Missense mutations of conserved glycine residues in fibrillin-1 highlight a potential subtype of cb-EGF-like domains


To cite this version:


HAL Id: hal-01669921
https://hal.archives-ouvertes.fr/hal-01669921
Submitted on 21 Dec 2017

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.
Missense Mutations of Conserved Glycine Residues in Fibrillin-1 Highlight a Potential Subtype of cb-EGF-like Domains

Philippe Khau Van Kien¹, David Baux¹, Nathalie Pallares-Ruiz¹, Corinne Baudoin¹, Aurélie Plancke¹, Nicolas Chassaing²,³,⁴, Patrick Collignon⁵, Valérie Drouin-Garraud⁶, Alain Hovnanian²,³,⁴, Dominique Martin-Coignard⁷, Gwenaelle Collod-Béroud⁸,⁹, Christophe Béroud¹,⁸,⁹, Anne-Françoise Roux¹, and Mireille Claustres¹,⁸,⁹

¹CHU Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Génétique Moléculaire, Montpellier, F-34000 France; ²CHU Toulouse, Hôpital Purpan, Service de Génétique Médicale, Toulouse, F-31300, France; ³INSERM, U563, Centre de Physiopathologie de Toulouse Purpan, Toulouse, F-31300, France; ⁴Université Toulouse III Paul-Sabatier, Toulouse, F-31400, France; ⁵CHU Marseille, Hôpital d'Enfants de la Timone, Département de Génétique Médicale, Marseille F-13000, France; ⁶CHU Rouen, Département de Génétique, Rouen, France; ⁷CH le Mans, Service de Génétique Médicale, le Mans, F-72000, France; ⁸INSERM, U827, Montpellier, F-34000 France; ⁹Université Montpellier1, UFR Médecine, Montpellier, F-34000 France

*Correspondence to Philippe Khau Van Kien, MD-PhD, CHU Montpellier, Laboratoire de Génétique Moléculaire, UFR Médecine site Nord UPM/IURC, 640 Avenue du doyen Gaston GIRAUD, 34295 MONTPELLIER CEDEX 5, FRANCE. Phone: +33467415360; Fax: +33467415365; E-mail: p-khau_van_kien@chu-montpellier.fr

ABSTRACT: In six index cases/families referred for Marfan syndrome (MFS) molecular diagnosis, we identified six novel mutations in the FBN1 gene: c.1753G>C (p.Gly585Arg), c.2456G>A (p.Gly819Glu), c.4981G>A (p.Gly1661Arg), c.5339G>A (p.Gly1780Glu), c.6418G>A (p.Gly2140Arg) and c.6419G>A (p.Gly2140Glu). These variants, predicted to result in Glycine substitutions are located at the third position of a 4 amino acids loop-region of calcium-binding Epidermal Growth Factor-like (cb-EGF) fibrillin-1 domains #5, #9, #24, #25 and #32. Familial segregation studies showing cosegregation with MFS manifestations or de novo inheritance in addition to in silico analyses (conservation, 3D modeling) suggest evidence for a crucial role of the respective Glycine positions. Extending these analyses to all Glycine residue at position 3 of this 4 residues loop in fibrillin-1 cb-EGF with the UMD predictor tool and alignment of 2038 available related sequences strongly support a steric strain that only allows Glycine or even Alanine residues for domain structure maintenance and for the fibrillin functions. Our data compared with those of the literature strongly suggest the existence of a cb-EGF domain subtype with implications for related diseases.

KEY WORDS: cb-EGF-like domain, Marfan Syndrome, Fibrillin 1 (FBN1) gene, fibrillinopathies
INTRODUCTION

Human fibrillin-1, the protein involved in Marfan syndrome (MFS; MIM#154700) is characterized by a specific modular organization with structural and non-structural functions in the extracellular matrix. The components of this organization consist in repeated cysteine rich structural modules that include: multiple copies of an epidermal growth factor-like (EGF) module of which 43/47 match with the consensus sequence of class I calcium-binding EGF (cb-EGF) (Downing, et al., 1996), TB (or 8-cysteine) or hybrid modules. About 60% of the reported mutations of the \(FBN1\) gene are missense. More than 75% of them are located in the cb-EGF modules of fibrillin-1 and involve mostly their obligatory residues (in particular, cysteine residues implicated in disulfide bond formation or other residues implicated in calcium-binding and/or intra-intermolecular interactions) (Collod-Beroud, et al., 2003; Downing, et al., 1996). Interpretation of the remaining missense mutations can be very challenging and requires additional arguments (familial segregation, epidemiological studies, functional data...) to assess their causality and thus, complicates the outcome of genetic testing with uninformative results. Reciprocally, the observation of deleterious mutations can help to identify additional key positions and therefore to assess clinical significance of missense variants.

To date, among 228 non-related index cases/families referred and analyzed in our laboratory for MFS molecular diagnosis, we have identified six missense mutations in the \(FBN1\) gene in six index cases, which were predicted to result in Glycine substitutions sharing an identical position in their corresponding cb-EGF domain (1.3% of tested chromosomes). This position corresponds to the p.Gly1127Ser mutation that has been intensively studied elsewhere (Francke, et al., 1995; Smallridge, et al., 2003; Whiteman, et al., 1998; Whiteman, et al., 2001). We provide here arguments to generalize previous findings concerning the p.Gly1127Ser mutation and demonstrate a crucial structural role for Glycine (or Alanine) at this position in class I cb-EGF domains.

PATIENTS AND METHODS

Patients

Pedigrees and main clinical data are presented in Fig. 1A and 1B. All the patients were of French Caucasian origin.

Case 1 / pedigree MFS 00021

A 39-year-old female, who was treated since 26 years of age with beta-blockers for a paroxysmal supraventricular tachycardia was referred for a MFS diagnosis (-1 Ghent criteria) because of a suggestive familial history of MFS and a bilateral ectopia lentis operated on at the age of 29 years. She also presented significant striae, positive wrist and thumb signs. A computed tomography (CT) of the hips and the lombosacral region revealed that she had a protrusio acetabulae. The rest of the investigation was normal. Her father and her brother (Fig. 1A) died from a type A aortic dissection, respectively at the age of 59 and 34 years. Both were reported to be of tall stature (194 and 202cm, respectively) with suggestive evidence of skeletal (arachnodactyly and dolichostenomelia) and ocular (strong myopia, which complicate in blindness in the father) involvements. There was no other familial antecedent.

Case 2 / pedigree MFS 00055

A 15-year-old female from healthy parents was referred for a MFS diagnosis (-1 Ghent criteria) after a comprehensive Ghent criteria assessment that identified a skeletal system involvement (scoliosis >20°, wrist and thumb signs, flexum at the elbows, characteristic face, high arched palate with crowding of teeth) and a major dural criterion on MRI. She had no familial history of MFS.

Case 3 / pedigree MFS 00126

A 60-year-old male was referred to confirm the diagnosis as well as for familial screening (Ghent criteria fulfilled). He had a familial history of MFS, a bilateral ectopia lentis operated on at the age of 16 years and 3 spontaneous pneumothorax events. Ghent criteria evaluation revealed that he also presented a scoliosis of more
than 20°, significant pes planus and a high arched palate with crowding of teeth. CT of the lumbosacral region also showed a significant meningocele on S3. Ectopia lentis and other suggestive MFS manifestations were also reported in his son, sisters and mother (Fig. 1A and 1B).

**Case 4 / pedigree MFS 00231**

A 59-year-old father and his 14-year-old son were both referred to confirm the diagnosis (Ghent criteria fulfilled). They showed a similar score for Ghent criteria: a major criterion for the skeletal system (reduced upper to lower segment ratio, wrist and thumb signs, reduced extension at the elbows, protrusio acetabulae, moderate pectus excavatum, joint hypermobility and a high arched palate with crowding of teeth), a major criterion for the cardiovascular system (type A aortic dissection at 40 years of age for the father and aortic root dilatation for the son) and recurrent herniae. There were no other evident MFS manifestations in the family (notably in the parents, who died from a breast cancer and Alzheimer disease at the age of 52 and 80 years, respectively for the mother and the father).

**Case 5 / pedigree MFS 00249**

A 41-year-old female with no familial history of MFS was referred for a MFS diagnosis (-1 Ghent criteria). Ghent criteria assessment identified a major criterion for the skeletal system (pectus carinatum, scoliosis >20°, wrist and thumb signs, protrusio acetabulae, high arched palate with crowding of teeth), a strong bilateral myopia (>6.5 diopters), a mitral insufficiency with a mild valve prolapse and significant striae.

**Case 6 / pedigree MFS 00405**

A 12-year-old female with a familial history of MFS manifestations including ectopia lentis (in the father and in an aunt) was referred for a MFS diagnosis (-1 Ghent criteria): major ocular criterion (ectopia lentis) and skeletal system involvement (wrist and thumb signs, high arched palate with crowding of teeth, joint hypermobility). Her paternal grandmother died suddenly at the age of 58 years. She was operated on heart and ascending aorta at the age of 34 and 48 years.

**FBN1 mutation screening in probands and family study**

Genomic DNA or blood samples from patients were sent to the “Laboratoire de Génétique Moléculaire, CHU Montpellier” from other French centers. DNA was harvested from peripheral blood using various methods. Mutation screening allowing the scanning of the 65 FBN1 exons and flanking regions including the splice sites up to the branching regions was performed in probands with a robot-assisted and single condition PCR/direct sequencing strategy as described elsewhere (Tjeldhorn, et al., 2006; Stheneur, et al., 2009). For several amplicons, at least one of the two primers was newly designed. Mutations were validated with two independent methods (performed from an independent dilution of the DNA and with other primers (Comeglio, et al., 2001; Sakai, et al., 2006; Howarth, et al., 2007)): 1- bidirectional sequencing and 2- dHPLC (performed on a WAVE DNA fragment analysis system (Transgenomic, Elancourt, France) for the c.1753G>C, c.2456G>A, c.4981G>A, c.6418G>A and the c.6419G>A mutations) or High-Resolution Melting analysis (HRM) (performed on a LightCycler 480 (Roche Diagnostics Corporation) for the c.5339G>A mutation). Both methods were then applied in proband relatives to establish familial segregation of each variant and in 95 French Caucasian controls. In all familial cases, haplotype analysis were performed in view to check for the good concordance of the samples origin and to complete Ghent criteria assessment using at least 8 microsatellites markers at the FBN1 locus (Judge, et al., 2001). Detailed protocols are available upon request.

Mutation numbering refers to the FBN1 cDNA GenBank reference sequence: NM_000138.3, with the A of the ATG translation initiation codon as nucleotide +1 (www.hgvs.org/mutnomen).

**Conservation patterns and in silico predictions**

Orthologs of the FBN1 gene were identified with the NCBI and EnsEMBL websites (http://www.ncbi.nlm.nih.gov/ and http://www.ensembl.org/) and sequences were aligned with the human fibrillin 1 precursor reference sequence accession NP_000129.2 using ClustalW (Chenna, et al., 2003). Jalview Alignment Editor (Clamp, et al., 2004), and homemade Perl (http://www.perl.org/) scripts were used to visualize and analyze
the respective alignments. The UMD-predictor tool (Frederic, et al., 2009) was used to assess the impact at both RNA and protein level for all possible single nucleotide substitutions of the 31 fibrillin-1 Glycine codons that correspond to the third amino acid position of all the fibrillin-1 cb-EGF modules identified with a 4 amino acid loops between obligatory Cysteine residues 2 and 3. This includes the six missense mutations reported here and the 6 previously described (see further).

3D Modeling of the fibrillin-1 Glycine substitution effects

Supramolecular structure and NMR data available in the Protein Data Bank (http://www.rcsb.org/, PDB ID:1EMN) (Downing, et al., 1996) was used for the cb-EGF#32 domain (p.Gly2140Arg and p.Gly2140Glu). For p.Gly585Arg, and p.Gly1661Arg, located in the cb-EGF#5 and #24 domains of the fibrillin-1, respectively, a homology model was created based on PDB template 1LMJ (40 and 45% sequence identity, respectively)(Smallridge, et al., 2003). PDB template 1UZP (40% and 54% sequence identity with cb-EGF-like domain #9 and #25, respectively)(Lee, et al., 2004) was used to build the model to analyze p.Gly819Glu and p.Gly1780Glu. Template search was performed with Domain Fishing 1.0 (Contreras-Moreira and Bates, 2002) and Modeller program (version 8.2, http://salilab.org/modeller/; (Eswar, et al., 2007)) was used for homology modeling. The quality of the resulting structures was assessed using the evaluation tools available on the CBS web site (http://bioserv.cbs.cnrs.fr/). Mutagenesis and visualization were performed with PyMOL (http://pymol.sourceforge.net).

Figure 1. Clinical data and familial investigation of the six Glycine missense mutations suggesting a class I cb-EGF subtype. (A)- Affected pedigrees with familial segregation of each Glycine mutation. Pedigree/Generation/subject number and status concerning Marfan (MFS) manifestations in each family member are indicated. Squares, male subjects; circles, female. Affected and unaffected subjects are represented by solid and open symbols, respectively. Slashes denote deceased. Subjects with samples available for this study are indicated with an asterisk. Presence and absence of the Glycine missense mutation is represented by +/- or -/-, respectively. (B)- Main clinical data and Ghent criteria evaluation for MFS in affected subjects. +: criteria are fulfilled (i.e. the diagnosis of MFS is established); -1: a criterion is missing before testing to fulfill the criteria.
Mutation identification and segregation studies

Six missense mutations were identified in the 6 unrelated French index cases, in whom no other mutation suspected to be deleterious was identified (Fig. 1). The c.1753G>C (p.Gly585Arg), c.2456G>A (p.Gly819Glu), c.4981G>A (p.Gly1661Arg), c.5339G>A (p.Gly1780Glu), c.6418G>A (p.Gly2140Arg) and c.6419G>A (p.Gly2140Glu) mutations are, to date, not reported in the literature, in the currently updating UMD-FBN1 online database (Collod-Beroud, et al., 2003) and were absent in 190 ethnically-matched control chromosomes. Data concerning MFS diagnosis criteria assessments (Ghent criteria) in positive tested relatives, familial investigations and mutation analyses are given in Fig. 1. With the exception of pedigree MFS 00249 for which no familial sample was available, each mutation cosegregated with manifestations of the disease in pedigrees MFS 00021, MFS 00126, MFS 00231 and MFS 00405. In pedigree MFS 00055, the c.6418G>A (p.Gly2140Arg) mutation was shown to arise de novo, although a possible germinal mosaicism in the parents could not be ruled out. Results of the haplotype study were concordant with the origin of the samples, data of the pedigrees and disease status.

Conservation patterns and in silico predictions

These missense mutations involved a conserved Glycine residue, located in a loop-region between obligatory Cysteine residues 2 and 3, that is present in 31/43 cb-EGF-like modules in fibrillin-1. Alignment of these 43 modules revealed two subtypes: one with a loop involving 4 residues (32 sequences including a Glycine (n=31) or an Alanine (n=1) at position 3 in cbEGF#42) and a second (the 11 remaining sequences) involving either a shorter or longer loop (i.e. cbEGF#15, #18-19, #26, #34-35, #38-41, #43). Orthologous alignments of fibrillin-1 in 19 different species (including Homo sapiens) revealed a similar pattern on a set of 833 sequences matching class I cb-EGF domains. A Glycine is at position 3 in 592/611 (97%) loops involving 4 residues. The only alternative residue is Alanine (100%, i.e. 19/611-592), the second smallest amino-acid after Glycine. A comparison of the respective distribution of Glycine and Alanine (excluding Glycine) from the position 1 (Glycine 115/611=19%, Fischer’s exact test: P<0.0001 and Alanine 42/611-115=8%, P<0.0001), from the position 2 (Glycine 120/611=20%, P<0.0001 and Alanine 5/611-120=1%, P<0.0001) and position 4 (Glycine 0/611=0%, P<0.0001 and Alanine 9/611-0=1.5%, P<0.0001) showed that both residues are specific of the third position. Amino-acid sequence alignments of fibrillin-2 and fibrillin-3 confirmed these findings. Finally, we assessed an NCBI alignment of 10421 cb-EGF sequences identified with the PFAM website (http://pfam.sanger.ac.uk, for PF07645). Only 8050/10421 contained the obligatory 6 Cysteines of the module and were taken into account. Of these, 2006 showed a loop composed of 4 residues between Cysteines 2 and 3. Glycine (1713/2006) and Alanine (143/2006) are found in position 3 of these loops in 85 and 7%, respectively (each alignment text file is available on request). The high conservation pattern of this Glycine position was also confirmed by the UMD-predictor analysis. Outpout of this analysis (Supp. Table S1) showed that all potential Glycine substitutions are predicted to be “pathogenous” except for the Alanine residues for which depending on the cb-EGF module predictions and scores were ranging from “probably polymorphism” to “probably pathogenous” or “pathogenous” and from 59 to 82/100, respectively. All these data suggest an important role for the domain maintenance of this position and a subtype of the fibrillin class I cb-EGF described by Downing et al (Downing, et al., 1996), Fig. 1.

3D Modeling of the fibrillin-1 Glycine substitution effects

The possible effect of these Glycine substitutions on the structure of their corresponding modules was assessed by in silico 3D Modeling analysis (Fig. 2). The Ramachadran plot obtained for the Glycine residue1661 showed
that the Phi-Psi angles are impossible to adopt for an Arginine residue. Therefore, the conformation of the mutant domain is likely to be strongly altered by the presence of this mutation. For the Glycine residues 585, 819, 1780 and 2140, the predicted change to Arginine or Glutamic Acid involves the same mechanism. Indeed, in each case, the mutant residue is predicted to induce steric clashes with its surroundings. The p.Arg585 breaks a loop involving p.Lys612, p.Lys599 and p.Pro600. Thus, the structure of the cb-EGF-like#5 domain is likely to be strongly affected to retain the volume of the Arginine residue (i.e.: van der Waals volume of Glycine is 48Å³ compared to 148Å³ for Arginine). The p.Arg2140 and p.Glu2140 mutants are predicted to interfere with a crucial structural component of the cb-EGF-like domains, a disulfide bridge involving p.Cys2153-p.Cys2164, and p.Cys2137-p.Cys2151, respectively (Fig. 2). A similar effect is predicted with p.Gly819Glu and p.Gly1780Glu, which are likely to disrupt respectively p.Cys832-p.Cys845 and p.Cys1806-p.Cys1793 as well as p.Cys1777-p.Cys1791 covalent bonds (data available on request).


DISCUSSION

To date, we have sequenced the 65 FBN1 exons and flanking regions in 228 non-related patients referred for molecular diagnosis of Marfan syndrome. We have identified in 6 of them (1.3 % of tested chromosomes) a single nucleotide substitution predicting a missense mutation of Glycine residues as the only unclassified and potentially deleterious variant. These mutations were absent in 190 ethnically-matched control chromosomes. They cosegregated with manifestations of the disease within the pedigrees or occured de novo. Study of conservation patterns showed that the corresponding Glycine position is highly conserved in orthologous, paralogous and also in
cb-EGF-like modules sequences that include a 4 amino-acids loop between Cysteine residues 2 and 3. These results were also sustained by the UMD-predictor (Frederic, et al., 2009) analysis that in addition showed no argument for an impact on splicing. Interestingly, the alternative residue in the alignments and for which the predictions made by UMD-predictor were variable at this position was Alanine (Supp. Table S1). Alanine is the second smallest amino-acid after Glycine. This suggests that the third position of the loop could have particular local strains with steric exclusion for larger residues. Thus, we studied the possible effect of these Glycine substitutions on the structure of their corresponding modules by in silico 3D Modeling analysis and we bring arguments for a key position in terms of steric hindrance. All this results suggest a potential cb-EGF type 1 subtype (Fig. 1) with potential implication for related diseases.

To our knowledge, six missense mutations in the FBN1 gene implicating a Glycine residue in such a position (p.Gly544Arg, p.Gly1042Ser, p.Gly1042Asp, p.Gly1127Ser, p.Gly1987Arg, p.Gly1987Val in cb-EGF#4, #11, #13, #30, respectively), in MFS or incomplete MFS subjects have been described and support our arguments (Attanasio, et al., 2008; Collod-Beroud, et al., 2003; Comeglio, et al., 2001; Francke, et al., 1995; Sakai, et al., 2006; Stheneur, et al., 2009). Only the p.Gly1127Ser mutation has been exhaustively studied: it was shown to result in a local cb-EGF like#13 domain folding disruption, a normal synthesis but a reduced fibrillin deposition into the extracellular matrix and proposed to cause disease via an extracellular dominant negative effect (Francke, et al., 1995; Whiteman, et al., 1998; Whiteman and Handford, 2003; Whiteman, et al., 2001). In their studies, Whiteman et al. (Whiteman, et al., 1998; Whiteman, et al., 2001) have previously noted that the p.Gly1127 position in cb-EGF like#13 was highly conserved and located in a turn of a two stranded antiparallel beta-sheet. However, neither Glycine nor Alanine were reported as mandatory at this position in the cb-EGF type I model (Downing, et al., 1996), thereby allowing to class variants and to facilitate the molecular diagnosis workflow of related diseases. This could be explained because of a variable composition in terms of type and number of amino acids in the loop between Cysteine residues 2 and 3. Here, the subtype of cb-EGF type I model, suggested by alignments of 8050 sequences (Fig. 1) and highlighted by 6 related missense mutations with some genetics/independent arguments for pathogenicity, shows how the findings previously made for p.Gly1127Ser could be extended. The p.Gly1127Ser mutation, associated with a mild effect on protein folding, was reported in 10 related-patients with a moderate phenotype (Francke, et al., 1995). None of the patients reported here as well as those described elsewhere when sufficient data are available to assess severity of the phenotype (Attanasio, et al., 2008; Comeglio, et al., 2001; Sakai, et al., 2006) have a severe and a typical MFS phenotype, suggesting that these FBN1 gene mutation class could be encountered in incomplete or atypical MFS. It is difficult to conclude definitely on the precise genotype/phenotype correlations associated with this mutation class. First, because of the marked intra- and inter-familial variability in MFS manifestations and age of onset of the symptoms, that is notably illustrated in pedigree MFS 00231 (c.2456G>A (p.Gly819Glu), Fig. 1). The father (subject II-2), who suffered from a type A aortic dissection at 40 years of age, fulfilled the Ghent criteria clinically. His 14-years-old affected son (subject III-3) displayed a similar phenotype and had an aortic root dilatation unlike her 23-year-old daughter (subject III-2), which also inherited the mutation. Following a comprehensive Ghent criteria assessment, she only presented a high arched palate with crowding of teeth, a suggestive MFS facial appearance, a wrist sign without thumb sign, a strong myopia and a dural ectasia on MRI (not fulfilling Ghent criteria). Considering that she is adult (but also that she may develop an aortic complication), the penetrance of the c.2456G>A mutation towards MFS is incomplete suggesting the importance of modifying factors. Second, we could not extend the clinical and molecular investigations because of a lack of participation of the relatives to fully assess the phenotypes associated with this class of mutation. However, from our study and as reported before we claim that this singular missense mutation class affecting Glycine residues located at the third position of this cb-EGF-like subtype (if resulting in a substitution other than Alanine) should be at least considered as a strong risk factor for a wide range of MFS manifestations, including aortic complications. Aortic dissection and sudden death are highly prevalent within families reported with such missense mutations. To our opinion, the medical management and follow-up should not differ from classic MFS.

We have shown a key residue-position in a potential subtype of class I cb-EGF like module. Both, Glycine and Alanine residues appear mandatory for the structure maintenance and therefore for the fibrillin-1 functions. These findings have implications for genetic testing in MFS, other fibrillinopathies and in other cb-EGF-like related-disorders.
ACKNOWLEDGMENTS

We gratefully acknowledge the clinical research team of the CHU of Montpellier, notably J. Dinet and S. Plagnol for their support. Corinne Baudoin, Aurélie Plancke and part of this work were funded through the « Programme de Soutien aux Techniques Innovantes Côtoeuses 2006 » from the French Ministry of Health.

REFERENCES


