Structure and kinetic characterization of human sperm-specific glyceraldehyde-3-phosphate dehydrogenase, GAPDS
Apirat Chaikuad, Naeem Shafqat, Ruby Al-Mokhtar, Gus Cameron, Anthony R Clarke, Leo Brady, Udo Oppermann, Jan Frayne, Wyatt W. Yue

To cite this version:

HAL Id: hal-00581535
https://hal.archives-ouvertes.fr/hal-00581535
Submitted on 31 Mar 2011

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.
Structure and kinetic characterization of human sperm-specific glyceraldehyde-3-phosphate dehydrogenase, GAPDS

Apirat Chaikuad¹, Naeem Shafqat¹, Ruby Al-Mokhtar¹, Gus Cameron¹, Anthony R. Clarke¹, R. Leo Brady¹, Udo Oppermann², Jan Frayne² and Wyatt W. Yue²

¹ Structural Genomics Consortium
University of Oxford
Oxford, UK OX3 7DU

² School of Biochemistry
University of Bristol
Bristol, UK BS8 1TD

¹ Botnar Research Centre
NIHR Biomedical Research Unit
Oxford, UK OX3 7LD

¹ These authors contribute equally to this work.
² Correspondence may be addressed to either of these authors (email jan.frayne@bristol.ac.uk or wyatt.yue@sgc.ox.ac.uk).

Short title: Structure and kinetics of human GAPDS
SYNOPSIS

Human sperm-specific glyceraldehyde-3-phosphate dehydrogenase (hGAPDS) is a glycolytic enzyme essential for the survival of spermatozoa, and constitutes a potential target for non-hormonal contraception. However, enzyme characterization of GAPDS has been hampered by the difficulty in producing soluble recombinant protein. Here, we have over-expressed in E. coli a highly-soluble form of hGAPDS truncated at the N-terminus (hGAPDS54), and crystallized the homo-tetrameric enzyme in two ligand complexes. The hGAPDS54-NAD⁺-phosphate structure maps the two anion recognition sites within the catalytic pocket that correspond to the conserved P₁ site and the newly recognised P₂ site identified in other organisms. The hGAPDS54-NAD⁺-glycerol structure shows serendipitous binding of glycerol at the P₁ and new P₂ sites, demonstrating the propensity of these anion recognition sites to bind non-physiologically-relevant ligands. A comparison of kinetic profiles between hGAPDS54 and its somatic equivalent reveals a moderate 3-fold increase in catalytic efficiency for hGAPDS54. This may be attributable to subtle amino acid substitutions peripheral to the active centre that influence the charge properties and protonation states of catalytic residues. Our data therefore elucidate structural and kinetic features of hGAPDS that might provide insightful information towards inhibitor development.

Abbreviations
GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GAPDS, glyceraldehyde-3-phosphate dehydrogenase, sperm-specific; G3P, glyceraldehyde-3-phosphate; PDB, protein data bank; RMSD, root-mean-square deviation.
INTRODUCTION

The glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH; E.C. 1.2.1.12) catalyzes the NAD⁺-dependent oxidative phosphorylation of d-glyceraldehyde-3-phosphate (d-G3P) to form 1,3-diphosphoglycerate (DPG). GAPDH is widely recognised as a drug target due to its central role in glycolysis, and in non-glycolytic processes such as nuclear RNA transport, DNA replication and repair, membrane fusion and apoptosis [1]. For example, GAPDHs from trypanosomatid species, which depend heavily on glycolysis for ATP production, are promising drug targets to combat tropical parasitic diseases [2-4]. In mammals, accumulation of GAPDH in the nucleus is linked to cell death [5], hence its inhibition might provide an anti-apoptotic application. In several neurodegenerative diseases associated with the expression of mutant poly-glutamine proteins, e.g. Huntington’s disease (HD), GAPDH has been shown to bind the mutant proteins [6]. GAPDH also reportedly binds the amyloid precursor protein (APP), implying a potential role in the development of Alzheimer’s disease [7].

Mammals possess not only the somatic GAPDH isozyme present in all tissues except spermatozoa, but also a sperm-specific isozyme (GAPDS) expressed at late stages of spermatogenesis [8, 9]. GAPDS shows significant sequence divergence from its somatic counterpart. For example, human GAPDS (hGAPDS) shares only 56% sequence identity with its somatic counterpart (hGAPDH; Figure S1A), and contains a unique N-terminal 72-residue proline extension [8] necessary for tight association with the cytoskeletal fibrous sheath of the spermatozoa flagellum [10]. The coordinated movements of sperm require substantial amounts of ATP provided by glycolysis, as evidenced by the compartmentalization of glycolytic enzymes in spermatogenic cells [11]. A GAPDS⁻/knockout mouse suffered defects in sperm motility, a crucial determinant for male fertility [12], revealing the potential of targeting GAPDS as a novel contraceptive method alternative to the hormone-based approaches. Chlorinated anti-fertility compounds have been shown to inhibit sperm motility by selectively targeting the sperm-specific and not the somatic enzyme [13-15], albeit with side effects from in vivo trials [16, 17].

To date, the wealth of structural data from archaeal [18], bacterial [19, 20] and eukaryotic GAPDHs [21-23] have revealed a common two-domain architecture and homo-tetramer assembly. Uniform cofactor and d-G3P substrate binding pockets featuring two anion recognition sites (traditionally known as P₁ and P₂ sites, for the binding of substrate and inorganic phosphate respectively) have also been well documented [20, 24]. Combined with extensive kinetic studies [25, 26], the structures are consistent with a well-accepted flip-flop reaction mechanism. This begins with a nucleophile attack by a cysteine sulfhydryl on the carbonyl of d-G3P to form a thio-hemicetal intermediate followed by hydride transfer from the thio-hemicetal to NAD⁺. The resulting thioacyl-enzyme is then attacked by an inorganic phosphate to form DPG. In this model, the C3 phosphate of substrate d-G3P is expected to bind to the P₁ site during the formation of the hemicetal intermediate, and flips to the P₂ site before hydride transfer, to vacate the P₁ site for an incoming inorganic phosphate [20, 27].

Compared to somatic GAPDHs which have been structurally and kinetically characterized (reviewed in [28]), the sperm-specific isozyme GAPDS has proven intractable to study largely due to the insolubility of native and recombinant proteins. Recently, the structure of a hetero-tetrameric complex of rat GAPDS-E. coli GDPH has been reported to moderate resolution, [29], and provided the first view of a mammalian GAPDS, albeit within an unnatural oligomer. Here, we have over-expressed human GAPDS truncated at the N-terminus (hGAPDS₅₅₉), to gain insights into its structural and kinetic properties. The hGAPDS₅₅₉ homo-tetrameric crystal structures, determined in complexes with phosphate and NAD⁺ (hGAPDS₅₅₉-NAD⁺-PO₄) and with glycerol and NAD⁺ (hGAPDS₅₅₉-NAD⁺-Gol), provide the first high-
resolution view of the two anion recognition sites in a human enzyme. Combined with a comparative kinetic analysis, these structures provide a rigorous description of the sperm-specific isozyme that may assist inhibitor design.
EXPERIMENTAL

Protein expression and purification
A DNA fragment encoding hGAPDS (hGAPDSΔN, aa 69-398; GenBank entry 7657116) was subcloned into the pNIC-CTHF vector (GenBank entry 124015079) incorporating a C-terminal TEV-cleavable His-tag. The corresponding plasmid was transformed into *Escherichia coli* BL21(DE3)-R3, cultured in 1 L of Terrific Broth at 37 °C, and induced with 0.5 mM IPTG overnight at 18 °C. Cells were homogenized in lysis buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 20 mM imidazole) and insoluble material was removed by centrifugation. The supernatant was purified by affinity (Ni-sepharose) and size-exclusion (Superdex 75) chromatography. Purified protein was treated with His-tagged TEV protease overnight at 4 °C, and then passed over 1 mL of Ni-Sepharose resin. The TEV-treated protein, containing the remnant vector-encoded sequence ‘AENLYFQ’ at the C-terminus, was concentrated to 12 mg/ml and stored in storage buffer (10 mM HEPES pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, 0.5 mM tris(2-carboxyethyl)phosphine hydrochloride (TECP) at -80 °C.

Crystallization and data collection
hGAPDSΔN was crystallized at 4 °C in 150-nl drops using the sitting drop vapour diffusion method. Viable crystals of the hGAPDSΔN-NAD⁺-PO₄ complex were obtained by mixing protein pre-incubated with 5mM NAD⁺ (Sigma) with the reservoir solution (20% PEG 3350, 0.2 M Na/K phosphate and 10% (v/v) ethylene glycol) in a 1:2 volume ratio. Crystals of the hGAPDSΔN-NAD⁺-Gol complex were fortuitously obtained from mixing protein alone with reservoir solution (20% PEG 3350, 0.2 M Na₂SO₄, 10% (v/v) ethylene glycol and 0.1 M bis-tris propane pH 6.5) in a 1:1 volume ratio. Attempts to crystallize hGAPDSΔN in the apo form have not been successful to date. All crystals were cryo-protected with reservoir solution supplemented with 25% (v/v) glycerol, and flash-cooled in liquid nitrogen. Diffraction data were collected on beamline I03 at the Diamond Light Source (hGAPDSΔN-NAD⁺-PO₄) and in-house on a Rigaku FR-E Superbright source (hGAPDSΔN-NAD⁺-Gol). Data were processed with MOSFLM and SCALA from the CCP4 Suite [30].

Structure determination
Both crystal forms belong to the C-centered monoclinic spacegroup C2 (Table 1). Structures were solved by molecular replacement with PHASER [31] using the rat sperm GAPDS coordinates [29] as a search model. Density modification and NCS averaging were performed with DM [32], and the improved phases were used for automated model building with Buccaneer [33]. Iterative cycles of manual model building using COOT [34] alternated with refinement in REFMACS [35] were performed. At the last refinement step TLS restrained refinement was applied using TLS tensor groups determined by the TLSMD server [36]. Data collection and refinement statistics are summarised in Table 1.

Steady-state enzyme kinetics
Recombinant hGAPDSΔN was dialyzed against storage buffer containing activated charcoal to remove bound cofactor, enabling accurate measurement of protein concentration. Examination of the protein spectra showed no NAD⁺ absorbance at 260 nm showing that the NAD⁺ had been successfully removed. Kinetic assays were based on the forward reaction, in which phosphorylation of d-G3P was catalyzed in the presence of NAD⁺ and inorganic phosphate [37]. All reactions were performed in 10 mM sodium pyrophosphate, 20 mM sodium phosphate pH 8.5, 3 μM dithiothreitol and 10 mM sodium arsenate, using 500 ng/ml enzyme. The *Kₘ* for d-G3P was measured using concentration range 0.3-4.0 mM at a fixed NAD⁺ concentration (0.5 mM). The *Kₘ* for NAD⁺ was determined using concentration range 0.02-1 mM at a fixed d-G3P concentration (1.5 mM). All reagents apart from hGAPDSΔN were dispensed in a 96-
well plate, and pre-incubated at 25 °C for 2 min. The reaction was initiated by the addition of hGAPDSΔN. Steady-state reactions were recorded at 340 nm for 3 min with readings taken every 2 sec, using a VERSAmax microplate reader (Molecular Devices Corp, UK). The tangent at time zero was well defined and accurately defined the initial velocity. Values for $K_{\text{m}}$, $V_{\text{max}}$ and $k_{\text{cat}}$ were calculated using GraphPad Prism (GraphPad Software Inc.) by non-linear regression analysis. The same process and parallel kinetic assays were also performed with human GAPDH from erythrocytes (Sigma).

**Accession numbers**

Atomic coordinates and structure factors have been deposited in the Protein Data Bank under the accession numbers 3h9e (hGAPDSΔN-NAD$^+$-PO$_4$) and 3pfw (hGAPDSΔN-NAD$^-$-GoI).
RESULTS

Recombinant production and structure determination of hGAPDS_{San}
To identify the optimal domain boundary of hGAPDS for structural studies, we expressed a series of full-length and truncated constructs in E. coli as N/C-terminally His-tagged proteins (Figure S1B). We observed that the full-length protein was predominantly insoluble when recombinantly expressed. High levels of soluble expression were found in several N-terminally truncated constructs harbouring the His-tag at the C-terminus, while fusing the affinity tag to the N-terminus of the equivalent constructs render them less soluble. We subsequently purified the C-terminal His-tagged construct encompassing aa 69-407 (referred hereafter as hGAPDS_{San}) to homogeneity (Figure S1C) and its crystal structure was determined in two complexes: one bound with NAD^+ and inorganic phosphate (hGAPDS_{San}-NAD^+-PO_4) at 1.72 Å resolution, and the other bound with NAD^+ and glycerol (hGAPDS_{San}-NAD^+-Gol) at 2.15 Å resolution (Table 1).

In both hGAPDS_{San} complexes, the asymmetric unit comprises two subunits O and P (C_{α}-RMSD between subunits ~0.1-0.2 Å) (Figure 1A), which generate another subunit pair O and R by a twofold symmetry operation (lying approximately on the RQ plane, using previous nomenclature [38]), resulting in a homotetramer with 222 point symmetry (Figure 1B). Similar to other GAPDHs, each subunit consists of an N-terminal nucleotide binding domain (N-domain; aa 69-220) and a C-terminal catalytic domain (C-domain; aa 223-407). The N-domain consists of an eight-stranded β-core flanked by helices on both sides in a classical Rossman fold. The C-domain is built from a five-stranded β-sheet core with helices on one side, and harboirs the active-site catalytic residues Cys224 and His251, as well as a conserved S-shaped loop (S-loop; aa 252-278) that forms the tetrameric core. The S-loop constitutes the main site of intersubunit contacts and has previously been suggested to determine NAD/NADP cofactor specificity [39]. The hGAPDS_{San} structures are similar to other GAPDHs e.g. human liver/placental (hGAPDH) [22], Bacillus stearothermophilus (BsGAPDH) [20, 27, 40], E. coli (EcGAPDH) [19] and Leishmania mexicana (LmGAPDH) [2-4] enzymes, as evidenced by C_{α} RMSD values ranging from 0.7-2.3 Å albeit sharing only 20-67% sequence identity.

The hGAPDS_{San} homo-tetramer
The quaternary arrangement and intersubunit interfaces of hGAPDS_{San} are consistent with the functional oligomeric state proposed for other GAPDHs, with subtle differences in two features namely the electrostatic surface potential and intersubunit selectivity cleft. The hGAPDS_{San} structures reveal three prominently charged patches at the dimeric and tetrameric interfaces, contributed from the positively-charged S-loop (aa 252-278), the negatively-charged 350-372 region and the mix-charged 106-130 region (Figure 1C). The widespread charge distribution in hGAPDS_{San} is more prominent than somatic hGAPDH (PDB id 1u8f, 1y8a) due to several non-conserved charged/polar residues (e.g. Glu112, Lys119, Tyr127, Lys128, Ser252, Tyr253, Arg265, His275, Asp351, Glu352, Lys370) (Figure S1A), and suggests that hGAPDS_{San} tetramerization is facilitated via intersubunit electrostatic complementarity. For example, at the dimeric interface the positively-charged S-loop of subunit P protrudes into a negatively-charged groove from the 350-372 region of subunit O, creating a number of salt bridges (Figure S2). At the tetrameric interface, close to the cofactor binding site, the mix-charged 106-130 region of subunit O packs against the positively-charged S-loop of subunit R and the negatively-charged 350-372 region of subunit Q.

A second differentiating feature between hGAPDH and hGAPDS_{San} is the selectivity cleft adjacent to the adenosine pocket of the NAD^+-binding site (Figure 1D). The selectivity cleft exhibits cross-species
variability in size, and has been exploited for small-molecule inhibitor development as exemplified by
the selective binding of an adenosine derivative NMDA [N^6-(1-naphthalenemethyl)-2′-deoxy-2′-(3,5-
dimethoxybenzamido)adenosine] to the selectivity cleft of the trypansomatid LmGAPDH (PDB id 1i32)
[2-4]. In our hGAPDSαN structures, loop β3-α3 residues (Phe108, Ile109 and Asp110) from one subunit
(e.g. O) form one side of the selectivity cleft, while S-loop residues (Pro263, Ser264, Arg265 and Lys266)
of another subunit (e.g. R) form the other side (Figure 1D). These residues are conserved between
hGAPDH and hGAPDS with the exception of Arg265 in hGAPDS (Gly in hGAPDH) (Figure S1A). This Arg
residue confers reduced conformational flexibility to the S-loop compared to the more adaptable glycine
in hGAPDH. As a result, while the hGAPDH S-loop adopts two flexible conformations (resulting in ‘open’
and ‘closed’ states of the cleft; PDB id 1u8f, 1znq) proposed to be inherent to the somatic isozyme [22,
23], only the ‘closed’ state is observed in our hGAPDSαN selectivity cleft. Furthermore, the Arg265 side-
chain forms electrostatic interactions and water-mediated hydrogen bonds with Asp110 on the opposite
side of the cleft, thereby constricting the cleft cavity to ~4 Å in width, compared with ~5 Å in hGAPDH
[23] and 7-8 Å in LmGAPDH [2, 3]. Modelling the NMDA ligand onto the NAD+ molecule in our structure
reveals severe steric clashes with Pro263 from S-loop and Phe108 from β3-α3 loop (Figure 1E). This,
together with a much reduced selectivity cleft cavity compared to LmGAPDH, might prevent the binding
of NMDA to hGAPDSαN.

Conservation of anion binding sites in the hGAPDSαN-NAD+·PO4 complex

Our hGAPDSαN complexes reveal an NAD+ molecule bound in each subunit (Figure 2A), adopting
essentially identical conformation and interactions as the rat GAPDS and somatic hGAPDS structures.
Additional electron density was observed in the substrate binding pocket adjacent to the cofactor-
binding site (Figure 2A) and was refined as two bound phosphate ions per subunit, presumably
originating from the crystallization buffer. The positions of the two phosphate ions are identical in both
subunits, and map two anion binding sites in hGAPDS that are commonly observed in other GAPDHs
(Figure 2B). The first phosphate site, known previously as the P1 site, is adjacent to the nicotinamide
and ribose moieties of NAD+. The bound phosphate interacts with residues enclosing the P1 site (Thr254,
Thr256, Arg306) as well as the nicotinamide ribose hydroxyl group of NAD+ (Figure 2C, left panel). The
second phosphate site is close to the catalytic residue Cys224 (closest phosphate oxygen is ~3.3 Å from
the cysteine thiol) (Figure 2C, right panel) and corresponds to the ‘new P1’ site previously observed in the
LmGAPDH (PDB id 1gyp) and EcpGAPDH structures (PDB id 1dc4) [19, 24]. This is at a distance of 2.5-3.0 Å
from the ‘original P1’ site reported for the R>GAPDH structure (PDB id 1gd1) [20] which is further away
from Cys224.

The two disparate P1 locations observed in various GAPDH structures are proposed to be commensurate
with the two different conformations of a nearby β-loop-α segment (referred hereafter as the active site
group) - namely an ‘out’ conformation for the ‘original P1’ site’ and an ‘in’ conformation for the ‘new P1
site’ (Figure 2B) [27]. In the hGAPDSαN-NAD+·PO4 structure, the active site segment (aa 280-292) adopts
the ‘in’ conformation for subunit O, and a mixed ‘in’-‘out’ conformation (50% occupancy each) for
subunit P (Figure 2D). The two conformers differ structurally in the orientation of the conserved
residues Thr283 and Gly284, which are displaced from the ‘in’ (closer to the second phosphate) to the
‘out’ (away from the second phosphate) conformations by ~ 2.0-3.8 Å (Cα RMSD). Interestingly, the
second phosphate from both subunits occupies the ‘new P1’ site (despite presence of the ‘out’-conformer
in subunit P), and are anchored by Ser223, Thr225 and His251 (Figure 2C, right panel). The ‘in’-
conformer of subunits O and P further contributes hydrogen bonds to the second phosphate ion via the
main-chain Thr283 and Gly284 atoms (~2.5-2.7 Å), which are not provided in the subunit P ‘out’
conformer as the two residues are more distant from the second phosphate (3.7 and 5.7 Å respectively)
(Figure 2D, compare left and right).
Binding of glycerol in the anion recognition sites
Structural characterization of a hGAPDS\textsubscript{SN}-NAD\textsuperscript{+} crystal grown in a sulphate-containing condition reveals electron density peaks at the two anion binding sites (Figure 3A) that do not fit the shape of sulphate ions often observed to bind in various GAPDH structures. Instead, this density is a good match to glycerol present in our purification and cryo-protectant buffers. As a result two glycerol molecules were refined into the P_1 and ‘new P’ sites of each subunit, and we referred to this structure as the hGAPDS\textsubscript{SN}-NAD\textsuperscript{+}-Gol complex. Neither of the two glycerol molecules overlaps with the expected location of the substrate D-G3P glyceraldehyde moiety between the two anion recognition sites (Figure 3A). The interactions between the two glycerol molecules and the enzyme are similar to the phosphate ions from the hGAPDS\textsubscript{SN}-NAD\textsuperscript{+}-PO_4 structure, albeit less extensive than the phosphate ions as reflected by a degree of structural degeneracy and flexibility in the binding conformation of glycerol. Although in this structure the active site segment adopts fully the ‘out’ conformation, the glycerol at the new P site is further away from the catalytic cysteine (compared to the phosphate in the hGAPDS\textsubscript{SN}-NAD\textsuperscript{+}-PO_4 complex) enabling it to form direct and water-mediated polar interactions with residues from the active site segment (Ser282, Thr283 and Gly284) (Figure 3B). The water-mediated hydrogen bond network bridging the two glycerol molecules is also conserved in the hGAPDS\textsubscript{SN}-NAD\textsuperscript{+}-PO_4 structure.

Comparative kinetic analysis of hGAPDS\textsubscript{SN}
The steady-state kinetic properties of hGAPDS\textsubscript{SN} were measured and compared with those of somatic hGAPDH (Table 2). The two isozymes share similar $K_m$ values for D-G3P substrate and rate constant. However, hGAPDS\textsubscript{SN} has a 3-fold lower $K_m$ for NAD\textsuperscript{+} (35 \textmu M) than hGAPDH (100 \textmu M), hence resulting in an approximately 3-fold increased catalytic efficiency ($k_{cat}/K_m$) with regards to NAD\textsuperscript{+}. The steady-state kinetics for hGAPDH in this study are comparable to those previously reported for hGAPDH purified from brain tissue [28], showing a $k_{cat}$ of $\sim$100 s$^{-1}$ and $K_m$ of 0.07 mM and 0.16 mM for NAD\textsuperscript{+} and D-G3P, respectively.
DISCUSSION

hGAPDS shares conserved structural features with other GAPDHs
Here we report the first detailed structural and kinetic characterization of human sperm-specific GAPDS. By fusing a C-terminal affinity tag and truncating the N-terminal proline region that hinders correct protein folding as shown in this work, and in a recent report [41], we have obtained high levels of recombinant soluble hGAPDS_{\text{sn}}. Our work significantly extends a recent structural study on rat GAPDS [29], where very low recombinant expression levels and solubility required complexation with native E. coli GAPDH for crystallisation. Cross-species hetero-tetramers in various stoichiometries have been reported between GAPDH and GAPDS subunits previously [29, 42]. Nevertheless, given the proximity of the active site cleft to the subunit interfaces, it is unlikely that studies of GAPDS in these unnatural hetero-tetrameric forms provide a realistic assessment of GAPDS structure and activity. Our high-resolution hGAPDS_{\text{sn}} structures confirm homo-tetrameric formation and show that electrostatic interactions and charge complementarity appear to play a substantial role in tetrameric assembly, more so than in hGAPDH, to facilitate the correct orientation of subunits. Our structures also reveal an effectively closed selectivity cleft in hGAPDS as compared to other GAPDHs, particularly the trypanosomal enzymes, and therefore lend further support to the strategy of designing species-specific inhibitors that fill the trypanosomal selectivity cleft without inhibiting the human enzymes [2, 3].

Mapping of anion recognition sites in hGAPDS
The binding positions of phosphate in our hGAPDS_{\text{sn}}-NAD^{+}-PO_{4} complex map the two anion recognition sites in hGAPDS, which are occupied by glycerol in our hGAPDS_{\text{sn}}-NAD^{+}-Gol structure. In addition, the trapping of active-site glycerol was not unprecedented for GAPDHs, since glycerol binding to the P_{1} site was previously reported in an inhibitor-bound T. cruzi GAPDH structure (PDB id 3dm7). It remains to be determined how glycerol, presumably derived from our purification or cryo-protection buffers, out-competes the binding of sulphate ions (closely similar to the physiological phosphates) that are present in high concentration in the crystallization solution. Our findings nevertheless reveal the flexible nature and capacity of the anion recognition sites to accommodate uncharged, polar and non-physiologically relevant ligands, and provide important information to their mode of interactions.

Previous structural characterization of GAPDHs structures has led to the widely-discussed P_{1} and P_{3} nomenclature and flip-flop mechanism [19, 27, 40]. In contrast to the P_{3} site, the true position of the P_{1} site remains unclear since the identification of a ‘new P_{1}’ site (at a distance from the ‘original’ P_{1} site) in a number of GAPDH structures. The divergence in the location and number of P_{1} site(s) is postulated to correlate with the conformations of the neighbouring ‘active site segment’ alternating during catalysis, to maintain interactions between the phosphate ion and the conserved threonine and glycine residues from the active site segment (Thr283 and Gly284 in hGAPDS) [27]. The hGAPDS_{\text{sn}}-NAD^{+}-PO_{4} structure clearly indicates that a phosphate ion is bound at the new P_{1} site when the active site segment adopts the fully ‘in’ conformation (presumably mimicking a ternary state of enzyme poised for catalysis) and the partially ‘out’ conformation (presumably allowing flipping of the hemi-acetal intermediate), while we did not observe phosphate binding at the ‘original’ P_{1} site. This contrasts with the BsGAPDH ternary complex, known to have both new and original P_{1} sites competent and partially occupied by two phosphate moieties, with the active site segment adopting two conformations [27]. It is not clear whether co-existence of two alternative P_{1} positions portrays a common catalytic property or a species-specific feature [27]. Since this is not the case in hGAPDS, a different mechanism for the human enzymes can be envisaged where the original P_{1} site may not be present.
Kinetic differences between hGAPDS and hGAPDH may arise from electrostatic properties
Disregarding the unique polyproline N-terminus in hGAPDS, the amino acid differences between hGAPDS and hGAPDH are mostly confined to the protein surface, contributing to differences in their surface electrostatic potentials. The active site architecture, ligand binding mode and catalytic residues, however, are highly conserved among the isozymes, in part reflected by their kinetic parameters being in the same order of magnitude. Nevertheless, subtle differences above standard deviations are notable, particularly the moderate 3-fold higher catalytic efficiency for hGAPDS than hGAPDH. This may be attributable to the effect of amino acid substitutions peripheral to the catalytic centre, often known as ‘2nd sphere’ residues for enzymes known to have identical active sites but different enzymatic properties [43, 44]. A number of amino acid positions lining the adenine and nicotinamide pockets of the NAD⁺-binding site (Figure 4) involve point charge changes between hGAPDS and hGAPDH. For example, Tyr173 and Asp311 in hGAPDS are substituted by Phe and Asn respectively in hGAPDH (Figure 4). These ‘2nd sphere’ residues, while not perturbing the structural integrity of the active site, could influence the catalytic efficiency of the enzyme by possibly removing a hydrogen bond with NAD⁺ (Tyr173Phe substitution) or by altering the pKₐ, and consequently the protonation state (Asp311Asn substitution) of nearby catalytic residues.

To conclude, isoform-specific inhibitor development for hGAPDS has been implemented for several decades as a strategy for non-hormonal contraception, though suffers from a lack of ligand-bound structural information for the human GAPDH isozymes. Our data highlight subtle yet significant differences in the structural and kinetic features between hGAPDS and other GAPDH enzymes, and therefore need to be considered for future design of inhibitors targeting the sperm-specific isozyme.

ACKNOWLEDGEMENTS
We thank Georgina Berridge and Rod Chalk for assistance in mass spectrometry analyses, and staff at the Diamond Light Source for access to synchrotron data collection facilities.

AUTHOR CONTRIBUTION
AC, NS performed the structural determination. RA, JF, GC, ARC, RLB performed the kinetic experiments and analyzed the data. WWY, UO, JF, RLB supervised the project. AC, WWY designed the experiments and wrote the manuscript. All authors read and approved the final manuscript.

FUNDING
The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust. The work was also supported by the NIHR Biomedical Research Unit to UO, and by the Wellcome Trust (Ref: 07746) and UK BBSRC (Ref: BB/F007256) to RLB.
REFERENCES


13
## TABLES

### Table 1 Data Collection and Refinement Statistics for hGAPDS\textsubscript{A}N structures

<table>
<thead>
<tr>
<th>Complex</th>
<th>hGAPDS\textsubscript{A}N-NAD\textsuperscript{+}-PO\textsubscript{4}</th>
<th>hGAPDS\textsubscript{A}N-NAD\textsuperscript{+}-GOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB id</td>
<td>3h9e</td>
<td>3pfw</td>
</tr>
<tr>
<td>Beamline</td>
<td>Diamond Light Source, I03</td>
<td>Rigaku FR-E Superbright</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9763</td>
<td>1.5418</td>
</tr>
<tr>
<td>Spacegroup</td>
<td>C2</td>
<td>C2</td>
</tr>
<tr>
<td>Resolution range(^a) (Å)</td>
<td>68.23 – 1.72 (1.81 – 1.72)</td>
<td>23.49 – 2.15 (2.27-2.15)</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>(a = 142.0, b = 71.8, c = 81.3) Å</td>
<td>(a = 143.6, b = 71.9, c = 81.1) Å</td>
</tr>
<tr>
<td>No. unique reflections(^a)</td>
<td>72,227 (10,230)</td>
<td>55,308 (5,001)</td>
</tr>
<tr>
<td>Completeness(^a) (%)</td>
<td>99.2 (96.3)</td>
<td>94.1 (92.0)</td>
</tr>
<tr>
<td>(I/\sigma) (^a)</td>
<td>11.4 (2.3)</td>
<td>7.0 (2.3)</td>
</tr>
<tr>
<td>R(_{merge}) (^a) (%)</td>
<td>7.7 (51.4)</td>
<td>17.1 (67.1)</td>
</tr>
<tr>
<td>Redundancy(^a)</td>
<td>4.4 (3.6)</td>
<td>5.3 (5.3)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. atoms in refinement (P/L/O)(^b)</td>
<td>5301/ 108/ 523</td>
<td>5194/ 124/ 310</td>
</tr>
<tr>
<td>(R\text{cryst}) (%)</td>
<td>14.3</td>
<td>16.8</td>
</tr>
<tr>
<td>(R\text{free}) (%)</td>
<td>17.6</td>
<td>22.2</td>
</tr>
<tr>
<td>B-factor (P/L/O)(^b) (Å(^2))</td>
<td>26/ 20/ 29</td>
<td>28/ 32/ 30</td>
</tr>
<tr>
<td>RMSD bond length(^c) (Å)</td>
<td>0.014</td>
<td>0.016</td>
</tr>
<tr>
<td>RMSD bond angle(^c) ((^\circ))</td>
<td>1.50</td>
<td>1.54</td>
</tr>
</tbody>
</table>

\(^a\)Values in brackets show the statistics for the highest resolution shells.

\(^b\) P/L/O denote protein, active-site ligands and other solvent molecules, respectively

\(^c\) RMSD, root-mean-square deviation.
Table 2 Steady-state kinetic properties hGAPDS_{2N} and hGAPDH.

<table>
<thead>
<tr>
<th></th>
<th>hGAPDS_{2N}</th>
<th>hGAPDH</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_M$ NAD* (μM)</td>
<td>35 ± 8</td>
<td>100 ± 38</td>
</tr>
<tr>
<td>$K_M$ d-G3P (mM)</td>
<td>0.46 ± 0.18*</td>
<td>0.25 ± 0.09*</td>
</tr>
<tr>
<td></td>
<td>(0.037)**</td>
<td>(0.020)**</td>
</tr>
<tr>
<td>$k_{cat}$ (s^{-1})</td>
<td>234 ± 24</td>
<td>199 ± 17</td>
</tr>
<tr>
<td>$V_{max}$ (µmole/min/mg)</td>
<td>405 ± 41</td>
<td>340 ± 31</td>
</tr>
</tbody>
</table>

Calculations of $K_M$ for d-G3P were based either on *an assumption of an equal concentration of d- and l-G3P in the racemic dl-G3P, or **an approximate 4% d-G3P as the active aldehyde in the racemic substrate according to [45], however the latter is not considered in other GAPDH kinetic analyses.
FIGURE LEGENDS

Figure 1 Structural overview, surface electrostatic potentials and selectivity cleft of hGAPDS\textsubscript{An}. (A) The asymmetric unit of hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-PO\textsubscript{4} complex comprises subunits O and P, each composed of the N-(green) and C-domains (cyan) (inset). (B) The hGAPDS\textsubscript{An} homo-tetramer is assembled from the OP dimer and the QR dimer generated by a two-fold crystallographic symmetry operator (grey plane) lying on the RQ plane. (C) Surface electrostatic potentials of hGAPDS\textsubscript{An} calculated by APBS program [46] reveal three prominently charged regions: the negatively-charged S-loop (blue), positively-charged 350-372 region (red) and mix-charged 110-130 region (orange). The inset shows calculated electrostatic potentials for somatic hGAPDH (PDB id 1u8f). (D) Superposition of the hGAPDS\textsubscript{An} selectivity cleft in closed conformation (yellow backbone) with two structures of somatic hGAPDHs showing the closed (blue, PDB id 1u8f) and open conformations (orange, PDB id 1znq). The closed conformation in hGAPDS\textsubscript{An} is stabilized by water-mediated hydrogen bonds (blue dashed lines, water in yellow sphere), similar to that seen in hGAPDH (green dashed lines, water in blue spheres). (E) Space-filling representation of the hGAPDS\textsubscript{An} selectivity cleft (yellow) reveals a much reduced cavity than LmGAPDH (magenta, PDB id 1j32) due to the bulky Phe108, Pro263 and Arg265. The trypanosomal GAPDH inhibitor NMDBA (orange) is modelled into the hGAPDS\textsubscript{An} structure based on the binding position of the adenine ring of the NAD\textsuperscript{+} cofactor and illustrates several clashes between the ligand and the protein.

Figure 2 Phosphate binding sites in hGAPDS\textsubscript{An}. (A) Refined 2F\textsubscript{o}-F\textsubscript{c} map (contoured at 1.0 o) for NAD- and phosphate ions in the hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-PO\textsubscript{4} complex. (B) Superposition of phosphate-bound LmGAPDH (semi-transparent ‘lm’, PDB id 1gyf) and sulphate-bound BgGAPDH (semi-transparent ‘bs’, PDB id 1gd1) onto hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-PO\textsubscript{4} subunit O (grey) maps the two bound phosphates in hGAPDS\textsubscript{An} (grey sticks) at the P\textsubscript{i} site, and at the new P\textsubscript{i} site. The positions of the original and new P\textsubscript{i} sites identified from various GAPDH structures are shown. (C) Interactions of phosphates at the P\textsubscript{i} (left) and new P\textsubscript{i} (right) sites. Note that at the new P\textsubscript{i} site, the closest phosphate oxygen atom is 3.3 Å away from the Cys244 thiol group. (D) Two conformations of the active site segment in hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-PO\textsubscript{4} complex: an ‘in’ conformation in subunit O (left), and mixed ‘in’ (semi-transparent) ‘out’ (grey) conformations in subunit P (right).

Figure 3 Glycerol binding in the hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-Gol complex. (A) Left, F\textsubscript{o}-F\textsubscript{c} difference density in the active site of both subunits (contoured at 3σ) was modelled as glycerol. Right, Phosphate anions from the hGAPDS\textsubscript{An}-NAD\textsuperscript{+}-PO\textsubscript{4} structure (semi-transparent), and two 0-G3P molecules (dotted transparent) one bound at the P\textsubscript{i} site (‘rs’, from rat GAPDS-NAD\textsuperscript{+}-G3P complex, PDB id 2vyv) and the other at the P\textsubscript{i} site (‘lm’, from LmGAPDH-NAD\textsuperscript{+}-G3P complex, PDB id 3cmc) are overlaid. (B) Interactions of the two glycerol molecules (Gol) in subunits O (top) and P (bottom) of hGAPDS\textsubscript{An}. Water molecules are represented as blue spheres, and hydrogen bond networks as dashed lines. Note the water-mediated hydrogen bond network bridging the two glycerol molecules is highly conserved among various binding conformations of glycerol.

Figure 4 Amino acid substitutions peripheral to the active sites between hGAPDS\textsubscript{An} and hGAPDH. (A) Amino acid differences between hGAPDS\textsubscript{An} and hGAPDH are clustered in two regions of the active site: one adjacent to the adenine binding pocket and the other at the nicotinamide binding pocket. Residue labels are coloured blue (basic), red (acidic), green (uncharged polar) or grey (non-polar). (B) Moderate differences in electrostatic potentials of the cofactor pocket and catalytic pocket between hGAPDS\textsubscript{An} and hGAPDH are circled.
Figure 1

A

B

C

D

E

18
Figure 2

A

B

C

D

NAD^+

T254

PO4

R306

H251

T225

C224

S223

Ps pocket

NAD^+

R306

H251

T283

C224

S223

Pi pocket

Pi

Ps

G284

T283

H251

S282

Im

mode

‘out’

mode

Ps'

‘Pl’

be

active

site

segment

new

original

Accepted Manuscript

Biochemical Journal Immediate Publication. Published on 26 Jan 2011 as manuscript BJ20101442

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101442

© 2011 The Authors Journal compilation © 2011 Portland Press Limited
Figure 3
Figure 4

A

B

cofactor pocket

catalytic pocket