon de dilution d'enzyme n°4 au pH désiré, soit une dilution d'environ  $\frac{1}{4}$ , on se trouve donc à environ 2,5 mM HEPES pH 7,6. De plus, on met 2  $\mu\text{L}$  d'enzyme dans 10  $\mu\text{L}$  d'essai, on est donc à 0,5 mM HEPES pH 7,6 dans les essais. Sachant que le tampon n°4 de pH désiré est à 25 mM dans les essais, on peut considérer que la quantité d'HEPES pH 7,6 est négligeable et que le pH des essais est uniquement dicté par le tampon n°4 de pH désiré.

On réalise les essais suivants :

- 1  $\mu\text{L}$  de duplexe cHyd/C à 100 cps. $\mu\text{L}^{-1}$ .
- 2  $\mu\text{L}$  de tampon d'incubation 4X n°4 (voir annexe) au pH voulu.
- 2  $\mu\text{L}$  de Fpg à 100  $\mu\text{M}$  dans du tampon de dilution d'enzyme n°4 (voir annexe) au pH voulu.
- 5  $\mu\text{L}$  d'eau ultrapure.

La composition finale de nos essais est donc :

- 20  $\mu\text{M}$  en Fpg.
- 2 nM en duplexe cHyd/C radiomarqué.
- 25 mM de tampon Phosphate ou Tris au pH voulu.
- Quantité de NaCl nécessaire pour avoir une conductivité identique dans tous nos essais.
- 2,2 mM EDTA.
- 2 mM  $\beta$ -mercaptoéthanol.
- 0,1 mM PMSF.
- 5 % glycérol.

Ensuite, les réactions sont incubées quatre jours à 37°C, puis arrêtées et analysées comme précédemment par SDS-PAGE-TRAP. Enfin, les résultats de quantification des

expériences ont été normalisés, en fixant une valeur de 100 unités arbitraires (UA) pour les pourcentages de coupure obtenus au pH physiologique de 7,5 (avec les deux types de Tampon, phosphate et Tris).

**- Tests de pontages covalents entre Fpg et les duplexes cHyd/C et cHyd/A pour voir l'effet de la base opposée à la lésion (correspondant à la Figure 77A)**

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de pontage covalent entre Fpg de type sauvage et les duplexes cHyd/C ou cHyd/A, ceci afin de tester l'effet du changement de base opposée à la lésion sur la possibilité d'avoir un pontage covalent. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que l'incubation dure ici quatre jours à 37°C, et que l'on réalise différents essais avec des concentrations finales en Fpg allant de 1,25 à 25 µM.

**- Tests de pontage covalent entre Fpg WT et le cHyd ou le nHyd, sous formes simple-brins ou appariés à une cytosine (correspondant à la Figure 77B)**

Ce sous-chapitre a pour but de préciser par SDS-PAGE-TRAP les efficacités de pontage de Fpg de type sauvage sur la 5-OH-5-Me-Hydantoïne naturelle ou greffée sur un cyclopentane, dans un contexte simple-brin ou appariés à une cytosine. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que les échantillons ont été incubés ici quatre jours à 37°C.

#### **IV.3.2.4.4. Bases fonctionnelles et sélectivité de la réaction suicide entre L/Fpg et le cHyd**

##### **- Tests de pontages covalents entre Fpg WT ou mutantes et le duplexe cHyd/C** **(correspondant à la Figure 83)**

Les conditions utilisées ici correspondent aux tests par SDS-PAGE-TRAP de pontage covalent entre Fpg de type sauvage ou mutantes  $\Delta$ P1, P1G, E2Q et E2D, et le duplexe cHyd/C. Les mutations de résidus de l'extrémité N-terminale de l'enzyme sont susceptibles de nous apporter des informations sur le résidu de Fpg impliqué dans le pontage covalent avec le cHyd, étant donné que l'on suspecte que le résidu N-terminal de Fpg soit impliqué dans ce pontage. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », hormis le fait que l'incubation dure ici

##### **- Tests de pontage covalent entre le duplexe cHyd/C et différentes protéines** **(correspondant à la Figure 83A)**

Ce sous-chapitre a pour but de tester par SDS-PAGE-TRAP la capacité de différentes protéines autres que Fpg à réaliser du pontage covalent avec le duplexe cHyd/C. Les conditions sont les mêmes que pour le sous-chapitre traitant de la « vérification du caractère nucléoprotéique du complexe covalent », et la concentration finale de chaque protéine est ici de 12,5  $\mu$ M.

##### **- Test de pontage covalent entre yOgg1 et le duplexe cHyd/C** **(correspondant à la Figure 83B)**

Ce sous-chapitre a pour but de tester par SDS-PAGE-TRAP la capacité de l'homologue fonctionnel yOgg1 de Fpg chez la levure à réaliser du pontage covalent avec le duplexe cHyd/C. Les conditions sont exactement identiques à celles du paragraphe précédent, hormis pour la concentration finale en yOgg1 qui est ici de 19,3  $\mu$ M.

**Solutions tampons et  
milieux utilisés**

## Solutions tampons et milieux utilisés

### Liste des solutions tampons et milieux utilisés :

- Milieu LB.
- Milieu NZY+.
- Tampon de dénaturation Formamide 2X.
- Tampon de dénaturation Laemmli 4X.
- Tampon de lyse des cellules.
- Tampon de migration et composition des gel-retards n°1.
- Tampon de migration et composition des gel-retards n°2.
- Tampon de migration et composition des gels dénaturant l'ADN.
- Tampon de migration et composition des gels dénaturant les protéines de type SDS-PAGE.
- Tampon de stockage de Fpg.
- Tampon d'hybridation de l'ADN n°1.
- Tampon d'hybridation de l'ADN n°2.
- Tampon d'incubation n°5.
- Tampon QMA.
- Tampons de dilution d'enzyme et d'incubation n°1.
- Tampons de dilution d'enzyme et d'incubation n°2.
- Tampons de dilution d'enzyme et d'incubation n°3.
- Tampons de dilution d'enzyme et d'incubation n°4.
- Tampons MonoQ.
- Tampons SP.
- Tampon TBE 5X (pH 7,6).
- Tampon TE 10X (pH 7,6).

### Composition des solutions tampons et milieux utilisés :

#### Milieu LB :

(Pour les cultures bactériennes).

- 20 g.L<sup>-1</sup> de milieu LB commercial.

Autoclaver (20 min à 120°C) et conserver à température ambiante.

Pour obtenir du milieu LB/Agar, ajouter 35 g.L<sup>-1</sup> final en agar au milieu LB avant autoclavage.

Pour obtenir du milieu LB/Ampicilline, ajouter 100 mg/L<sup>-1</sup> final en ampicilline après autoclavage.

#### Milieu NZY+ :

(Pour les transformations bactériennes).

- 10 g.L<sup>-1</sup> de N-Z amine (hydrolysate de caséine).
- 5 g.L<sup>-1</sup> d'extrait de levure.
- 5 g.L<sup>-1</sup> de NaCl.

pH à 7,5, autoclaver (20 min à 120°C), aliquoter par 10 mL et conserver à 4°C.

Juste avant utilisation, ajouter en conditions stériles dans un aliquot :

- 1,250 mM final en MgCl<sub>2</sub> (filtré à 0,2 µm en conditions stériles).
- 1,250 mM final en MgSO<sub>4</sub> (filtré à 0,2 µm en conditions stériles).
- 20 mM en glucose (filtré à 0,2 µm en conditions stériles).

- Tampon de dénaturation Formamide 2X :

(Pour les gels dénaturant l'ADN).

- 80 % de Formamide déionisé (v/v).
  - 10 mM EDTA.
  - 0,1 % de Xylène cyanol (p/v).
  - 0,1 % de Bleu de bromophénol (p/v).
  - 100 mM de NaOH (pour la coupure des sites abasiques).
- Filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampon de dénaturation Laemmli 4X :

(Pour les gels dénaturant les protéines de type SDS-PAGE).

- 25 mM de Tris/HCl (pH 6,8).
  - 25 % de glycérol (v/v).
  - 4 % de SDS.
  - 4 % de β-mercaptoéthanol (v/v), soit environ 574 mM.
  - 0,2 % de Bleu de bromophénol (p/v).
- pHer à 6,8, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampon de lyse des cellules :

- 25 mM d'Hepes/NaOH pH 7,6.
  - 250 mM de NaCl.
  - 1 mM d'EDTA.
  - 5 mM de β-mercaptoéthanol.
  - 100 µM de PMSF.
  - 5 % glycérol.
- pHer à 7,6, filtrer à 0,2 µm et conserver à 4°C. Juste avant utilisation, ajouter 0,1 % de Triton X-100 (v/v).

- Tampon de migration et composition des gel-retards de type n°1 :

(Sauf pour les manipulations comparatives avec Nth).

Tampon de migration (1X) :

- 0,09 M Tris/base.
  - 0,08 M acide borique.
- Filtrer à 0,2 µm et conserver à température ambiante.

Composition des gels retards :

- 10 % d'acrylamide/bisacrylamide 19,76/0,24 (v/v).
- 1X en Tampon de migration.
- 0,8 % de Persulfate d'ammonium (p/v).
- 0,05 % de TEMED (v/v).

- Tampon de migration et composition des gel-retards de type n°2 :

(Pour les manipulations comparatives avec Nth).

Tampon de migration :

- TBE 0,5X.

Composition des gels retards :

- 10 % d'acrylamide/bisacrylamide 37,5/1 (v/v).
- 0,5X en Tampon de migration.
- 0,08 % de Persulfate d'ammonium (p/v).
- 0,05 % de TEMED (v/v).

- Tampon de migration et composition des gels dénaturant l'ADN :

Tampon de migration :

- TBE 0,5X dans le compartiment supérieur de la cuve, et 1,5 X dans le compartiment inférieur.

Composition des gels dénaturant l'ADN :

- 20 % d'acrylamide/bisacrylamide 19/1 (v/v).

- Gradient de Tampon de migration allant de 0,5X à 1,5X du haut en bas du gel.

- 7 M d'Urée.

- 0,015 % de Persulfate d'ammonium (p/v).

- 0,05 % de TEMED (v/v).

- Tampon de migration et composition des gels dénaturant les protéines de type SDS-PAGE:

Tampon de migration (1X) :

- 0,1 M de Tris/base.

- 0,2 M de Glycine.

- 0,1 % de SDS (p/v).

Filtrer à 0,2 µm et stocker à température ambiante (le SDS est ajouté après filtration).

Composition des gels dénaturant l'ADN :

Gel de concentration :

- 4,5 % d'acrylamide/bisacrylamide 29/1 (v/v).

- 125 mM de Tris/HCl (pH 6,8).

- 0,1 % de SDS (p/v).

- 0,1 % de Persulfate d'ammonium (p/v).

- 0,1 % de TEMED (v/v).

Gel de séparation :

- 10 à 16 % d'acrylamide/bisacrylamide 29/1 (v/v).

- 0,375 mM de Tris/HCl (pH 8,8).

- 0,1 % de SDS (p/v).

- 0,075 % de Persulfate d'ammonium (p/v).

- 0,075 % de TEMED (v/v).

- Tampon de stockage de Fpg :

- 20 mM de Hepes/NaOH pH 7,6.

- 5 mM de TCEP.

- 100 µM de PMSF.

- 5 % de glycérol pour le stockage à -80°C, et 50 % de glycérol pour celui à -20°C.

- 100 mM de NaCl.

pHer à 7,6, filtrer à 0,2 µm et conserver à -20°C.

- Tampon d'hybridation de l'ADN n°1 :

(Pour la biochimie).

- TE 1X (pH 7,6).

- 500 mM NaCl.

pHer à 7,6, filtrer à 0,2 µm et conserver à température ambiante.

- Tampon d'hybridation de l'ADN n°2 :

(Pour la cristallographie).

- 5 mM de Cacodylate (pH 7).

- 200 mM de KCl.

- 0,5 mM d'EDTA.

pHer à 7, filtrer à 0,2 µm et conserver à 4°C.

- Tampon d'incubation n°5 :

(Pour les tests de pontage covalent avec la 5-OH-5-Me-Hydantoïne).

Tampon d'incubation (4X) :

- 100 mM d'HEPES/NaOH (pH 7,6).

- 4 mM d'EDTA.

- 8 mM de β-mercaptoéthanol.

- 0,4 mM de PMSF.

- 28 % de glycérol.

pHer à 7,6, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampon QMA :

(Pour la purification de Fpg par chromatographie sur colonne QMA ).

- 25 mM d'Hepes/NaOH pH 7,6.

- 250 mM de NaCl.

- 1 mM d'EDTA.

- 5 mM de β-mercaptoéthanol.

- 100 µM de PMSF.

- 5 % glycérol.

pHer à 7,6, filtrer à 0,2 µm et conserver à 4°C.

- Tampons MonoQ :

(Pour la purification d'oligonucléotides par chromatographie sur colonne MonoQ 10/100 GL).

- 10 mM de NaOH.

- 10 mM de NaCl pour le tampon MonoQ de type A, et 1 M de NaCl pour le type B.

Filtrer à 0,2 µm et conserver à température ambiante.

- Tampons de dilution d'enzyme et d'incubation n°1 :

(Pour toutes les manipulations, sauf pour les gel-retards comparatifs de Fpg et Nth, pour les tests d'effet du pH de réaction, pour les tests sur les mutants de Fpg, et pour les tests de pontage covalent avec la 5-OH-5-Me-Hydantoïne).

Tampon de dilution de Fpg :

- 25 mM d'HEPES/NaOH (pH 7,6).

- 200 mM de NaCl.

- 2 mM d'EDTA.

- 2 mM de β-mercaptoéthanol.

- 0,1 mM de PMSF.

- 5 % de glycérol.

- 3X en BSA (soit 300 µg.mL<sup>-1</sup>).

pHer à 7,6, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

Tampon d'incubation (4X) :

- 100 mM d'HEPES/NaOH (pH 7,6).

- 200 mM de NaCl.

- 8 mM d'EDTA.

- 8 mM de β-mercaptoéthanol.

- 0,4 mM de PMSF.

- 40 % de glycérol.

- 12X en BSA (soit 1,2 mg.mL<sup>-1</sup>).

pHer à 7,6, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampons de dilution d'enzyme et d'incubation n°2 :

(Pour les gel-retards comparatifs de Fpg et Nth).

Tampon de dilution d'enzyme :

- 20 mM d'HEPES/NaOH (pH 7,6).
- 100 mM de NaCl.
- 0,5 mM d'EDTA.
- 0,5 mM de DTT.
- 5 % de glycérol.

pHer à 7,8, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

Tampon d'incubation (4X) :

- 80 mM d'HEPES/NaOH (pH 7,6).
- 400 mM de NaCl.
- 2 mM d'EDTA.
- 2 mM de DTT.
- 20 % de glycérol.

pHer à 7,8, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampons de dilution d'enzyme et d'incubation n°3 :

(Pour les tests sur les mutants de Fpg).

Tampon de dilution de Fpg :

- 25 mM d'HEPES/NaOH (pH 7,6).
- 100 mM de NaCl.
- 5 mM d'EDTA.
- 5 mM de β-mercaptoéthanol.
- 0,1 mM de PMSF.
- 5 % de glycérol.
- 1X en BSA (soit 100 µg.mL<sup>-1</sup>).

pHer à 7,6, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

Tampon d'incubation (4X) :

- 75 mM d'HEPES/NaOH (pH 7,6).
- 15 mM d'EDTA.
- 15 mM de β-mercaptoéthanol.
- 0,3 mM de PMSF.
- 30 % de glycérol.
- 3X en BSA (soit 300 µg.mL<sup>-1</sup>).

pHer à 7,6, filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

- Tampons de dilution d'enzyme et d'incubation n°4 :

(Pour les tests d'effet du pH de réaction).

Tampons de dilution de Fpg :

- 25 mM de Tampon phosphate ou Tris au pH désiré.
- Quantité nécessaire de NaCl pour avoir la même conductivité que dans le tampon de dilution d'enzyme n°1.
- 2 mM d'EDTA.
- 2 mM de β-mercaptoéthanol.
- 0,1 mM de PMSF.
- 5 % de Glycérol.

Filtrer à 0,2 µm, aliquoter par 2 mL et conserver à -20°C.

Tampons d'incubation (4X) :

- 100 mM de Tampon phosphate ou Tris au pH désiré.
  - Quantité nécessaire de NaCl pour avoir la même conductivité que dans le tampon d'incubation 4X n°1.
  - 8 mM d'EDTA.
  - 8 mM de  $\beta$ -mercaptoéthanol.
  - 0,4 mM de PMSF.
  - 20 % de Glycérol.
- Filtrer à 0,2  $\mu$ m, aliquoter par 2 mL et conserver à -20°C.

- Tampons SP :

(Pour la purification de Fpg par chromatographie sur colonnes SPFF ou SPHP).

- 25 mM Hepes/NaOH pH 7,6.
  - 5 mM  $\beta$ -mercaptoéthanol.
  - 100  $\mu$ M de PMSF.
  - 5 % glycérol.
  - pas de NaCl pour le tampon SP de type A, et 1 M de NaCl pour le type B.
- pHer à 7,6, filtrer à 0,2  $\mu$ m et conserver à 4°C.

- Tampon TBE 5X (pH 7,6) :

- 450 mM de Tris/base.
  - 450 mM d'Acide borique.
  - 10 mM d'EDTA.
- pHer à 7,6, filtrer à 0,2  $\mu$ m et conserver à température ambiante.

- Tampon TE 10X (pH 7,6) :

- 100 mM de Tris/base.
  - 10 mM EDTA.
- pHer à 7,6, filtrer à 0,2  $\mu$ m et conserver à température ambiante.

**Références  
bibliographiques**

## Références bibliographiques

### **-A-**

- Agarwal, R.C. (1978). A new least-squares refinement technique based on the Fast Fourier Transform algorithm. *Acta. Cryst.* **A34**, 791-809.
- Alekseyev, Y.O., Hamm, M.L., and Essigmann, J.M. (2004). Aflatoxin B<sub>1</sub> formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway *in vivo*. *Carcinogenesis* **25**, 1045-1051.
- Alseth, I., Rognes, T., Lindbäck, T., Solberg, I., Robertsen, K., Kristiansen, K.I., Mainieri, D., Lillehagen, L., Kolstø, A.-B., and Bjørås, M. (2006). A new protein superfamily includes two novel 3-methyladenine DNA glycosylases from *Bacillus cereus*, AlkC and AlkD. *Molecular Microbiology* **59**, 1602-1609.
- Aravind, L., and Koonin, E.V. (2000). The  $\alpha/\beta$  fold uracil DNA glycosylases: a common origin with diverse fates. *Genome Biology* **1**, 1-8.
- Arczewska, K.D., and Kuśmierk, J.T. (2007). Bacterial DNA repair genes and their eukaryotic homologues: 2. Role of bacterial mutator gene homologues in human disease. Overview of nucleotide pool sanitization and mismatch repair systems. *Acta Biochimica Polnica* **54**, 435-457.
- Asagoshi, K., Terato, H., Ohyama, Y., and Ide, H. (2002). Effects of a Guanine-derived Formamidopyrimidine Lesion on DNA Replication. *J. Biol. Chem.* **277**, 14589-14597.
- Asagoshi, K., Yamada, T., Terato, H., Ohyama, Y., Monden, Y., Arai, T., Nishimura, S., Aburatani, H., Lindahl, T., and Ide, H. (2000). Distinct repair activities of human 7,8-dihydro-8-oxoguanine DNA Glycosylase and Formamidopyrimidine DNA Glycosylase for Formamidopyrimidine and 7,8-dihydro-8-oxoguanine. *J. Biol. Chem.* **275**, 4956-4964.
- Asselt Van, E.J., Perrakis, A., Kalk, K.H., and Lamzin, V.S. (1998). Accelerated X-ray structure elucidation of a 36 kDa muramidase/transglycosylase using wARP. *Acta. Cryst.* **D54**, 58-73.
- Au, K.G., Clark, S., Miller, J.H., and Modrich, P. (1989). *Escherichia coli mutY* gene encodes an adenine glycosylase active on G-A mispairs. *Proc. Natl. Acad. Sci. USA.* **86**, 8877-8881.

### **-B-**

- Bailly, V., Verly, W.G., O'Connor, T.R., and Laval, J. (1989). Mechanism of DNA nicking at apurinic/apyrimidinic sites by *Escherichia coli* [formamidopyrimidine]DNA glycosylase. *J. Biochem.* **262**, 581-589.
- (a) Banerjee, A., and Verdine, G.L. (2006). A nucleobase lesion remodels the interaction of its normal neighbor in a DNA glycosylase complex. *Proc. Natl. Acad. Sci. USA.* **103**, 15020-15025.
- (b) Banerjee, A., Santos, W.L., and Verdine, G.L. (2006). Structure of a DNA glycosylase searching for lesions. *Science* **311**, 1153-1157.
- Banerjee, A., Yang, W., Karplus, M., and Verdine, G.L. (2005). Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* **434**, 612-618.
- Bebenek, K., Kunkel, T.A. (2004). Functions of DNA polymerases. *Adv. Prot. Chem.* **69**, 137-165.
- Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929-6948.

- Blackburn, C.M., and Gait, M.J. (1996). Nucleic acids in chemistry and biology. *Oxford University Press New York*.
- Bogenhagen, D.F., Pinz, K.G., and Perez-Jannotti, R.M. (2001). Enzymology of mitochondrial base excision repair. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 257-271.
- Boiteux, S., and Guillet, M. (2004). Abasic sites in DNA: repair and biological consequences in *Saccharomyces cerevisiae*. *DNA Repair (Amsterdam)* **3**, 1-12.
- Boiteux, S., and Laval, J. (1982). Coding properties of poly(deoxycytidylic acid) templates containing uracil or apyrimidinic sites: in vitro modulation of mutagenesis by deoxyribonucleic acid repair enzymes. *Biochemistry* **21**, 6746-6751.
- Boiteux, S., and Laval, J. (1983). Imidazole open ring 7-methylguanine : An inhibitor of DNA synthesis. *Biochem. Biophys. Res. Com.* **110**, 552-558.
- Boiteux, S., Bichara, M., Fuchs, R.P.P., and Laval, J. (1989). Excision of the imidazole ring-opened form of *N*-2-aminofluorene-C(8)-guanine adduct in poly(dG-dC) by *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Carcinogenesis* **10**, 1905-1909.
- Boiteux, S., Gajewski, E., Laval, J., and Dizdaroglu, M. (1992). Substrate specificity of the *Escherichia coli* Fpg protein (Formamidopyrimidine-DNA glycosylase): Excision of purine lesions in DNA produced by ionizing radiation or photosensitization. *Biochemistry* **31**, 106-110.
- Boiteux, S., O'Connor, T.R., and Laval, J. (1987). Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the *fpg* structural gene and overproduction of the protein. *EMBO J.* **6**, 3177-3183.
- Boiteux, S., O'Connor, T.R., Lederer, F., Gouyette, A., and Laval, J. (1990). Homogeneous *Escherichia coli* FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites. *J. Biol. Chem.* **265**, 3916-3922.
- Boulton, S.J., and Jackson, S.P. (1996). *Saccharomyces cerevisiae* Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways. *EMBO J.* **15**, 5093-5103.
- Bourdat, A.-G., Gasparutto, D., and Cadet, J. (1999). Synthesis and enzymatic processing of oligonucléotides containing tandem base damage. *Nucleic Acids Res.* **27**, 1015-1024.
- Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
- Breimer, L.H., and Lindahl, T. (1984). DNA glycosylase activities for thymine residues damaged by ring saturation, fragmentation, or ring contraction are functions of Endonuclease III in *Escherichia coli*. *J. Biol. Chem.* **259**, 5543-5548.
- Brotzel, F., and Mayr, H. (2007). Nucleophilicities of amino acids and peptides. *Org. Biomol. Chem.* **5**, 3814-3820.
- Bruner, S.D., Nash, H.M., Lane, W.S., and Verdine, G.L. (1998). Repair of oxidatively damaged guanine in *Saccharomyces cerevisiae* by an alternative pathway. *Curr. Biol.* **8**, 393-403.
- Bruner, S.D., Norman, D.P.G., and Verdine, G.L. (2000). Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature* **403**, 859-866.
- Budman, J., and Chu, G. (2005). Processing of DNA for nonhomologous end-joining by cell-free extract. *EMBO J.* **24**, 849-860.
- Burgers, P.M.J., Koonin, E.V., Bruford, E., Blanco, L., Burtis, K.C., Christman, M.F., Copeland, W.C., Friedberg, E.C., Hanaoka, F., Hinkle, D.C., Lawrence, C.W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C-A., Sugino, A.,

- Todo, T., Wang, Z., Weill, J.-C., and Woodgate, R. (2001). Eukaryotic DNA polymerases: Proposal for a revised nomenclature. *J. Biol. Chem.* **276**, 43487-43490.
- Burkovics, P., Szukacsov, V., Unk, I., and Haracska, L. (2006). Human Ape2 protein has a 3'-5' exonuclease activity that acts preferentially on mismatched base pairs. *Nucleic Acids Res.* **34**, 2508-2515.

-C-

- Cabrera, M., Nghiem, Y., and Miller, J.H. (1988). *mutM*, a second mutator locus in *Escherichia coli* that generates G•C → T•A transversions. *J. Bacteriol.* **170**, 5405-5407.
- Cadet, J., Douki, T., Frelon, S., Sauvaigo, S., Pouget, J.-P., and Ravanat, J.-L. (2002). Assessment of oxidative base damage to isolated and cellular DNA by HPLC-MS/MS measurement. *Free Radic. Biol. Med.* **33**, 441-449.
- Cadet, J., Douki, T., Gasparutto, D., and Ravanat, J.-L. (2003). Oxidative damage to DNA: formation, measurement and biochemical features. *Mutation Research* **531**, 5-23.
- Cadet, J., and Teoule, R. (1978). Comparative study of oxidation of nucleic acid components by hydroxyl radicals, singlet oxygen and superoxide anion radicals. *Photochem. Photobiol.* **28**, 661-665.
- Castaing, B., Boiteux, S., and Zelwer, C. (1992). DNA containing a chemically reduced apurinic site is a high affinity ligand for the *E. coli* formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* **20**, 389-394.
- Castaing, B., Fourrey, J.-L., Hervouet, N., Thomas, M., Boiteux, S., and Zelwer, C. (1999). AP site structural determinants for Fpg specific recognition. *Nucleic Acids Res.* **27**, 608-615.
- Castaing, B., Geiger, A., Seliger, H., Nehls, P., Laval, J., Zelwer, C., and Boiteux, S. (1993). Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant Fpg proteins. *Nucleic Acids Res.* **21**, 2899-2905.
- Chen, J.-H., Hales, C.N., and Ozanne, S.E. (2007). DNA damage, cellular senescence and organismal ageing: causal or correlative? *Nucleic Acids Research* **35**, 7417-7428.
- Chepanoske, C.L., Porello, S.L., Fujiwara, T., Sugiyama, H., and David, S.S. (1999). Substrate recognition by *Escherichia coli* MutY using substrate analogs. *Nucleic Acids Res.* **27**, 3197-3204.
- Chetsanga, C.J., and Frenette, G.P. (1983). Excision of aflatoxin B<sub>1</sub>-imidazole ring opened guanine adducts from DNA by formamidopyrimidine-DNA glycosylase. *Carcinogenesis* **4**, 997-1000.
- Chestanga, C.J., and Lindahl, T. (1979). Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. *Nucleic Acids Res.* **6**, 3673-3683.
- Chetsanga, C.J., Polidori, G., and Mainwaring, M. (1982). Analysis and excision of ring-opened phosphoramidate mustard-deoxyguanine adducts in DNA. *Cancer Res.* **42**, 2616-2621.
- Chung, S.J., and Verdine, G.L. (2004). Structures of end products resulting from lesion processing by a DNA glycosylase/lyase. *Chemistry & Biology* **11**, 1643-1649.
- Costa de Oliveira, R., Ribeiro, D.T., Nigro, R.G., Di Mascio, P., and Menck, C.F.M. (1992). Singlet oxygen induced mutation spectrum in mammalian cells. *Nucleic Acids Research* **20**, 4319-4323.
- Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C., and Castaing, B. (2004). Structural basis for the recognition of the FapydG lesion (2,6-diamino-4-hydroxy-5-

- formamidopyrimidine) by Formamidopyrimidine-DNA Glycosylase. *J. Biol. Chem.* **279**, 44074-44083.
- Coste, F., Ober, M., Le Bihan, Y.-V., Izquierdo, M.A., Hervouet, N., Mueller, H., Carell, T., and Castaing, B. (2008). Bacterial base excision repair enzyme Fpg recognizes bulky N7-substituted-FapydG lesion via unproductive binding mode. *Chemistry & Biology* **15**, 706-717.
  - Culp, S.J., Cho, B.P., Kadlubar, F.F., and Evans, F.E. (1989). Structural and conformational analyses of 8-Hydroxy-2'-deoxyguanosine. *Chem. Res. Toxicol.* **2**, 416-422.

#### -D-

- Dalhus, B., Helle, I.H., Backe, P.H., Alseth, I., Rognes, T., Bjørås, M., and Laerdahl, J.K. (2007). Structural insight into repair of alkylated DNA by a new superfamily of DNA glycosylases comprising HEAT-like repeats. *Nucleic Acids Res.* **35**, 2451-2459.
- Damia, G., and D'Incalci, M. (2007). Targeting DNA repair as a promising approach in cancer therapy. *Eur. J. Cancer* **43**, 1791-1801.
- David, S.S., and Williams, S.D. (1998). Chemistry of glycosylases and endonucleases involved in Base-Excision Repair. *Chem. Rev.* **98**, 1221-1261.
- De Bont, R., and van Larebeke, N. (2004). Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* **19**, 169-185.
- Delaney, M. O., Wiederholt, C. J., and Greenberg, M. M. (2002). Fapy-dA induces nucleotide misincorporation translesionally by a DNA polymerase. *Angew. Chem. Int. Ed. Engl.* **41**, 771-775.
- Demple, B., and DeMott, M.S. (2002). Dynamics and diversions in base excision DNA repair of oxidized abasic lesions. *Oncogene* **21**, 8926-8934.
- Dedon, P.C., Jiang, Z.W., Goldberg, I.H. (1992). Neocarzinostatin-mediated DNA damage in a model AGT.ACT site: mechanistic studies of thiol-sensitive partitioning of C4' DNA damage products. *Biochemistry* **31**, 1917-1927.
- DeLano, W.L. (2002). The PyMOL molecular graphics system. *DeLano Scientific, San Carlos, CA, USA* (<http://www.pymol.org>).
- Deng, L., Schäfer, O.D., and Verdine, G.L. (1997). Unusually strong binding of a designed transition-state analog to a base-excision DNA Repair Protein. *J. Am. Chem. Soc.* **119**, 7865-7866.
- De Souza-Pinto, N.C., Wilson III, D.M., Stevnsner, T.V., and Bohr, V.A. (2008). Mitochondrial DNA, base excision repair and neurodegeneration. *DNA Repair (Amst)*. **7**, 1098-1109.
- D'Ham, C., Romieu, A., Jaquinod, M., Gasparutto, D., and Cadet, J. (1999). Excision of 5,6-dihydroxy-5,6-dihydrothymine, 5,6-dihydrothymine, and 5-hydroxycytosine from defined sequence oligonucleotides by *Escherichia coli* Endonuclease III and Fpg proteins: Kinetic and mechanistic aspects. *Biochemistry* **38**, 3335-3344.
- Dianov, G.L., Souza-Pinto, N., Nyaga, S.G., Thybo, T., Stevnsner, T., and Bohr, V.A. (2001). Base excision repair in nuclear and mitochondrial DNA. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 285-297.
- Diaz, M., and Lawrence, C. (2005). An update on the role of translesion synthesis DNA polymerases in Ig hypermutation. *Trends Immunol.* **26**, 216-220.
- Diederichs, K., and Karplus, P.A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* **4**, 269-275.
- Dizdaroglu, M. (2005). Base-excision repair of oxidative DNA damage by DNA glycosylases. *Mutation Research* **591**, 45-59.

- Dizdaroglu, M., Kirkali, G., and Jaruga, P. (2008). Formamidopyrimidines in DNA: Mechanisms of formation, repair, and biological effects. *Free Radic. Biol. Med.* **45**, 1610-1621.
- Dizdaroglu, M., Laval, J., and Boiteux, S. (1993). Substrate specificity of the *Escherichia coli* endonuclease III: Excision of thymine- and cytosine-derived lesions in DNA produced by radiation-generated free radicals. *Biochemistry* **32**, 12105-12111.
- Dou, H., Theriot, C.A., Das, A., Hegde, M.L., Matsumoto, Y., Boldogh, I., Hazra, T.K., Bhakat, K.K., and Mitra, S. (2008). Interaction of the human DNA glycosylase NEIL1 with Proliferating Cell Nuclear Antigen. The potential for replication-associated repair of oxidized bases in mammalian genomes. *J. Biol. Chem.* **283**, 3130-3140.
- Dronkert, M.L.G., and Kanaar, R. (2001). Repair of DNA interstrand cross-links. *Mutation Research* **486**, 217-247.
- Duarte, V., Gasparutto, D., Jaquinod, M., and Cadet, J. (2000). *In vitro* DNA synthesis opposite oxazolone and repair of this DNA damage using modified oligonucleotides. *Nucleic Acids Res.* **28**, 1555-1563.
- Duarte, V., Gasparutto, D., Jaquinod, M., Ravanat, J., and Cadet, J. (2001). Repair and mutagenic potential of oxaluric acid, a major product of singlet oxygen-mediated oxidation of 8-oxo-7,8-dihydroguanine. *Chem. Res. Toxicol.* **14**, 46-53.
- Duncan, B.K., Weiss, B. (1982). Specific mutator effects of ung (uracil-DNA glycosylase) mutations in *Escherichia coli*. *J. Bacteriol.* **151**, 750-755.

#### -E-

- Espejel, S., Franco, S., Rodriguez-Perales, S., Bouffler, S.D., Cigudosa, J.C., and Blasco, M.A. (2002). Mammalian Ku86 mediates chromosomal fusions and apoptosis caused by critically short telomeres. *EMBO J.* **21**, 2207-2219.
- Evans, M.D., Dizdaroglu, M., and Cooke, M.S. (2004). Oxidative DNA damage and disease: induction, repair and significance. *Mutation Research* **567**, 1-61.
- Evans, J., Maccabee, M., Hatahet, Z., Courcelle, J., Bockrath, R., Ide, H., and Wallace S. (1993). Thymine ring saturation and fragmentation products: lesion bypass, misinsertion and implications for mutagenesis. *Mutat. Res.* **299**, 147-156.

#### -F-

- Francis, A.W., and David, S.S. (2003). *Escherichia coli* MutY and Fpg utilize a processive mechanism for target location. *Biochemistry* **42**, 801-810.
- Freisinger, E., Grollman, A.P., Miller, H., and Kisker, C. (2004) Lesion (in)tolerance reveals insights into DNA replication fidelity. *EMBO J.* **23**, 1494-1505.
- French, G.S., and Wilson, K.S. (1978). On the treatment of negative intensity observations. *Acta. Cryst.* **A34**, 517-525.
- Frit, P., Kwon, K., Coin, F., Auriol, J., Dubaele, S., Salles, B., and Egly, J.-M. (2002). Transcriptional Activators Stimulate DNA Repair. *Molecular Cell* **10**, 1391-1401.
- Fromme, J.C., Banerjee, A., and Verdine, G.L. (2004). DNA glycosylase recognition and catalysis. *Curr. Opin. Struct. Biol.* **14**, 43-49.
- Fromme, J.C., and Verdine, G.L. (2002). Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat. Struct. Biol.* **9**, 544-552.
- (a) Fromme, J.C., and Verdine, G.L. (2003). DNA lesion recognition by the bacterial repair enzyme MutM. *J. Biol. Chem.* **278**, 51543-51548.
- (b) Fromme, J.C., and Verdine, G.L. (2003). Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J.* **22**, 3461-3471.

- Frosina, G. (2006). Prophylaxis of oxidative DNA damage by formamidopyrimidine-DNA glycosylase. *Int. J. Cancer* **119**, 1-7.
- Fukae, J., Mizuno, Y., and Hattori, N. (2007). Mitochondrial dysfunction in Parkinson's disease. *Mitochondrion* **7**, 58-62.
- Furge, L.L., and Guengerich, F.P. (1997). Analysis of nucleotide insertion and extension at 8-oxo-7,8-dihydroguanine by replicative T7 polymerase exo(-) and human immunodeficiency virus-1 reverse transcriptase using steady-state and pre-steady-state kinetics. *Biochemistry* **36**, 6475-6487.
- Fuxreiter, M., Luo, N., Jedlovsky, P., Simon, I., and Osman, R. (2002). Role of base flipping in specific recognition of damaged DNA by repair enzymes. *J. Mol. Biol.* **323**, 823-834.

### -G-

- Gasparutto, D., Muller, E., Boiteux, S., and Cadet, J. (2009). Excision of the oxidatively formed 5-hydroxyhydantoin and 5-hydroxy-5-methylhydantoin pyrimidine lesions by *Escherichia coli* and *Saccharomyces cerevisiae* DNA N-glycosylases. *Biochimica et Biophysica Acta* **1790**, 16-24.
- Gasparutto, D., Ait-Abbas, M., Jaquinod, M., Boiteux, S., and Cadet, J. (2000). Repair and coding properties of 5-Hydroxy-5-methylhydantoin nucleosides inserted into DNA oligomers. *Chem. Res. Toxicol.* **13**, 575-584.
- Gilboa, R., Zharkov, D.O., Golan, G., Fernandes, A.S., Gerchman, S.E., Matz, E., Kycia, J.H., Grollman, A.P., and Shoham, G. (2002). Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA. *J. Biol. Chem.* **277**, 19811-19816.
- Graziewicz, M. A., Zastawny, T. H., Olinski, R., Speina, E., Siedlecki, J., and Tudek, B. (2000). Fapyadenine is a moderately efficient chain terminator for prokaryotic DNA polymerases. *Free Radic. Biol. Med.* **28**, 75-83.
- Golan, G., Zharkov, G.O., Grollman, A.P., Dodson, M.L., McCullough, A.K., Lloyd, R.S., and Shoham, G. (2006). Structure of T4 pyrimidine dimer glycosylase in a reduced imine covalent complex with abasic site-containing DNA. *J. Mol. Biol.* **362**, 241-258.
- Goodman, M. F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu. Rev. Biochem.* **71**, 17-50.
- Gopalakrishnan, K., Sowmiya, G., Sheik, S.S., and Sekar, K. (2007). RP(2.0): Ramachandran plot on the web. *Protein Peptide Letters* **14**, 669-671.
- Gorbunova, V., Seluanov, A., Mao, Z., and Hine, C. (2007). Changes in DNA repair during aging. *Nucleic Acids Research* **35**, 7466-7474.
- Goswami, S.C., Yoon, J.H., Abramczyk, B.M., Pfeifer, G.P., and Postel, E.H. (2006). Molecular and functional interactions between *Escherichia coli* nucleoside-diphosphate kinase and the uracil-DNA glycosylase Ung. *J. Biol. Chem.* **281**, 32131-32139.
- Gouet, P., Courcelle, E., Stuart, D.I., and Metz, F. (1999). ESPript : analysis of multiple sequence alignments in Postscript, *Bioinformatics* **15**, 305-308.
- (a) Graves, R.J., Felzenszwalb, I., Laval, J., and O'Connor, T.R. (1992). Excision of 5'-terminal deoxyribose phosphate from damaged DNA is catalyzed by the Fpg protein of *Escherichia coli*. *J. Biol. Chem.* **267**, 14429-14435.
- (b) Graves, R., Laval, J., and Pegg, A.E. (1992). Sequence specificity of DNA repair by *Escherichia coli* Fpg protein. *Carcinogenesis* **13**, 1455-1459.
- Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.

## -H-

- Halliwell, B., and Aruoma, O.I. (1991). DNA damage by oxygen-derived species. Its mechanism and measurement in mammalian systems. *FEBS letters* **281**, 9-19.
- Haraguchi, K., and Greenberg, M.M. (2001). Synthesis of oligonucleotides containing Fapy•dG (N6-(2-Deoxy- $\alpha,\beta$ -D-erythro-pentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine). *J. Am. Chem. Soc.* **123**, 8636-8637.
- Haraguchi, K., Delaney, M.O., Wiederholt, C.J., Sambandam, A., Hantosi, Z., and Greenberg, M.M. (2002). Synthesis and characterization of oligodeoxynucleotides containing Formamidopyrimidine lesions and nonhydrolyzable analogues. *J. Am. Chem. Soc.* **124**, 3263-3269.
- Harbut, M.B., Meador, M., Dodson, M.L. and Lloyd, R.S. (2006). Modulation of the turnover of Formamidopyrimidine DNA glycosylase. *Biochemistry* **45**, 7341-7346.
- Hashimoto, M., Greenberg, M.M., Kow, Y.W., Hwang, J.-T., and Cunningham, R.P. (2001). The 2-deoxyribonolactone lesion produced in DNA by neocarzinostatin and other damaging agents forms cross-links with the base-excision repair enzyme Endonuclease III. *J. Am. Chem. Soc.* **123**, 3161-3162.
- Hatahet, Z., Kow, Y.W., Purmal, A.A., Cunningham, R.P., and Wallace, S.S. (1994). New substrates for old enzymes. 5-hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for *Escherichia coli* endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is substrate for uracil DNA N-glycosylase. *J. Biol. Chem.* **269**, 18814-18820.
- Haushalter, K.A., Stukenberg, P.T., Kirschner, M.W., and Verdine, G.L. (1999). Identification of a new uracil-DNA glycosylase family by expression cloning using synthetic inhibitors. *Curr. Biol.* **9**, 174-185.
- Hayakawa, H., Taketomi, A., Sakumi, K., Kuwano, M., and Sekiguchi, M. (1995). Generation and elimination of 8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-triphosphate, a mutagenic substrate for DNA synthesis, in human cells. *Biochemistry* **34**, 89-95.
- Hazra, T.K., Hill, J.W., Izumi, T., and Mitra, S. (2001). Multiple DNA glycosylases for repair of 8-Oxoguanine and their potential *in vivo* functions. *Prog. Nucleic Acid Res.* **68**, 193-205.
- Hazra, T.K., Izumi, T., Mardt, L., Floyd, R.A., and Mitra, S. (1998). The presence of two distinct 8-oxoguanine repair enzymes in human cells: their potential complementary roles in preventing mutation. *Nucleic Acids Res.* **26**, 5116-5122.
- Hazra, T.K., Izumi, T., Venkataraman, R., Kow, Y.W., Dizdaroglu, M., and Mitra, S. (2000). Characterization of a novel 8-Oxoguanine-DNA glycosylase activity in *Escherichia coli* and identification of the enzyme as Endonuclease VIII. *J. Biol. Chem.* **275**, 27762-27767.
- Hazra, T.K., and Mitra, S. (2006). Purification and characterization of NEIL1 and NEIL2, members of a distinct family of mammalian DNA glycosylases for repair of oxidized bases. *Methods Enzymol.* **408**, 33-48.
- Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000). DNA bending and a flip-out mechanism for base excision by the helix- hairpin-helix DNA glycosylase, *Escherichia coli* AlkA. *EMBO J.* **19**, 758-766.
- Hsu, G.W., Ober, M., Carell, T., and Beese1, L.S. (2004) Error-prone replication of oxidatively damaged DNA by a high-fidelity DNA polymerase. *Nature* **431**, 217-221.
- Huffman, J.L., Sundheim, O., and Tainer, J.A. (2005). DNA base damage recognition and removal: New twists and grooves. *Mutation Research* **577**, 55-76.

-I-

- Imoto, S., Patro, J.N., Jiang, Y.L., Oka, N., and Greenberg, M.M. (2006). Synthesis, DNA polymerase incorporation, and enzymatic phosphate hydrolysis of Formamidopyrimidine nucleoside triphosphates. *J. Am. Chem. Soc.* **128**, 14606-14611.

-J-

- Jurado, J., Sapparbaev, M., Matray, T.J., Greenberg, M.M., and Laval, J. (1998). The ring fragmentation product of thymidine C5-hydrate when present in DNA is repaired by the *Escherichia coli* Fpg and Nth proteins. *Biochemistry* **37**, 7757-7763.

-K-

- Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637.
- Kalam, M. A., Haraguchi, K., Chandani, S., Loechler, E. L., Moriya, M., Greenberg, M. M., and Basu, A. K. (2006). Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res.* **34**, 2305-2315.
- Karakaya, A., Jaruga, P., Bohr, V.A., Grollman, A.P., and Dizdaroglu, M. (1997). Kinetics of excision of purine lesions from DNA by *Escherichia coli* Fpg protein. *Nucleic Acids Res.* **25**, 474-479.
- Katafuchi, A., Nakano, T., Masaoka, A., Terato, H., Iwai, S., Hanaoka, F., and Ide, H. (2004). Differential specificity of human and *Escherichia coli* Endonuclease III and VIII homologues for oxidative base lesions. *J. Biol. Chem.* **279**, 14464-14471.
- Kino, K., and Sugiyama, H. (2005). UVR-induced G-C to C-G transversions from oxidative DNA damage. *Mutation Research* **571**, 33-42.
- Kleywegt, G.J., Harris, M.R., Zou, J.Y., Taylor, T.C., Wahlby, A., and Jones, T.A. (2004). The uppsala electron density server. *Acta. Cryst.* **D60**, 2240-2249.
- Korniyushyna, O., Berges, A.M., Muller, J.G., and Burrows C.J. (2002). In vitro nucleotide misinsertion opposite the oxidized guanosine lesions Spiroiminodihydantoin and Guanidinohydantoin and DNA synthesis past the lesions using *Escherichia coli* DNA Polymerase I (Klenow fragment). *Biochemistry* **41**, 15304-15314.
- Kouchakdjian, M., Bodepudi, V., Shibusaki, S., Eisenberg, M., Johnson, F., Grollman, A.P., and Patel, D.J. (1991). NMR structural studies of the ionizing radiation adduct 7-Hydro-8-oxodeoxyguanosine (8-oxo-7H-dG) opposite deoxyadenosine in a DNA duplex. 8-oxo-7H-dG(syn)•dA(anti) alignment at lesion site. *Biochemistry* **30**, 1403-1412.
- Kroeger, K.M., Hashimoto, M., Kow, Y.W., and Greenberg, M.M. (2003). Cross-linking of 2-deoxyribonolactone and its  $\beta$ -elimination product by base excision repair enzymes. *Biochemistry* **42**, 2449-2455.
- Krwawicz, J., Arczewska, K.D., Speina, E., Maciejewska, A., and Grzesiuk, E. (2007). Bacterial DNA repair genes and their eukaryotic homologues: 1. Mutations in genes involved in base excision repair (BER) and DNA-end processors and their implication in mutagenesis and human disease. *Acta Biochimica Polnica* **54**, 413-434.
- Kuchino, Y., Mori, F., Kasai, H., Inoue, H., Iwai, S., Miura, K., Ohtsuka, K., and Nishimura, S. (1987). Misreading of DNA templates containing 8-

hydroxydeoxyguanosine at the modified base and at adjacent residues. *Nature* **327**, 77-79.

-L-

- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK : a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291.
- Lau, A.Y., Schäfer, O.D., Samson, L., Verdine, G.L., and Ellenberger, T. (1998). Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: Mechanisms for nucleotide flipping and base excision. *Cell* **95**, 249-258.
- Lavrukhin, O.V., and Lloyd, R.S. (2000). Involvement of phylogenetically conserved acidic amino acid residues in catalysis by an oxidative DNA damage enzyme Formamidopyrimidine glycosylase. *Biochemistry* **39**, 15266-15271.
- LeDoux, S.P., and Wilson, G.L. (2001). Base excision repair of mitochondrial DNA damage in mammalian cells. *Prog. Nucleic Acid Res. Mol. Biol.* **68**, 273-284.
- Lee, S., Radom, C.T., and Verdine, G.L. (2008). Trapping and structural elucidation of a very advanced intermediate in the lesion-extrusion pathway of hOGG1. *J. Am. Chem. Soc.* **130**, 7784-7785.
- Leipold, M.D., Muller, J.G., Burrows, C.J., and David, S.S. (2000). Removal of Hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, Fpg. *Biochemistry* **39**, 14984-14992.
- Leonard, G.A., McAuley-Hecht, K.E., Gibson, N.J., Brown, T., Watson, W.P., and Hunter, W.N. (1994) Guanine-1,N6-ethenoadenine base pairs in the crystal structure of d(CGCGAATT(edA)GCG). *Biochemistry* **33**, 4755-4761.
- Leppard, J.B., Champoux, J.J. (2005). Human DNA topoisomerase I: relaxation, roles, and damage control. *Chromosoma* **114**, 75-85.
- Leslie, A.G.W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**.
- Li, G.-M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Research* **18**, 85-98.
- Li, Q., Laval, J., and Ludlum, D.B. (1997). Fpg protein releases a ring-opened N-7 guanine adduct from DNA that has been modified by sulfur mustard., *Carcinogenesis* **18**, 1035-1038.
- Lindahl, T. (1974). An N-glycosidase from *Escherichia coli* that releases free Uracil from DNA containing deaminated Cytosine residues. *Proc. Natl. Acad. Sci. USA.* **71**, 3649-3653.
- Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715.
- Lipscomb, L.H., Peek, M.E., Morningstar, M.L., Verghis, S.M., Miller, E.M., Richs, A., Essigmann, J.M., and Williams, L.D. (1995). X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc. Natl. Acad. Sci. USA.* **92**, 719-723.
- Lopez, B.S. (2006). La recombinaison homologue: l'équilibre entre la stabilité, la diversité et l'instabilité génétique. *Biofutur* **271**, 22-26.
- Lovell, M.A., and Markesbery, W.R. (2007). Oxidative DNA damage in mild cognitive impairment and late-stage Alzheimer's disease. *Nucleic Acids Research* **35**, 7497-7504.
- Lowe, L.G., and Guengerich, F.P. (1996). Steady-State and Pre-Steady-State Kinetic Analysis of dNTP Insertion Opposite 8-Oxo-7,8-dihydroguanine by *Escherichia coli* Polymerases I  $\text{exo}^-$  and II  $\text{exo}^-$ . *Biochemistry* **35**, 9840-9849.

**-M-**

- Maki, H., and Sekiguchi, M. (1992). MutT protein specifically hydrolyzes a potent mutagenic substrate for DNA synthesis. *Nature (London)* **355**, 273-275.
- Mao, G., Pan, X., Zhu, B.-B., Zhang, Y., Yuan, F., Huang, J., Lovell, M.A., Lee, M.P., Markesbery, W.R., Li, G.-M., and Gu, L. (2007). Identification and characterization of OGG1 mutations in patients with Alzheimer's disease. *Nucleic Acids Research* **35**, 2759-2766.
- Martomo, S.A., and Gearhart, P.J. (2006). Somatic hypermutation: subverted DNA repair. *Curr. Opin. Immunol.* **18**, 243-248.
- Matsumoto, Y., Zhang, Q.M., Takao, M., Yasui, A., and Yonei, S. (2001). Escherichia coli Nth and human hNTH1 DNA glycosylases are involved in removal of 8-oxoguanine from 8-oxoguanine/guanine mispairs in DNA. *Nucleic Acids Res.* **29**, 1975-1981.
- McAuley-Hecht, K.E., Leonard, G.A., Gibson, N.J., Thomson, J.B., Watson, W.P., Hunter, W.N., and Brown, T. (1994). Crystal structure of a DNA duplex containing 8-Hydroxydeoxyguanine-adenine base pairs. *Biochemistry* **33**, 10266-10270.
- McBride, T.J., Preston, B.D., and Loeb, L.A. (1991). Mutagenic spectrum resulting from DNA damage by oxygen radicals. *Biochemistry* **30**, 207-213.
- McHugh, P.J., Knowland, J. (1995). Novel reagents for chemical cleavage at abasic sites and UV photoproducts in DNA. *Nucleic Acids Res.* **23**, 1664-1670.
- McMurray, C.T. (2008). Hijacking of the mismatch repair system to cause CAG expansion and cell death in neurodegenerative disease. *DNA repair* **7**, 1121-1134.
- (a) Michaels, M.L., Cruz, C., Grollman, A.P., and Miller, J.H. (1992). Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. *Proc. Natl. Acad. Sci. USA.* **89**, 7022-7025.
- (b) Michaels, M.L., and Miller, J.H. (1992). The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-Hydroxyguanine (7,8-Dihydro-8-Oxoguanine). *J. Bacteriol.* **174**, 6321-6325.
- (c) Michaels, M.L., Tchou, J., Grollman, A.P., and Miller, J.H. (1992). A repair system for 8-Oxo-7,8-dihydrodeoxyguanine. *Biochemistry* **31**, 10964-10968.
- Miller, H., Prasad, R., Wilson, S.H., Johnson, F., and Grollman, A.P. (2000). 8-oxodGTP incorporation by DNA polymerase beta is modified by active-site residue Asn279. *Biochemistry* **39**, 1029-1033.
- Mo, J-Y., Maki, H., and Sekiguchi, M. (1992). Hydrolytic elimination of a mutagenic nucleotide, 8-oxodGTP, by human 18-kilodalton protein: Sanitization of nucleotide pool. *Proc. Natl. Acad. Sci. USA.* **89**, 11021-11025.
- Moore, J.K., and Haber, J.E. (1996). Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 2164-2173.
- Mortusewicz, O., Rothbauer, U., Cardosol, M.C., and Leonhardt, H. (2006). Differential recruitment of DNA Ligase I and III to DNA repair sites. *Nucleic Acids Res.* **34**, 3523-3532.
- Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Cryst.* **D53**, 240-255.

-N-

- Nakano, T., Terato, H., Asagoshi, K., Masaoka, A., Mukuta, M., Ohyama, Y., Suzuki, T., Makino, K and Ide, H. (2003). DNA-protein cross-link formation mediated by oxanine. A novel genotoxic mechanism of nitric oxide-induced DNA damage. *J. Biol. Chem.* **278**, 25264-25272.
- Navaza, J. (1994). AMoRe : an automated package for molecular replacement. *Acta. Cryst.* **A50**, 157-163.
- Neuberger, M.S., Di Noia, J.M., Beale, R.C., Williams, G.T., Yang, Z., and Rada, C. (2005). Somatic hypermutation at A•T pairs: polymerase error versus dUTP incorporation. *Nat. Rev. Immunol.* **5**, 171-178.
- Neuberger, M.S., Harris, R.S., Di Noia, J., and Petersen-Mahrt, S.K. (2003). Immunity through DNA deamination. *Trends. Biochem. Sci.* **28**, 305–312.
- Nghiem, Y., Cabrera, M., Cupples, C.G., and Miller, J.H. (1988). The *mutY* gene: A mutator locus in *Escherichia coli* that generates G•C → T•A transversions. *Proc. Natl. Acad. Sci. USA.* **85**, 2709-2713.
- Nimesh, J., Duppatla, V., and Rao, D.N. (2006). Prokaryotic DNA mismatch repair. *Prog. Nucleic Acid Res. Mol. Biol.* **81**, 1-49.
- Nospikel, T. (2008). Nucleotide excision repair and neurological diseases. *DNA Repair* **7**, 1155-1167.

-O-

- Ober, M., Linne, U., Gierlich, J., and Carell, T. (2003). The two main oxidative DNA lesions 8-oxo-7,8-dihydroguanine and 2,6-diamino-5-formamido-4-hydroxypyrimidine exhibit strongly different pairing characteristics. *Angew. Chem. Int. Ed. Engl.* **42**, 4947-4951.
- Ober, M., Müller, H., Pieck, C., Gierlich, J., and Carell, T. (2005). Base pairing and replicative processing of the Formamidopyrimidine-dG DNA lesion. *J. Am. Chem. Soc.* **127**, 18143-18149.
- O'Connor, T.R., and Laval, J. (1989). Physical association of the 2,6-diamino-4-hydroxy-5N-formamidopyrimidine-DNA glycosylase of *Escherichia coli* and an activity nicking DNA at apurinic/apyrimidinic sites. *Proc. Natl. Acad. Sci. USA* **86**, 5222-5226.
- O'Connor, T.R., Boiteux, S., and Laval, J. (1988). Ring-opened 7-methylguanine residues in DNA are a block to *in vitro* DNA synthesis. *Nucleic Acids Research* **16**, 5879-5894.
- Oda, Y., Uesugi, S., Ikehara, M., Nishimura, S., Kawase, Y., Ishikawa, H., Inoue, H., and Ohtsuka, E. (1991). NMR studies of a DNA containing 8-hydroxydeoxyguanosine. *Nucleic Acids Research* **19**, 1407-1412.
- Ohmori, H., Friedberg, E. C., Fuchs, R. P., Goodman, M. F., Hanaoka, F., Hinkle, D., Kunkel, T. A., Lawrence, C. W., Livneh, Z., Nohmi, T., Prakash, L., Prakash, S., Todo, T., Walker, G. C., Wang, Z., and Woodgate, R. (2001). The Y-family of DNA polymerases. *Mol. Cell* **8**, 7-8.
- Olinski, R., Gackowski, D., Foksinski, M., Rozalski, R., Roszkowski, K, and Jaruga, P. (2002). Oxidative DNA damage: assessment of the role in carcinogenesis, atherosclerosis, and acquired immunodeficiency syndrome. *Free Radic. Biol. Med.* **33**, 192-200.
- Olinski, R., Gackowski, D., Rozalski, R., Foksinski, M., and Bialkowski, K. (2003). Oxidative DNA damage in cancer patients: a cause or a consequence of the disease development? *Mutation Research* **531**, 177-190.

- Ono, T., Negishi, K., and Hayatsu, H. (1995). Spectra of superoxide-induced mutations in the *lacI* gene of a wild-type and a *mutM* strain of *Escherichia coli* K-12. *Mutation Research* **326**, 175-183.

### -P-

- Pagés, V., and Fuchs, R. P. (2002). How DNA lesions are turned into mutations within cells. *Oncogene* **21**, 8957-8966.
- Parikh, S.S., Mol, C.D., Slupphaug, G., Bharati, S., Krokan, H.E., and Tainer, J.A. (1998). Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J.* **17**, 5214-5226.
- Parikh, S.S., Walcher, G., Jones, G.D., Slupphaug, G., Krokan, H.E., Blackburn, G.M., and Tainer, J.A. (2000). Uracil-DNA glycosylase–DNA substrate and product structures: Conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl. Acad. Sci. USA.* **97**, 5083-5088.
- Parsons, J.L., Zharkov, D.O., and Dianov, G.L. (2005). NEIL1 excises 3' end proximal oxidative DNA lesions resistant to cleavage by NTH1 and OGG1. *Nucleic Acids Research* **33**, 4849-4856.
- Patro, J. N., Wiederholt, C. J., Jiang, Y. L., Delaney, J. C., Essigmann, J. M., and Greenberg, M. M. (2007). Studies on the replication of the ring opened Formamidopyrimidine, Fapy.dG in *Escherichia coli*. *Biochemistry* **46**, 10202-10212.
- Pereira de Jésus, K., Serre, L., Hervouet, N., Bouckson-Castaing, V., Zelwer, C., and Castaing, B. (2002). Crystallization and preliminary X-ray crystallographic studies of a complex between the *Lactococcus lactis* Fpg DNA-repair enzyme an abasic site containing DNA. *Acta. Cryst.* **D58**, 679-682.
- Pereira de Jésus, K., Serre, L., Zelwer, C., and Castaing, B. (2005). Structural insights into abasic site for Fpg specific binding and catalysis: comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogues-containing DNA. *Nucleic Acids Res.* **33**, 5936-5944.
- Perlow-Poehnelt, R.A., Zharkov, D.O., Grollman, A.P., and Broyde, S. (2004). Substrate discrimination by Formamidopyrimidine-DNA Glycosylase: Distinguishing interactions within the active site. *Biochemistry* **43**, 16092-16105.
- Pinz, K.G., Shibutani, S., and Bogenhagen, D.F. (1995). Action of mitochondrial DNA polymerase  $\gamma$  at sites of base loss or oxidative damage. *J. Biol. Chem.* **270**, 9202-9206.
- Pouget, J.-P., Douki, T., Richard, M.-J., and Cadet, J. (2000). DNA damage induced in Cells by  $\gamma$  and UVA radiation as measured by HPLC/GC-MS and HPLC-EC and comet assay. *Chem. Res. Toxicol.* **13**, 541-549.
- Purmal, A.A., Lampman, G.W., Bond, J.P., Hatahet, Z., and Wallace, S.S. (1998). Enzymatic processing of uracil glycol, a major oxidative product of DNA cytosine. *J. Biol. Chem.* **273**, 10026-10035.
- Prasad, R., Liu, Y., Deterding, L.J., Poltoratsky, V.P., Kedar, P.S., Horton, J.K., Kanno, S., Asagoshi, K., Hou, E.W., Khodyreva, S.N., Lavrik, O.I., Tomer, K.B., Yasui, A., and Wilson, S.H. (2007). HMGB1 is a cofactor in mammalian base excision repair. *Mol. Cell* **27**, 829-841.

### -R-

- Ramstein, J., and Lavery, R. (1988). Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci. USA.* **85**, 7231-7235.
- Read, R.J., and Schierbeek, A.J. (1988). A phased translation function. *J. Appl. Cryst.* **21**, 490-495.

- Roussel, A., and Cambillau, C. (1991). Turbo-Frodo. *Silicon Graphics Geometry (Ed. Silicon Graphics Mountain View, CA)*, 86-86.
- Rydberg, B., and Lindahl, T. (1982). Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* **1**, 211-216.

-S-

- Saenger, W. (1984). Principles of Nucleic Acid Structure. *Springer Advanced Texts in Chemistry*, Springer-Verlag, New York.
- Schärer, O.D., and Jiricny, J. (2001). Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *BioEssays* **23**, 270-281.
- Schärer, O.D., and Verdine, G.L. (1995). A designed inhibitor of base-excision repair. *J. Am. Chem. Soc.* **117**, 10781-10782.
- Schärer, O.D., Kawate, T., Gallinari, P., Jiricny, J., and Verdine, G.L. (1997). Investigation of the mechanisms of DNA binding of the human G/T glycosylase using designed inhibitors. *Proc. Natl. Acad. Sci. USA* **94**, 4878-4883.
- Schärer, O.D., Nash, H.M., Jiricny, J., Laval, J., and Verdine, G.L. (1998). Specific binding of a designed pyrrolidine abasic site analog to multiple DNA glycosylases. *J. Biol. Chem.* **273**, 8592-8597.
- Seibert, E., Ross, J.B.A, and Osman, R. (2002). Role of DNA flexibility in sequence-dependent activity of Uracil DNA Glycosylase. *Biochemistry* **41**, 10976-10984.
- Serre, L., Peirera de Jésus, K., Boiteux, S., Zelwer, C., and Castaing, B. (2002). Crystal structure of the *Lactococcus lactis* formamidopyrimidine-DNA glycosylase bound to an abasic site analog-containing DNA. *EMBO J.* **21**, 2854-2865.
- Shibutani, S., Takeshita, M., and Grollman, A.P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* **349**, 431-434.
- Sidorenko, V.S., Mechetin, G.V., Nevinsky, G.A, and Zharkov, D.O. (2008). Ionic strength and magnesium affect the specificity of *Escherichia coli* and human 8-oxoguanine-DNA glycosylases. *FEBS J.* **275**, 3747-3760.
- Sidorkina, O.M., and Laval, J. (2000). Role of the N-terminal proline residue in the catalytic activities of the *Escherichia coli* Fpg protein. *J. Biol. Chem.* **275**, 9924-9929.
- Sidorkina, O., Dizdaroglu, M., and Laval, J. (2001). Effect of single mutations on the specificity of *Escherichia coli* Fpg protein for excision of purine lesions from DNA damaged by free radicals. *Free Radic. Biol. Med.* **31**, 816-823.
- Simonelli, V., Narciso, L., Dogliotti, E., and Fortini, P. (2005). Base excision repair intermediates are mutagenic in mammalian cells. *Nucleic Acids Res.* **33**, 4404-4411.
- Simon, P., Gasparutto, D., Gambarelli, S., Saint-Pierre, C., Favier, A., and Cadet, J. (2006). Formation of isodialuric acid lesion within DNA oligomers via one-electron oxidation of 5-hydroxyuracil: characterization, stability and excision repair. *Nucleic Acids Res.* **34**, 3660-3669.
- Slupphaug, G., Mol, C.D., Kavli, B., Arvai, A.S., Krokan, H.E., and Tainer, J.A. (1996). A nucleotide flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA. *Nature* **384**, 87-92.
- Sobol, R.W., Kartalou, M., Almeida, K.H., Joyce, D.F., Engelward, B.P., Horton, J.K., Prasad, R., Samson, L.D., and Wilson, S.H. (2003). Base excision repair intermediates induce p53-independent cytotoxic and genotoxic responses. *J. Biol. Chem.* **278**, 39951-39959.
- Speina, E., Ciesla, J.M., Wójcik, J., Bajek, M., Kusmierk, J.T., and Tudek, B. (2001). The pyrimidine ring-opened derivative of 1,N<sup>6</sup>-ethenoadenine is excised from DNA by the *Escherichia coli* Fpg and Nth proteins. *J. Biol. Chem.* **276**, 21821-21827.

- Stivers, J.T., and Jiang, Y.L. (2003). A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.* **103**, 2729-2759.
- Sugahara, M., Mikawa, T., Kumasaka, T., Yamamoto, M., Kato, R., Fukuyama, K., Inoue, Y., and Kuramitsu, S. (2000). Crystal structure of a repair enzyme of oxidatively-damaged-DNA, MutM (Fpg), from an extreme thermophile, *Thermus Thermophilus* Hb8", *EMBO J.* **19**, 3857-3869.

#### -T-

- Tchou, J., and Grollman, A.P. (1995). The catalytic mechanism of Fpg protein. Evidence for Schiff base intermediate and amino terminus localization of the catalytic site. *J. Biol. Chem.* **269**, 15318-15324.
- Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A.P., and Johnson, F. (1994). Substrate Specificity of Fpg Protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem.* **269**, 15318-15324.
- Tchou, J., Kasai, H., Shibutani, S., Chung, M.-H., Laval, J., Grollman, A.P., and Nishimura, S. (1991). 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl. Acad. Sci. USA* **88**, 4690-4694.
- Thivyanathan, V., Somasunderam, A., Hazra, T.K., Mitra, S., and Gorenstein, D.G. (2003). Solution structure of a DNA duplex containing 8-hydroxy-2'-deoxyguanosine opposite deoxyguanosine. *J. Mol. Biol.* **325**, 433-442.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680.
- Tiwari, A., and Panigrahi, S.K. (2007). *HBAT*: A complete package for analysing strong and weak hydrogen bonds in macromolecular crystal structures. *In Silico Biology* **7**, 651-661.
- Tudek, B. (2003). Imidazole ring-opened DNA purines and their biological significance. *J. Biochem. Mol. Biol.* **36**, 12-19.
- Tudek, B., Boiteux, S., and Laval, J. (1992). Biological properties of imidazole ring-opened N7-methylguanine in M13mp18 phage DNA. *Nucleic Acids Research* **20**, 3079-3084.
- Tudek, B., Graziewicz, M. A., Kazanova, O., Zastawny, T. H., Obtulowicz, T., and Laval, J. (1999). Mutagenic specificity of imidazole ring-opened N7-methylguanine in M13mp18 phage DNA. *Acta Biochimica Polnica* **46**, 785-799.
- Tudek, B., VanZeeland, A.A., Kusmierek, J.T., and Laval, J. (1998). Activity of *Escherichia coli* DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. *Mutation Res.* **407**, 169-176.

#### -V-

- Vidal, A.E., Hickson, I.D., Boiteux, S., and Radicella, J.P. (2001). Mechanism of stimulation of the DAN glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step. *Nucleic Acids Research* **29**, 1285-1292.

-W-

- Wang, J.-S., and Groopman, J.D. (1999). DNA damage by mycotoxins. *Mutat. Res.* **424**, 167-181.
- Weiss, J.M., Goode, E.L., Ladiges, W.C., and Ulrich, C.M. (2005). Polymorphic variation in *hOGG1* and risk of cancer: A review of the functional and epidemiologic literature. *Mol. Carcinog.* **42**, 127-141.
- (a) Wiederholt, C.J., Delaney, M.O., and Greenberg, M.M. (2002). Interaction of DNA containing Fapy•dA or its C-nucleoside analogues with base excision repair enzymes. Implications for mutagenesis and enzyme inhibition. *Biochemistry* **41**, 15838-15844.
- (b) Wiederholt, C.J., and Greenberg, M.M. (2002). Fapy•dG instructs Klenow exo- to misincorporate deoxyadenosine. *J. Am. Chem. Soc.* **124**, 7278-7279.
- Wiederholt, C.J., Delaney, M.O., Pope, M.A., David, S.S., and Greenberg, M.M. (2003). Repair of DNA containing Fapy•dG and Its  $\beta$ -C-nucleoside analogue by Formamidopyrimidine DNA Glycosylase and MutY. *Biochemistry* **42**, 9755-9760.
- Wiederholt, C.J., Patro, J.N., Jiang, Y.L., Haraguchi, K., and Greenberg, M.M. (2005). Excision of formamidopyrimidine lesions by endonucleases III and VIII is not a major DNA repair pathway in *Escherichia coli*. *Nucleic Acids Research* **33**, 3331-3338.
- Williams, S.D., and David, S.S. (1998). Evidence that MutY is a monofunctional glycosylase capable of forming a covalent Schiff base intermediate with substrate DNA. *Nucl. Acids Res.* **26**, 5123-5133.
- Wilson, T.E., and Lieber, M.R. (1999). Efficient processing of DNA ends during yeast nonhomologous end joining. Evidence for a DNA polymerase  $\beta$  (Pol4)-dependent pathway. *J. Biol. Chem.* **274**, 23599-23609.
- Wilson III, D.M., and Bohr, V.A. (2007). The mechanics of base excision repair, and its relationship to aging and disease. *DNA Repair* **6**, 544-559.
- (a) Winter, R. B., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. II. The *E. coli* lac repressor-operator interaction: equilibrium measurements. *Biochemistry* **20**, 6948-6960.
- (b) Winter, R. B., Berg, O. G., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. III. The *E. coli* lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961-6977.
- Wood, R.D., Mitchell, M., and Lindahl, T. (2005). Human DNA repair genes, 2005. *Mutation Research* **577**, 275-283.
- Wuenschell, G.E., O'Connor, T.R., and Termini, J. (2003). Stability, miscoding potential, and repair of 2'-deoxyxanthosine in DNA: Implications for nitric oxide-induced mutagenesis. *Biochemistry* **42**, 3608-3616.

-Y-

- Yanofsky, C., Cox, E.C., and Horn, V. (1966). The unusual mutagenic specificity of an *E. coli* mutator gene. *Proc. Natl. Acad. Sci. USA.* **55**, 274-281.

-Z-

- Zaika, E.I., Perlow, R.A., Matz, E., Broyde, S., Gilboa, R., Grollman, A.P., and Zharkov, D.O. (2004). Substrate discrimination by Formamidopyrimidine-DNA glycosylase. A mutational analysis. *J. Biol. Chem.* **279**, 4849-4861.
- Zhang, Q.-M., Ishikawa, N., Nakahara, T., and Yoney, S. (1998). *Escherichia coli* MutY protein has a guanine-DNA glycosylase that acts on 7,8-dihydro-8-

- oxoguanine:guanine mispair to prevent spontaneous G:C→C:G transversions. *Nucleic Acids Research* **26**, 4669-4675.
- Zhang, Q.-M., Miyabe, I., Matsumoto, Y., Kino, K., Sugiyama, H., and Yonei, S. (2000). Identification of repair enzymes for 5-formyluracil in DNA. Nth, Nei, and MutM proteins of *Escherichia coli*. *J. Biol. Chem.* **275**, 35471-35477.
  - Zharkov, D.O. (2008). Base excision repair. *Cell. Mol. Life Sci.* **65**, 1544-1565.
  - Zharkov, D.O., Golan, G., Gilboa, R., Fernandes, A.S., Gerchman, S.E., Kycia, J.H., Rieger, R.A., Grollman, A.P. and Shoham, G. (2002). Structural analysis of an *Escherichia coli* endonuclease VIII covalent reaction intermediate. *EMBO J.* **21**, 789-800.
  - Zharkov, D.O., Rieger, R.A., Iden, C.R., and Grollman, A.P. (1997). NH<sub>2</sub>-terminal proline acts as a nucleophile in the glycosylase/AP-lyase reaction catalyzed by *Escherichia coli* formamidopyrimidine-DNA glycosylase (Fpg) protein. *J. Biol. Chem.* **272**, 5335-5341.
  - Zharkov, D.O., Rosenquist, T.A., Gerchman, S.E., and Grollman, A.P. (2000). Substrate specificity and reaction mechanism of murine 8-oxoguanine-DNA glycosylase. *J. Biol. Chem.* **275**, 28607-28617.
  - Zharkov, D.O., Shoham, G., and Grollman, A.P. (2003). Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair* **2**, 839-862.

**Yann-Vaï LE BIHAN**

**Etude structurale et fonctionnelle de la reconnaissance et de la métabolisation de lésions puriques et pyrimidiques dans l'ADN par la Formamidopyrimidine-ADN glycosylase**

**Résumé :**

Les oxydations sur les bases nucléiques constituent l'une des sources principale d'apparition de lésions sur l'ADN, qui peuvent être mutagènes ou létales pour les cellules en l'absence de réparation de l'ADN. La Formamidopyrimidine-ADN glycosylase (Fpg), une enzyme procaryote du système de réparation de l'ADN par excision de base (BER), initie la réparation d'un large panel de lésions de ce type *via* ses activités ADN glycosylase (excision de la base oxydée) et AP lyase (clivage du site abasique par  $\beta,\delta$ -élimination). Nous avons réalisé des études fonctionnelles par des techniques biochimiques et structurales par cristallographie des rayons X afin de préciser la spécificité de substrat et le mécanisme catalytique de Fpg. Ainsi, nous avons pu mettre en évidence des déterminants structuraux permettant à cette enzyme d'accommoder des lésions de tailles très différentes dans son site actif, en l'occurrence des résidus 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) substitués ou non en N7 par des adduits encombrants. D'autre part, nous avons caractérisé structuralement et fonctionnellement la reconnaissance et l'excision par Fpg d'une lésion pyrimidique, la 5-hydroxy-5-méthyle-hydantoïne (Hyd). Ainsi, nous avons montré que cette lésion appariée à une cytosine était un bon substrat pour l'enzyme, et nous avons précisé structuralement le mode de reconnaissance de l'Hyd par Fpg. D'autre part, nous avons mis en évidence un comportement inattendu de l'enzyme sur ce substrat. En l'occurrence, nous avons montré biochimiquement et structuralement qu'un pontage covalent se formait en quantités non négligeables entre Fpg et l'Hyd dans des conditions physiologiques.

**Mots clés :** Réparation de l'ADN; Réparation par excision de base; Formamidopyrimidine-ADN glycosylase; 2,6-diamino-4-hydroxy-5-formamidopyrimidine; 7,8-dihydro-8-oxo-guanine; 5-hydroxy-5-méthyle-hydantoïne.

**Structural and functional study of the recognition and metabolization of puric and pyrimidic DNA lesions by the Formamidopyrimidine-DNA glycosylase**

**Abstract:**

Oxidations on nucleic bases constitute one of the major sources of DNA lesions appearance, which can be mutagenic or lethal for cells in the absence of DNA repair. The prokaryotic Formamidopyrimidine-DNA glycosylase (Fpg), a base excision DNA repair (BER) enzyme, initiate the repair of a wide range of such lesions *via* its DNA glycosylase (excision of the oxidized base) and AP lyase (cleavage of the AP site by  $\beta,\delta$ -elimination) activities. We carried out functional studies by biochemical techniques and structural studies by X-ray crystallography so as to state Fpg's substrate specificity and catalytic mechanism. Thus, we have been able to underline the structural determinants enabling this enzyme to accommodate lesions of very different sizes in its active site, in this case 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG) residues N7-substituted or not by bulky adducts. On the other hand, we structurally and functionally characterized the recognition and excision by Fpg of a pyrimidic lesion, the 5-hydroxy-5-methyl-hydantoin (Hyd). Thus, we have shown that this lesion paired with a cytosine was a good substrate for the enzyme, and stated structurally the recognition mode of Hyd by Fpg. On the other hand, we have underlined an unexpected behaviour of the enzyme on this substrate. In this case, we have biochemically and structurally shown that a covalent link was formed in sizeable quantities between Fpg and Hyd in physiological conditions.

**Keywords:** DNA repair; Base-excision-repair; Formamidopyrimidine-DNA glycosylase; 2,6-diamino-4-hydroxy-5-formamidopyrimidine; 7,8-dihydro-8-oxo-guanine; 5-hydroxy-5-méthyl-hydantoïne.



**Centre de Biophysique Moléculaire**  
**Équipe « Interactions nucléoprotéiques : Approches Structurales et Fonctionnelles de la Réparation de l'ADN »**  
**Rue Charles Sadron**  
**45071 Orléans cedex 2**

