Functional and structural characterisation of 5 missense mutations of the phenylalanine hydroxylase

Martina Pecimonova[1,2], Emil Polak[1], Frantisek Csicsay[3], Kamila Reblova[4], Maja Stojiljkovic[5], Zdenko Levarski[2], Ludovit Skultety[3], Ludevit Kadasi[1,3,6] and Andrea Soltysova[1,3,6]

[1] *Department of Molecular Biology, Faculty of Natural Sciences, Comenius University, Ilkovicova 6, 842 15 Bratislava, Slovakia*
[2] *Comenius University Science Park, Ilkovicova 8, 841 04 Bratislava, Slovakia*
[3] *Institute of Virology, Biomedical Research Center, Slovak Academy of Sciences, Dubravska cesta 9, 845 05 Bratislava, Slovakia*
[4] *Central European Institute of Technology, Masaryk University, Brno, Czech Republic*
[5] *Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia*
[6] *Institute for Clinical and Translational Research, Biomedical Research Center, Slovak Academy of Sciences, Dubravska cesta 9, 845 05 Bratislava, Slovakia*

Abstract. Phenylketonuria (PKU) and hyperphenylalaninemia (HPA) are a group of genetic disorders predominantly caused by mutations in the phenylalanine hydroxylase (PAH) gene. To date, more than 950 variants have been identified, however the pathogenic mechanism of many variants remains unknown. In this study, *in silico* prediction and *in vitro* prokaryotic and eukaryotic expression systems were used to functionally characterise five PAH missense variants (p.F233I, p.R270I, p.F331S, p.S350Y, and p.L358F) previously identified in Slovak and Czech patients. p.F233I, p.R270I, and p.S350Y were classified as deleterious mutations since they showed no specific activity in functional assay and no response to chaperone co-expression. Protein levels of these PAH variants were very low when expressed in HepG2 cells, and only p.S350Y responded to BH4 precursor overload by significant increase in PAH monomer, probably due to reduced rate of protein degradation as the result of proper protein folding. Variants p.F331S and p.L358F exerted residual enzymatic activity *in vitro*. While the first can be classified as probably pathogenic due to its very low protein levels in HepG2 cells, the latter is considered to be mild mutation with protein levels of approximately 17.85% compared to wt PAH. Our findings contribute to better understanding of structure and function of PAH mutated enzymes and optimal treatment of PKU patients carrying these mutations using BH4 supplementation.

Key words: PAH mutations — Phenylketonuria — Specific activity — Chaperons — Functional study

Introduction

Phenylketonuria (PKU, OMIM 261600) and its milder forms, hyperphenylalaninemias (HPA), are a group of metabolic disorders inherited in an autosomal recessive

Correspondence to: Andrea Soltysova, Department of Molecular Biology, Faculty of Natural Sciences, Comenius University, Ilkovicova 6, Mlynska dolina, 842 15 Bratislava, Slovakia
E-mail: soltysova@fns.uniba.sk

manner. PKU and HPA are predominantly caused by the impairment of phenylalanine hydroxylase (PAH, EC 1.14.16.1). This enzyme catalyses conversion of the essential amino acid L-phenylalanine (L-Phe) to L-tyrosine (L-Tyr) in the presence of the cofactor 6(R)-L-erythro-5,6,7,8-tetrahydrobiopterin (BH4), molecular oxygen and iron (Kobe et al. 1999). In PKU patients, uncatabolised L-Phe and its toxic metabolic by-products accumulate in the blood, resulting in disturbed brain development, microcephaly, seizures, and other neurological symptoms (Williams et al. 2008). Therefore, phenylalanine dietary intake in PKU patients must be strictly controlled (Guldberg et al. 1998; Blau et al. 2011).

PAH protein is expressed in the liver and kidneys (Wang et al. 1992; Lichter-Konecki et al. 1999) and it exists in cells in pH-dependent equilibrium between a homodimeric and homotetrameric form (Martinez et al. 1995). Each monomer consists of three domains: the N-terminal regulatory domain (residues 1–142), the catalytic domain (residues 143-410) and the C-terminal oligomerization domain (411–452) (Fusetti et al. 1998; Kobe et al. 1999).

At present, there are more than 950 identified PAH variants which are distributed along the entire gene, of which approximately 60% are missense mutations (Phenylalanine Hydroxylase Gene Locus-Specific Database. http://biopku. org/home/pah.asp, 2006- 2015 (accessed 29.09.16)). Only a small percentage of PAH missense mutations have been functionally characterised. The impact of amino acid substitution on the protein function and stability can be predicted *in silico* by various software or it can be tested *in vitro* using prokaryotic and eukaryotic expression systems. Previous studies show that more than 50 PAH missense mutations do not affect PAH gene transcription and *in vitro* expression of PAH mutants showed normal levels of mRNA compared to the wild type (wt), therefore mutation impact on PAH is mostly studied on the protein level (Dworniczak et al. 1991; Weinstein et al. 1993; Desviat et al. 1995; Waters et al. 1998). In *E. coli*, increased expression of mutant proteins leads to aggregate formation and subsequent protein degradation (Bjørgo et al. 1998). Co-expression with GroEL and GroES bacterial chaperones can promote the proper folding of misfolded mutant proteins preventing them from degradation (Gámez et al. 2000). Co-expression of mutated PAH enzymes in the presence with these chaperones, which are homologues of eukaryotic Hsp60 and Hsp10 (Feldman and Frydman 2000), provides useful information whether the mutation is associated with decreased PAH protein stability caused by protein misfolding (Stojiljkovic et al. 2009).

Mutations with residual enzymatic activity are usually associated with mild PKU and non-PKU HPA, and these can respond to BH4, which has a chaperone-like effect on mild mutants and promotes correct folding and re-establishing of the oligomerisation equilibrium from dimer to tetramer (Blau et al. 2009; Cerreto et al. 2011). Functional protein testing in eukaryotic expression systems in the presence of sepiapterin, which is a BH4 precursor, provides a useful tool for the assessment of whether the mutation is BH4 responsive. This information could be applied as a predictor in PKU treatment using BH4 supplementation.

We previously identified five PAH variants (p.F233I, p.R270I, p.F331S, p.S350Y, and p.L358F) in Slovak PKU patients with an unknown impact on the resultant protein (Polak et al. 2013). The variants p.F233I, p.R270I, p.F331S, and p.S350Y were identified for the first time in the Slovak population while the p.L358F variant was identified in Czech PKU patients contemporaneously with our findings, and it

was classified as a mild mutation by *in silico* analysis (Réblová et al. 2013). In the positions F233, R270, F331, and S350, other amino acid substitutions ere also noted (Kleiman et al. 1993; Benit et al. 1994; Kuzmin et al. 1995; Guldberg et al. 1996, 1998; Tyfield et al. 1997; Leuzzi et al. 2006; Zhu et al. 2010), but only two of them, p.R270S and p.R270K, were functionally assayed *in vitro* (Bjørgo et al. 1998; Trunzo et al. 2016).

Recent studies underline the importance of PAH functional testing in terms of better understanding the genotype-phenotype correlations and the elucidation of BH4 responsiveness (Blau 2016). Thus, the aim of this work is to functionally and structurally characterise p.F233I, p.R270I, p.F331S, p.S350Y, and p.L358F PAH variants and contribute to a better understanding of PKU molecular basis.

Functional analysis of PAH mutated variants was performed in *in vitro* prokaryotic and eukaryotic expression systems. To extend the results, all variants were expressed in the presence of GroEL and GroES chaperones and sepiapterin, which has a chaperone-like effect on PAH protein. Moreover, functional and structural analysis of all variants was performed using molecular dynamics simulations and two consensus tools, Meta-SNP and PredictSNP.

Material and Methods

In silico analysis

The functional impact of each mutation on protein function was predicted using a Meta-SNP server (Capriotti et al. 2013) and consensus Classifier PredictSNP (Bendl et al. 2014). These are web-based tools which combine the following nine methods: Panther, PhD-SNP, SIFT, SNAP, Meta-SNP, MAPP, PredictSNP, PolyPhen-1, and PolyPhen-2.

The structural effect of four missense variants was also investigated using molecular dynamics (MD) simulations. We used our previous 3D model of human PAH (Reblova et al. 2013), where the catalytic domain (residues 143–410) and oligomerisation domain (residues 411–452) were built based on the X-ray structures of truncated forms of human PAH (pdb codes: 1PAH (Erlandsen et al. 1997) and 2PAH (Fusetti et al. 1998)) while the N-terminal regulatory domain (residues 19–142) was built based on the homology modelling of rat PAH (pdb code: 1PHZ (Kobe et al. 1999). Homology modelling was carried out using the Modeller 9.11 program (Sali and Blundell 1993). Based on the model of the human PAH monomer, we built the structure of the mutants p.F233I, p.R270I, p.F331S, and p.S350Y, which carried the mutations c.697T>A, c.809G>T, c.992T>C, and c.1049C>A, using the Xleap module of Amber 12 (Case et al. 2012) and the force field parm99SB (Hornak et al. 2006). Using the Xleap module wild type (wt), mutant

structures were neutralised by K^+ ions (Dang and Koll-man 1995) and solvated by an octahedral SPC/E water box extending 10 Å away from the solute. The equilibration and production phases were carried out using standard protocol (Réblová et al. 2015). Wt and mutant proteins were simulated for 100 nanoseconds (ns). Trajectories were analysed using the Ptraj module of AMBER and visualised using VMD program (Humphrey et al. 1996). We analysed the structural contacts of the studied residues (direct H-bonds with these criteria: the distance between the H-bond donor and acceptor below 3.2 Å and angle from donor atom over the hydrogen to the acceptor atom >120°, salt bridges, and aromatic interactions between amino acids) and their buriedness corresponding to the relative accessible surface area (RSA) ≤ 10%, which was calculated by measuring the accessible surface area of wt amino acid using the Stride program (Stride Web interface. http://webclu.bio.wzw.tum. de/cgi-bin/stride/stridecgi.py, 1995 (accessed 07.10.16); Frishman and Argos 1995) and normalised by the maximum accessible surface area for the corresponding amino acid residue type (Chothia 1976).

Vectors and E. coli protein expression

The mutations c.697T>A, c.809G>T, c.992T>C, c.1049C>A, and c.1074A>T were introduced into the vector pMAL-MBP-c2 containing human wild type PAH cDNA by site-directed mutagenesis using a Phusion Site-Directed Mutagenesis Kit (Thermo Fisher Scientific) with specific primers (Supplementary Table S1). The pMAL-PAH constructs were transformed into *E. coli* Xl1 blue cells, which were used for plasmid propagation and PAH proteins expression, and selected by the appropriate antibiotics. Each clone for subsequent analyses was verified by sequence analysis. All mutant plasmids were also co-transformed with pGroELS plasmids coding GroEL and GroES bacterial chaperones. Each clone for subsequent analysis was verified by sequencing analysis.

Expression of mutated PAH proteins, fused with maltose binding protein (MBP), was performed as previously described (Martinez et al. 1995). Briefly, expression of mutated enzymes was triggered by adding 100 mM IPTG simultaneously with 0.2 mM ferrous ammonium sulphate into the cultivation medium. After 18 hours of incubation, the bacterial cells were collected and sonicated in a buffer containing 200 mM NaCl, 20 mM Hepes, and cOmplete Protease Inhibitor Cocktail (Roche), pH 7.0.

Purification of recombinant PAH fusion proteins

Wt PAH and each mutated form were purified by affinity chromatography using the ÄKTA avant 25 system (GE Healthcare Life Sciences), and MBPTrap™ HP 1 ml column

(GE Healthcare Life Science) loaded with a column buffer containing 200 mM NaCl and 20 mM HEPES, pH 7.0. Elution of PAH-MBP was performed using 10mM maltose dissolved in the column buffer and proteins were subsequently purified by size-exclusion chromatography using a Superdex 200 HR 10/30 column (GE Healthcare Life Sciences) in the same column buffer. Protein concentration was then spectrophotometrically determined using Quick Start™ Bradford Protein Assay (Bio-Rad).

PAH functional assay

PAH functional assay was performed as previously described (Martinez et al. 1995). It comprised of 1 μg of protein, pre-incubated for 4 minutes, with 10 mM L-phenylalanine in the presence of 1 mg/ml catalase and 0.1 M NaHEPES. Next, 100 μM ferrous ammonium sulphate was added and the mixture was incubated for a further minute. The reaction was initiated by adding 75 μM BH4 diluted in 5 mM DTT and terminated with 1% (v/v) acetic acid in ethanol after 1 or 5 minutes. Enzymes were assayed in triplicates at 25°C in 50 μl reaction volumes.

Tyrosine and phenylalanine were separated by HPLC (Agilent 1100 series) in an analytical column Zorbax XDB-C18; 150 × 4.6 mm; 5 μm. The mobile phase was isocratic with the column temperature set at 30°C. Signals were detected by a fluorescence detector with an excitation wavelength of 210 nm and analysed by Chemstation software. Percentual specific activity was calculated from the molar concentration of forming L-tyrosine in the reaction of each PAH mutant in comparison to the wt PAH.

Eukaryotic expression system

pFLAG-PAH vectors were prepared by cloning PAH cDNA into the pFLAG-CMV vector through SalI and NotI restriction sites. Wt and mutant PAH cDNAs were obtained by PCR amplification of corresponding pMAL-PAH constructs using the primers 5'-TTTTGCGGCCGCGATGTCCACTGCG-GTCCTGG-3' and 5'-AAAAGTCGACGGCTTTACTT-TATTTTCTGGAG-3' (Aguado et al. 2006) carrying recognition sites. Each clone selected for subsequent analyses was verified by sequence analysis.

Cell culture and transfection

Human hepatoma cells HepG2 were cultured in Eagle's Minimum Essential Medium (Sigma-Aldrich) supplemented with 10% Fetal Bovine Serum (Sigma-Aldrich). The cells were plated in 12-well plates at 5×10^4 cells *per* well and transfected with Lipofectin® Reagent (Sigma-Aldrich) for transient expression of pFLAG-PAH constructs (0.7 μg of each recombinant plasmid *per* well). The transfecting mix-

ture was replaced after 12 hours of incubation with a fresh medium in the presence or absence of 100 μM sepiapterin (Sigma-Aldrich). The transfected HepG2 cells were collected 36 hours after transfection.

Western blot analysis

The transfected HepG2 cells were lysed in NP-40 lysis buffer containing 150 mM sodium chloride, 1.0% Nonidet P-40, and 50 mM Tris, pH 8.0. Protein concentration was determined spectrophotometrically and 50 μg of proteins were separated by electrophoresis on 10% SDS polyacrylamide gels. Protein transfer on PVDF membrane was subsequently analysed by Western blot with Monoclonal ANTI-FLAG® M2 antibody produced in mice (Sigma Aldrich) and Anti-Mouse IgG (H+L) HRP Conjugate (Promega). Signal was determined using Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore) and ImageQuant LAS 500 (GE Healthcare Life Science). The optical density of individual bands was quantified by GelQuant.NET software from biochemlabsolutions.com.

Statistical analysis

Each protein was expressed in *E.coli* and purified in two independent experiments. Functional assay was performed using proteins from two independent affinity or size exclusion chromatography purifications in three replicates, each for proteins with or without co-transformed GroEL and GroES chaperones. Three independent Western blot analyses of PAH proteins expressed in HepG2 cells were performed. The results are presented as mean ± S.D. Statistical differences between multiple comparisons were performed using adjusted t-test.

Results

In silico analysis

The three p.R270I, p.F331S, and p.S350Y PAH variants were classified as deleterious mutations by all nine prediction methods, and the p.F233I and p.L358F variants were predicted as disease-causing mutations by five of the programs (Table 1).

We also investigated mutant proteins by using MD simulations. The behaviour of mutated residues was analysed based on 100-ns long simulations and compared with wt simulation. Stability of the simulations was detected using Root Mean Square deviation (RMSd). This parameter reveals structural change from the initial starting structure, which was very small for wild-type and also for the mutated proteins on 100 ns time scale, particularly, averaged RMSd values of all systems were below 2 Å (Figure 1). The changes we detected in the mutant simulations were rather local structural rearrangements, see below.

Residue F233 is positioned close to the protein surface (buriedness is 18%) in the hydrophobic pocket created by residues L212, I209, and L213. Simulation of the p.F233I mutant revealed that mutated isoleucine fits well into this position (Figure 2).

Residue R270 (buriedness is 5%) is positioned in a long turn, which is the entrance to the active site. The arginine forms a salt bridge with D282 and H-bond with P275 and both stabilise the long turn (Figure 2 and Figure 3). MD simulations of wt revealed both contacts to be stable. The substitution of large/charged arginine by small/hydrophobic isoleucine prevents such contacts (Figure 2). Conformational change of the long turn was observed in the p.R270I simulation (not seen in the wt simulation), which is probably due

Table 1. The results of *in silico* prediction of the severity of PAH variants using Meta-SNP and PredictSNP tools

Mutation	Meta-SNP					PredictSNP						
	Panther	PhD-SNP	SIFT	SNAP	Meta-SNP	PredictSNP	MAPP	PhD-SNP	PolyPhen-1	PolyPhen-2	SIFT	SNAP
p.F233I	Disease 0.846	Disease 0.543	Neutral 0.410	Neutral 0.375	Disease 0.568	Neutral 60%	Neutral 78%	Disease 68%	Neutral 67%	Disease 40%	Disease 79%	Neutral 58%
p.R270I	Disease 0.924	Disease 0.911	Disease 0.010	Disease 0.815	Disease 0.888	Disease 87%	Disease 86%	Disease 82%	Disease 74%	Disease 81%	Disease 79%	Disease 89%
p.F331S	Disease 0.924	Disease 0.952	Disease 0.000	Disease 0.680	Disease 0.817	Disease 87%	Disease 77%	Disease 88%	Disease 74%	Disease 81%	Disease 79%	Disease 72%
p.S350Y	Disease 0.971	Disease 0.896	Disease 0.000	Disease 0.790	Disease 0.854	Disease 87%	Disease 81%	Disease 88%	Disease 74%	Disease 68%	Disease 79%	Disease 85%
p.L358F	Disease 0.701	Neutral 0.362	Disease 0.000	Disease 0.540	Disease 0.574	Neutral 68%	Neutral 77%	Neutral 58%	Neutral 67%	Disease 45%	Disease 46%	Neutral 67%

Neutral, neutral variants; Disease, disease causing variants. Outputs of Meta-SNP: Panther, PhD-SNP, SNAP, Meta-SNP – reported value between 0 and 1, if > 0.5 mutation is predicted disease causing variant; SIFT – if > 0.05 mutation is predicted neutral variant. Outputs of PredictSNP: reported percentage represents normalized confidence of used tool.
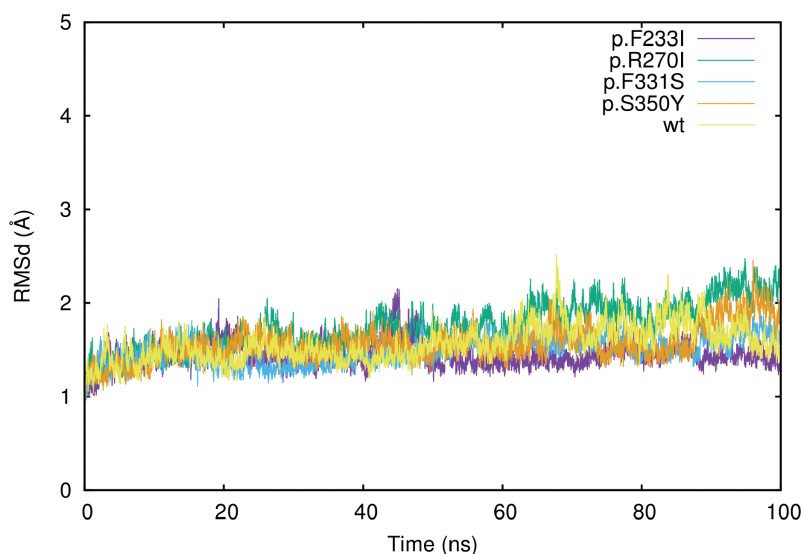
Figure 1. Time development of RMSd values calculated for protein Cα atoms with respect to the starting structures along the molecular dynamics simulations of wild type (wt) and mutants proteins p.F233I, p.R270I, p.F331S and p.S350Y. Terminal parts (residues 19–22 and 425–432) and flexible loop in the middle of the structure (residues 132–152), which was partially modelled into the protein structure, were not included in the calculations of root mean square deviation (RMSd).

to the missing contacts (Figure 3). This kind of change may impact the functionality of the protein.

Residue F331 is positioned in turn in the active site (buriedness is 9%). It stacks onto W326 and it is adjacent to E330, which together with H285 coordinate Fe^{3+} ion (Figure 2). MD simulation of wt revealed that the stacking between F331 and W326 is perfectly stable. The substitution F331 for serine abolishes the stacking interaction, additionally the presence

of serine disrupts hydrophobicity of this area. Even though MD simulation of p.F331S mutant extended up to 200 ns did not show any large structural rearrangement in the active site, it is more likely that p.F331S mutation results in significant impairment of the protein structure/function, which would probably be visible on a longer (microsecond) time scale.

S350 is positioned at the beginning of the active site near protein surface (buriedness is 32%). In the X-ray structure, it
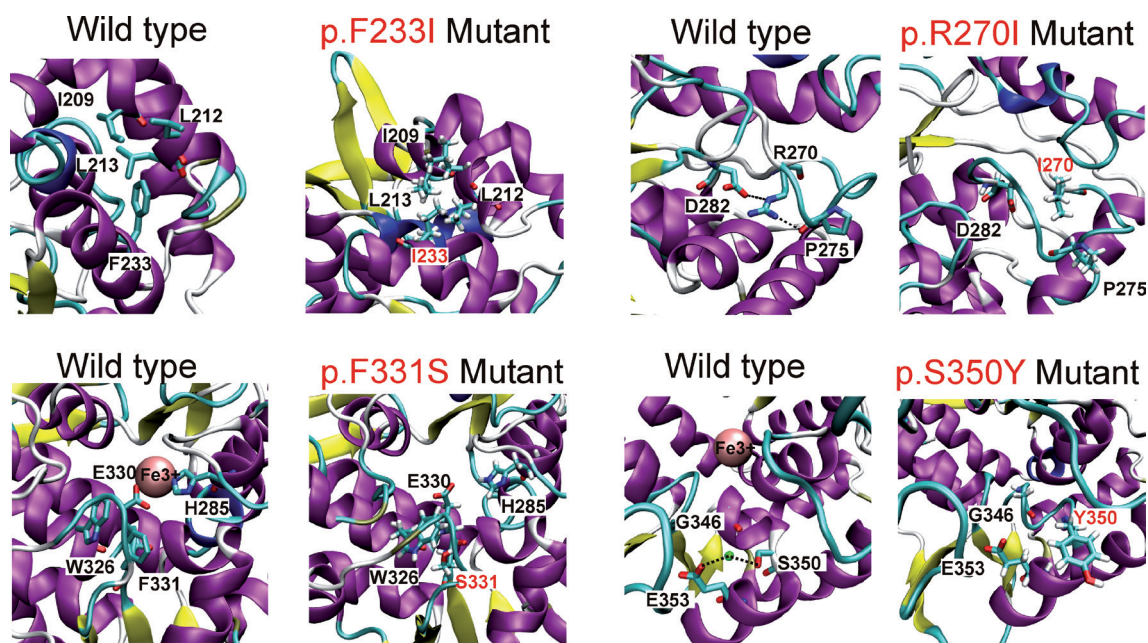
Figure 2. Detailed views on tertiary structures of mutant phenylalanine hydroxylase (PAH) proteins carrying newly discovered missense mutations from molecular dynamics simulations and the corresponding regions in the wild type X-ray protein structure. Wild type and substituting amino acids are highlighted in licorice representation and marked. H-bonds are marked by black dashed lines. Water molecule forming H-bonds to E353 and S350 is in green. Some pictures also show Fe^{3+} coordinated in the active site of the X-ray structure.

A

B
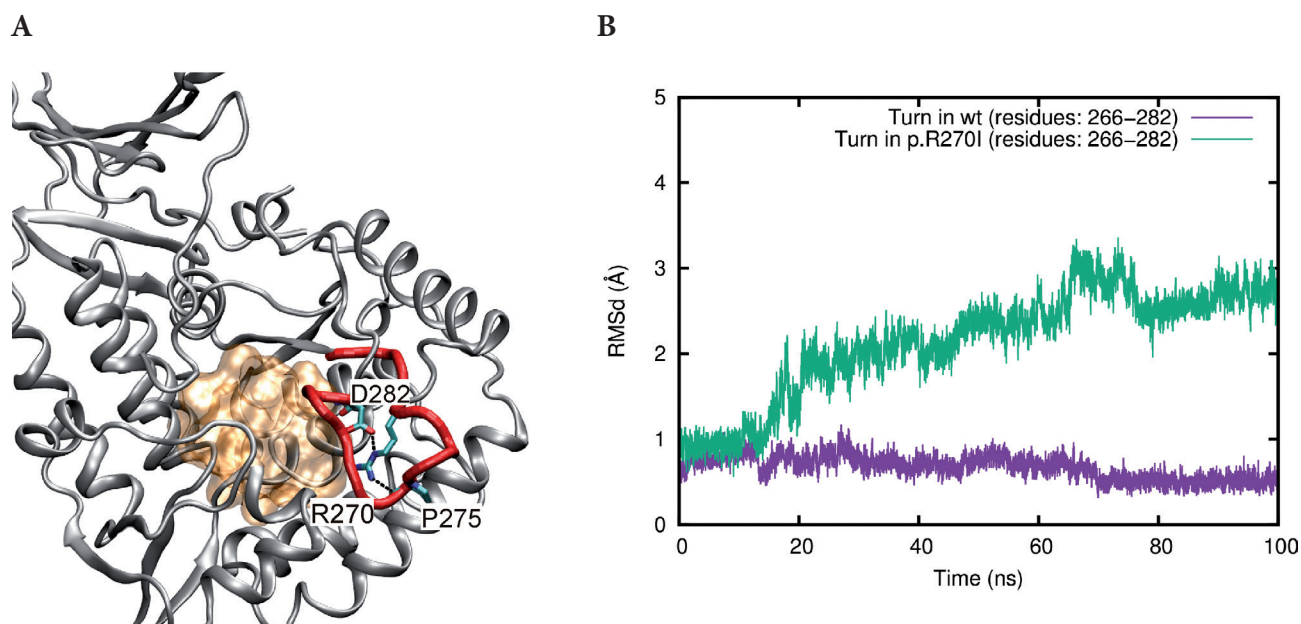
Figure 3. A. Detail of tertiary structure of wt PAH with active site cavity highlighted by orange surface and long turn (in red) at the entrance to the active site formed by residues (266–282). **B.** Time development of RMSd values calculated for turn structure (residues: 266–282) in wt PAH and p.R270I mutant simulations. Increased RMSd in p.R270I simulation indicates structural change. PAH, phenylalanine hydroxylase; RMSd, root mean square deviation; wt, wild type.

creates H-bond to E353 via a water molecule (Figure 2). In the wt simulation, this contact changed to fluctuating direct H-bond between E353 and S350. Additionally, in the wt simulation, side chain hydroxyl of S350 formed fluctuating H-bond to the main chain oxygen of G346. These contacts stabilise the position of the S350 in the active site. In the mutant p.S350Y simulation, neither of these contacts were

formed (Figure 2), nor any new interaction between S350 and the surrounding amino acid was created.

Prokaryotic PAH expression

Human wt PAH and its mutant forms were expressed in *E.coli* from the pMAL vector fused with MBP. After af-
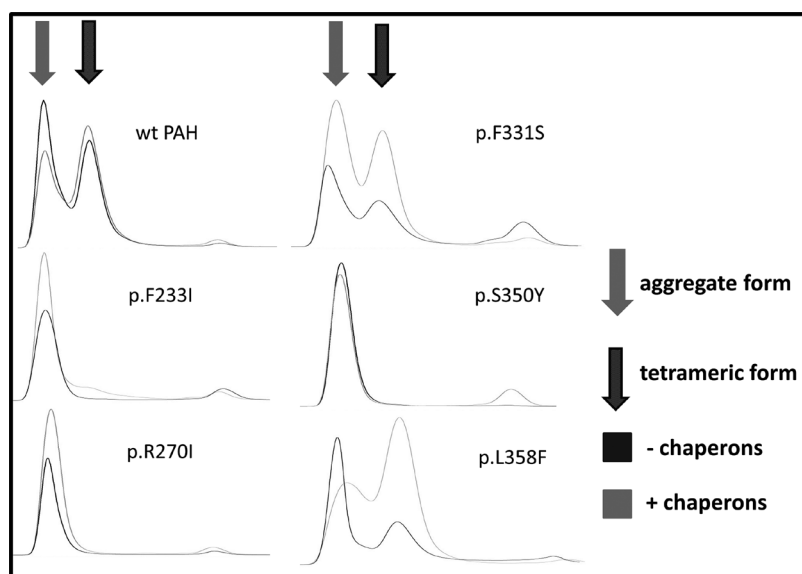
Figure 4. Chromatograms of affinity and SEC purified wt and mutant PAH proteins expressed in *E. coli* using the absorbance at 280 nm. The grey chromatogram shows the proteins co-expressed with GroEL and GroES bacterial chaperones, the black chromatogram represents the proteins without the GroEL and GroES chaperone co-expression. The peaks represent three PAH enzyme forms, from left to right: aggregates, tetramers, and dimers, respectively. SEC, size exclusion chromatography; wt, wild type; PAH, phenylalanine hydroxylase.

Table 2. The results of expression analyses of wt PAH and five mutant variants in the presence or absence GroEL/ES chaperones in *E. coli*
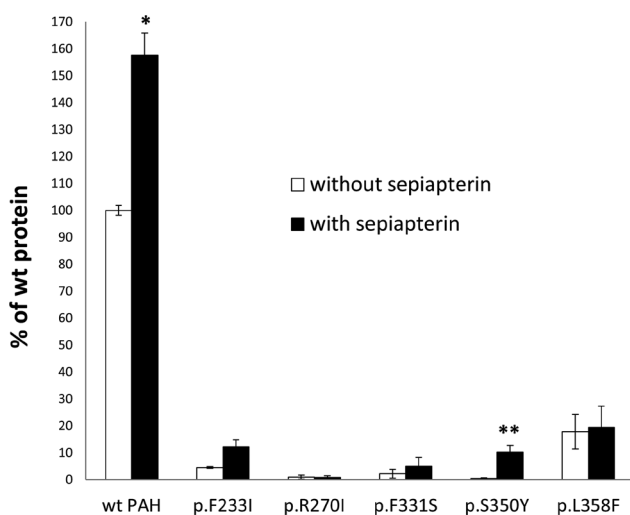
	Expression in E. coli							
	– GroEL/ES chaperones				+ GroEL/ES chaperones			
	Aggregates (%)	Tetramers (%)	Dimers (%)	Specific activity (%)	Aggregates (%)	Tetramers (%)	Dimers (%)	Specific activity (%)
wt	67	32	1	100 ± 1.24	39	59	2	112 ± 8.3
p.F233I	88	0	12	0	83	12	5	0.6 ± 0.2
p.R270I	96	0	4	0	95	0	5	0
p.F331S	48	34	18	99 ± 12.7	51	46	3	115 ± 17.9
p.S350Y	88	0	12	0	87	0	13	0
p.L358F	62	34	4	52 ± 0.7	35	64	1	57.7 ± 4.4

For each protein analysis, the percentage of the various forms was calculated based on peak area after size exclusion chromatography. Percentual specific activity was calculated from molar concentration of forming tyrosine in PAH assay reaction for every mutated variant compared to the wt PAH activity measured by HPLC analysis. Data represent the mean of two independent protein purifications and six independent PAH assays. PAH, phenylalanine hydroxylase; wt, wild type.

finity and size-exclusion chromatography (Figure 4), all three forms (aggregates, tetramers, and dimers) of wt PAH were obtained, of which catalytically active tetrameric form accounted for 32% of the total affinity purified protein. The p.F233I, p.R270I or p.S350Y PAH mutants predominantly formed aggregates (87.74%, 96.22%, and 87.90%, respectively) and lacked tetramers. Tetrameric forms of the remaining p.F331S and p.L358F PAH variants

Figure 5. The levels of PAH wild-type and mutant proteins expressed in HepG2 cells with or without sepiapterin. Data represent the mean ± S.D of three independent Western blot analyses. * statistically significant level changes with the *p*-value of 0.02 between wt PAH proteins expressed in the presence and absence of sepiapterin; ** statistically significant level changes with the *p*-value of 0.007 between p.S350Y PAH proteins expressed in the presence and absence of sepiapterin. PAH, phenylalanine hydroxylase; wt, wild type.

accounted for approximately a third of the total purified protein (Table 2).

Wt PAH and three mutant proteins were partially rescued in the presence of GroEL and GroES chaperones (Table 2). The p.F233I mutant tetramer formation increased to almost 12% of the total purified protein, while the p.F331S and p.L358F mutant proteins had increased tetrameric forms to 46.04% and 63.72%, respectively. For the two mutant proteins, p.R270I and p.S350Y, the percentage ratio in the presence of chaperones remained unchanged.

PAH assay

Functional *in vitro* testing of PAH proteins was performed with tetrameric forms and affinity purified proteins used for mutants lacking tetramers. While p.F233I, p.R270I, and p.S350Y mutant proteins showed no detectable specific activity, the mutant p.L358F exhibited 52 ± 0.7% residual activity and p.F331S PAH specific activity was 99 ± 12.7% compared to wt PAH.

Specific activity was recovered by GroEL and GroES chaperones only in p.F233I mutant, whose residual activity increased to 0.6 ± 0.2% compared to wt PAH. While p.L358F and p.F331S PAH showed no significant change in specific activity (57.7 ± 4.4% and 115 ± 17.9%), those of p.R270I and p.S350Y remained null (Table 2).

Eukaryotic expression system

Human hepatic cells HepG2 were used for PAH expression to test the sepiapterin effect on the protein level. Proteins isolated from non-transfected HepG2 cells were controls for non-specific antibody binding. All mutant PAH levels differed significantly compared to wt PAH (all *p*-values < 0.00013). We observed a significant increase of the wt PAH

protein level, up to 157.67 ± 8% after sepiapterin treatment (p = 0.02), as well as in the case of the p.S350Y mutant protein where the protein level increased from 0.38 ± 0.3% up to 10.2 ± 2.5% (p = 0.007) compared to wt PAH. While p.F233I and p.F331S exert a non-significant increase in the protein level, the protein level of p.L358F remained unchanged. The amount of mutant protein p.R270I was nearly undetectable and there was also no increase in the protein level after cultivation in the presence of sepiapterin (Figure 5).

Discussion

The most frequent missense mutations were functionally and structurally characterised in several studies many years ago (Waters 2003). However, large-scale PKU studies continuously reveal novel variants with an unknown impact on PAH enzyme function (Groselj et al. 2013; Polak et al. 2013; Réblová et al. 2013; Trunzo et al. 2015). Therefore, it is highly important to constantly update and extend these data which contribute to a better understanding of the complex PKU nature. Moreover, data obtained in PAH functional analyses provide a beneficial tool for clinical praxis and BH4 cofactor treatment.

We functionally characterised five missense PAH variants p.F233I, p.R270I, p.F331S, p.S350Y, and p.L358F previously identified in the Slovak and Czech populations (Polak et al. 2013; Reblova et al. 2013). All mutations are located in the catalytic domain and these were subjected to *in silico* analysis by Meta-SNP and PredictSNP web-based tools combining nine prediction programs. The three p.R270I, p. F331S, and p.S350Y mutations were identified as disease-causing mutations by all nine programs and the remaining p.F233I and p.L358F variants were classified as deleterious by five programs.

In silico prediction of p.R270I and p.S350Y PAH mutants was confirmed by *in vitro* testing, neither forming catalytically active tetramers nor exhibiting enzymatic activity. These same results were obtained after their co-expression with GroEL and GroES chaperones. These findings are consistent with Western blot results where cell protein levels of both PAH variants were close to zero, probably due to improper folding and consequent protein degradation (Gjetting et al. 2001; Waters 2003; Kim et al. 2006; Pey et al. 2007). While cultivation with sepiapterin significantly affected only the p.S350Y PAH (p = 0.007), where protein quantity increased to 10.2 ± 2.5% compared to untreated p.S350Y, the p.R270I protein level remained unchanged (Figure 5). We assume that observed increase in p.S350Y level occurs through decreased protein degradation rate due to the increased steady-state level of PAH monomers caused by the BH4 chaperone-like effect. As our MD simulation indicates, the substitution of hydrophobic isoleucine for

the electrically charged arginine abolishes the salt bridge and hydrogen bond formation and subsequent interactions with D282 and P275 required for the stabilisation of loop structure, which is the entrance to the active site. Furthermore, substitution of the hydrophobic tyrosine for the polar serine in the p.S350Y mutant abolishes structural contact with E353, which exists in the X-ray wt protein and was observed in the wt simulation. Therefore, our results suggest that the PAH mutations p.R270I and p.S350Y are null PAH mutations with a severe impact on protein function. These data are consistent with genotype-phenotype correlations in our previous study (Polak et al. 2013). Mutation p.R270I was identified in two unrelated patients in a compound heterozygous state with mutations p.R408W, manifesting severe classical PKU (cPKU), and p.I306V associated with mild PKU (mPKU) phenotype, respectively. The mutation p.S350Y was found in a cPKU patient with one allele uncharacterised. It is interesting to note that this patient was also a carrier of the newly discovered PAH promoter variant (c.−30A>T) with an unknown impact (Polak et al. 2013).

According to the *in silico* prediction analysis, p.F233I should exhibit residual enzymatic activity. The residue F233 is positioned in a catalytic domain and completes the hydrophobic pocket on the protein surface (Flatmark and Stevens 1999). Our MD simulation indicates that it forms no direct structural contacts with neighbouring amino acids, and therefore phenylalanine substitution for isoleucine should not affect the protein folding. However, its predicted milder effect was not confirmed. Since p.F233I PAH does not form any tetramers, it did not exert any specific activity in *in vitro* testing. Although co-expression with chaperones enhanced the tetramer assembling, the specific activity was only 0.6 ± 0.2% compared to wt PAH (Table 2, Figure 4). This is consistent with the *in vitro* eukaryotic expression, where the protein levels were significantly lower than wt PAH (p = 8E-7). Therefore, both *in vitro* results indicate that the mutation p.F233I has a severe impact on PAH protein function. Interestingly enough, clinical manifestation correlates with these *in vitro* findings as well, since this mutation was identified in compound heterozygosity with the well-characterised null mutation p.R408W, and the phenotype in this patient corresponded to severe PKU (Polak et al. 2013).

The PAH variants p.F331S and p.L358F were identified in one patient who suffered from hyperphenylalaninemia (Polak et al. 2013). MD simulation indicates that the substitution of phenylalanine for serine in position 331 prevents the forming of the π-stacking interaction with W326, reported essential for stabilisation of the active site wall (Erlandsen and Stevens 1999), and consequent incorrect packaging. Moreover, F331 is adjacent to E330, which together with H285, coordinate ferric ion at the active site. Despite classification as a deleterious mutation by *in silico* analysis, our results suggest that p.F331S PAH displays *in vitro* enzymatic

activity in PAH assay approximately equal to that of wt PAH (Table 2); both solely and together with co-expressed chaperones. This contradicts *in vitro* eukaryotic analysis where the p.F331S PAH protein levels were approximately 2.23 ± 1.6% compared to wt PAH. While the protein quantity increased to 5 ± 3.2% transfected HepG2 cells cultivated with sepiapterin, these low percentages may indicate defective oligomerisation, decreased stability or high rate of proteolytic degradation (Gjetting et al. 2001; Waters et al. 2001; Waters 2003; Kim et al. 2006; Pey et al. 2007). Due to its mutation-induced PAH level reduction, p.F331S PAH variant could be classified as possibly pathogenic.

The p.L358F PAH variant was in silico predicted disease-causing mutation by five programs of all nine used, and it was previously characterised by MD simulation as a mutation with a milder protein impact (Reblova et al. 2013), which is consistent with our *in vitro* assay results. MD simulation also indicates that leucine substitution for phenylalanine initiates new interaction where F358 stacks on the H201, which is in interaction with Thr200 through hydrogen bond. Although this new interaction does not disrupt the protein function, it reduces its activity (Reblova et al. 2013). *In vitro*, this mutant exhibited approximately 50% residual activity compared to wt PAH with no significant change in activity when co-expressed with GroEL and GroES chaperones (Table 2). Moreover, the protein level expressed in HepG2 cells was the highest from all tested mutations with approximately 17.85 ± 6.4% compared to wt PAH, and it was BH4-nonresponsive since it showed no significant difference in protein quantity in the presence of sepiapterin. Such ambiguous results highlight the necessity of functional testing in mammalian cell cultures because these are more aligned with the natural environment (Waters 2003). Our combined results confirmed the *in silico* prediction of a milder p.L358F impact on the enzymatic function. Overall, since this variant exerted high residual activity of purified proteins in *in vitro* PAH assay and low expression level in the hepatoma cellular model, it can be considered a mild mutation.

In view of this, we functionally assayed five previously uncharacterised missense mutations in the *PAH* gene. Our results indicate that the mutations p.F233I, p.R270I, and p.S350Y are severe mutations lacking residual enzymatic activity, even in co-expression with the GroEL and GroES bacterial chaperones. The protein levels of these three proteins were very low when expressed in HepG2 cells and only p.S350Y responded to BH4 precursor by significant increase in PAH protein level. Considering these results these mutations are classified as null mutations. Despite its deleterious prediction, the PAH variant p.F331S retains its enzymatic activity in *in vitro* functional testing, however, on the other hand, showed very low protein levels in human hepatoma cells; therefore it can be concluded as a probable pathogenic PAH variant. Our results suggest that p.L358F is a mild muta-

tion with retained residual activity. In conclusion, all protein levels of the studied missense mutations significantly differed when expressed in human hepatoma cells HepG2 (Figure 5) compared to wt PAH, probably through their pathogenic effects on protein folding and/or stability. Our findings obtained by protein analysis in the presence of BH4 precursor will have an implication on the optimal treatment of PKU patients carrying these previously uncharacterised mutations using BH4 supplementation. The presented results agree with patients' PKU phenotypes and are a plausible contribution to accumulating data of the structure and function of PAH mutated enzymes.

Conflict of interest. The authors have no conflict of interest that could influence the content or processing of this manuscript.

References

Aguado C., Pérez B., Ugarte M., Desviat L. R. (2006): Analysis of the effect of tetrahydrobiopterin on PAH gene expression in hepatoma cells. FEBS Lett. **580,** 1697–1701 https://doi.org/10.1016/j.febslet.2006.02.005

Bendl J., Stourac J., Salanda O., Pavelka A., Wieben E. D., Zendulka J., Brezovsky J., Damborsky J. (2014): PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput. Biol. **10,** e1003440 https://doi.org/10.1371/journal.pcbi.1003440

Benit P., Rey F., Melle D., Munnich A., Rey J. (1994): Five novel missense mutations of the phenylalanine hydroxylase gene in phenylketonuria. Hum. Mutat. **4,** 229–231 https://doi.org/10.1002/humu.1380040311

Bjørgo E., Knappskog P. M., Martinez A., Stevens R. C., Flatmark T. (1998): Partial characterization and three-dimensional-structural localization of eight mutations in exon 7 of the

human phenylalanine hydroxylase gene associated with phenylketonuria. Eur. J. Biochem. **257,** 1–10
https://doi.org/10.1046/j.1432-1327.1998.2570001.x

Blau N. (2016): Genetics of phenylketonuria: then and now. Hum. Mutat. **37,** 508–515
https://doi.org/10.1002/humu.22980

Blau N., Belanger-Quintana A., Demirkol M., Feillet F., Giovannini M., MacDonald A., Trefz F. K., van Spronsen F. J. (2009): Optimizing the use of sapropterin (BH(4)) in the management of phenylketonuria. Mol. Genet. Metab. **96,** 158–163
https://doi.org/10.1016/j.ymgme.2009.01.002

Blau N., Hennermann J. B., Langenbeck U., Lichter-Konecki U. (2011): Diagnosis, classification, and genetics of phenylketonuria and tetrahydrobiopterin (BH4) deficiencies. Mol. Genet. Metab. **104,** S2–S9
https://doi.org/10.1016/j.ymgme.2011.08.017

Capriotti E., Altman R. B., Bromberg Y. (2013): Collective judgment predicts disease-associated single nucleotide variants. BMC Genomics 14 (Suppl. 3), S2
https://doi.org/10.1186/1471-2164-14-S3-S2

Case D. A., Darden T. A., Cheatham III T. E., Simmerling C. L., Wang J., Duke R. E., Luo R., Walker R. C., Zhang W., Merz K. M., et al. (2012): AMBER 12. University of California, San Francisco

Cerreto M., Cavaliere P., Carluccio C., Amato F., Zagari A., Daniele A., Salvatore F. (2011): Natural phenylalanine hydroxylase variants that confer a mild phenotype affect the enzyme's conformational stability and oligomerization equilibrium. Biochim. Biophys. Acta **1812,** 1435–1445
https://doi.org/10.1016/j.bbadis.2011.07.012

Chothia C. (1976): The nature of the accessible and buried surfaces in proteins. J. Mol. Biol. **105,** 1–12
https://doi.org/10.1016/0022-2836(76)90191-1

Dang L. X., Kollman P. A. (1995): Free energy of association of the K+: 18-crown-6 complex in water: a new molecular dynamics study. J. Phys. Chem. **99,** 55–58
https://doi.org/10.1021/j100001a011

Desviat L. R., Perez B., De Lucca M., Cornejo V., Schmidt B., Ugarte M. (1995): Evidence in Latin America of recurrence of V388M, a phenylketonuria mutation with high in vitro residual activity. Am. J. Hum. Genet. **57,** 337–342

Dworniczak B., Aulehla-Scholz C., Kalaydjieva L., Bartholome K., Grudda K., Horst J. (1991): Aberrant splicing of phenylalanine hydroxylase mRNA: the major cause for phenylketonuria in parts of southern Europe. Genomics **11,** 242–246
https://doi.org/10.1016/0888-7543(91)90129-3

Erlandsen H., Fusetti F., Martinez A., Hough E., Flatmark T., Stevens R. C. (1997): Crystal structure of the catalytic domain of human phenylalanine hydroxylase reveals the structural basis for phenylketonuria. Nat. Struct. Biol. **4,** 995–1000
https://doi.org/10.1038/nsb1297-995

Erlandsen H., Stevens R. C. (1999): The structural basis of phenylketonuria. Mol. Genet. Metab. **68,** 103–125
https://doi.org/10.1006/mgme.1999.2922

Feldman D. E., Frydman J. (2000): Protein folding in vivo: the importance of molecular chaperones. Curr. Opin. Struct. Biol. **10,** 26–33
https://doi.org/10.1016/S0959-440X(99)00044-5

Flatmark T., Stevens R. C. (1999): Structural insight into the aromatic amino acid hydroxylases and their disease-related mutant forms. Chem. Rev. **99,** 2137–2160
https://doi.org/10.1021/cr980450y

Frishman D., Argos P. (1995): Knowledge-based protein secondary structure assignment. Proteins **23,** 566–579
https://doi.org/10.1002/prot.340230412

Fusetti F., Erlandsen H., Flatmark T., Stevens R. C. (1998): Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. J. Biol. Chem. **273,**16962–16967
https://doi.org/10.1074/jbc.273.27.16962

Gámez A., Pérez B., Ugarte M., Desviat L. R. (2000): Expression analysis of phenylketonuria mutations. Effect on folding and stability of the phenylalanine hydroxylase protein. J. Biol. Chem. **275,** 29737–29742
https://doi.org/10.1074/jbc.M003231200

Gjetting T., Petersen M., Guldberg P., Güttler F. (2001): In vitro expression of 34 naturally occurring mutant variants of phenylalanine hydroxylase: correlation with metabolic phenotypes and susceptibility toward protein aggregation. Mol. Genet. Metab. **72,** 132–143
https://doi.org/10.1006/mgme.2000.3118

Groselj U., Tansek M. Z., Podkrajsek K. T., Battelino T. (2013): The IVS8-2A>G (c.913-2A>G) mutation and the PAH deficiency populations of Central Europe. J. Inherit. Metab. Dis. **36,** 157
https://doi.org/10.1007/s10545-012-9562-2

Guldberg P., Levy H. L., Hanley W. B., Koch R., Matalon R., Rouse B. M., Trefz F., de la Cruz F., Henriksen K. F., Guttler F. (1996): Phenylalanine hydroxylase gene mutations in the United States: report from the Maternal PKU Collaborative Study. Am. J. Hum. Genet. **59,** 84–94

Guldberg P., Rey F., Zschocke J., Romano V., Francois B., Michiels L., Ullrich K., Hoffmann G. F., Burgard P., Schmidt H., et al. (1998): A European multicenter study of phenylalanine hydroxylase deficiency: classification of 105 mutations and a general system for genotype-based prediction of metabolic phenotype. Am. J. Hum. Genet. **63,** 71–79
https://doi.org/10.1086/301920

Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C. (2006): Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins **65,** 712–725
https://doi.org/10.1002/prot.21123

Humphrey W., Dalke A., Schulten K. (1996): VMD: visual molecular dynamics. Model J. Mol. Graph. **14,** 33–38
https://doi.org/10.1016/0263-7855(96)00018-5

Kim S.-W., Jung J., Oh H.-J., Kim J., Lee K.-S., Lee D.-H., Park C., Kimm K., Koo S. K., Jung S.-C. (2006): Structural and functional analyses of mutations of the human phenylalanine hydroxylase gene. Clin. Chim. Acta. **365,** 279–287
https://doi.org/10.1016/j.cca.2005.09.019

Kleiman S., Li J., Schwartz G., Eisensmith R. C., Woo S. L., Shiloh Y. (1993): Inactivation of phenylalanine hydroxylase by a missense mutation, R270S, in a Palestinian kinship with phenylketonuria. Hum. Mol. Genet. **2,** 605–606
https://doi.org/10.1093/hmg/2.5.605

Kobe B., Jennings I. G., House C. M., Michell B. J., Goodwill K. E., Santarsiero B. D., Stevens R. C., Cotton R. G., Kemp B. E.

(1999): Structural basis of autoregulation of phenylalanine hydroxylase. Nat. Struct. Biol. **6,** 442–448
https://doi.org/10.1038/8247

Kuzmin A. I., Eisensmith R. C., Goltsov A. A., Sergeeva N. A., Schwartz E. I., Woo S. L. (1995): Complete spectrum of PAH mutations in Tataria: presence of Slavic, Turkic and Scandinavian mutations. Eur. J. Hum. Genet. **3,** 246–255

Leuzzi V., Carducci C., Chiarotti F., Artiola C., Giovanniello T., Antonozzi I. (2006): The spectrum of phenylalanine variations under tetrahydrobiopterin load in subjects affected by phenylalanine hydroxylase deficiency. J. Inherit. Metab. Dis. **29,** 38–46
https://doi.org/10.1007/s10545-006-0096-3

Lichter-Konecki U., Hipke C. M., Konecki D. S. (1999): Human phenylalanine hydroxylase gene expression in kidney and other nonhepatic tissues. Mol. Genet. Metab. **67,** 308–316
https://doi.org/10.1006/mgme.1999.2880

Martinez A., Knappskog P. M., Olafsdottir S., Døskeland A. P., Eiken H. G., Svebak R. M., Bozzini M., Apold J., Flatmark T. (1995): Expression of recombinant human phenylalanine hydroxylase as fusion protein in Escherichia coli circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. Biochem. J. **306,** 589–597
https://doi.org/10.1042/bj3060589

Pey A. L., Stricher F., Serrano L., Martinez A. (2007): Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases. Am. J. Hum. Genet. **81,**1006–1024
https://doi.org/10.1086/521879

Polak E., Ficek A., Radvanszky J., Soltysova A., Urge O., Cmelova E., Kantarska D., Kadasi L. (2013): Phenylalanine hydroxylase deficiency in the Slovak population: genotype-phenotype correlations and genotype-based predictions of BH4-responsiveness. Gene **526,** 347–355
https://doi.org/10.1016/j.gene.2013.05.057

Réblová K., Hrubá Z., Procházková D., Pazdírková R., Pouchlá S., Zeman J., Fajkusová L. (2013): Hyperphenylalaninemia in the Czech Republic: genotype-phenotype correlations and in silico analysis of novel missense mutations. Clin. Chim. Acta **419,** 1–10
https://doi.org/10.1016/j.cca.2013.01.006

Reblova K., Kulhanek P., Fajkusova L. (2015): Computational study of missense mutations in phenylalanine hydroxylase. J. Mol. Model. **21,** 70
https://doi.org/10.1007/s00894-015-2620-6

Sali A., Blundell T. L. (1993): Comparative protein modeling by satisfaction of spatial restraints. J. Mol. Biol. **234,** 779–815
https://doi.org/10.1006/jmbi.1993.1626

Stojiljkovic M., Pérez B., Desviat L. R., Aguado C., Ugarte M., Pavlovic S. (2009): The Missense p.S231F phenylalanine hy-

droxylase gene mutation causes complete loss of enzymatic activity in vitro. Protein J. **28,** 294–299
https://doi.org/10.1007/s10930-009-9194-z

Trunzo R., Santacroce R., D'Andrea G., Longo V., De Girolamo G., Dimatteo C., Leccese A., Bafunno V., Lillo V., Papadia F., Margaglione M. (2015): Phenylalanine hydroxylase deficiency in south Italy: Genotype-phenotype correlations, identification of a novel mutant PAH allele and prediction of BH4 responsiveness. Clin. Chim. Acta **450,** 51–55
https://doi.org/10.1016/j.cca.2015.07.014

Trunzo R., Santacroce R., Shen N., Jung-Klawitter S., Leccese A., De Girolamo G., Margaglione M., Blau N. (2016): In vitro residual activity of phenylalanine hydroxylase variants and correlation with metabolic phenotypes in PKU. Gene **594,** 138–143
https://doi.org/10.1016/j.gene.2016.09.015

Tyfield L. A., Stephenson A., Cockburn F., Harvie A., Bidwell J. L., Wood N. A., Pilz D. T., Harper P., Smith I. (1997): Sequence variation at the phenylalanine hydroxylase gene in the British Isles. Am. J. Hum. Genet. **60,** 388–396

Wang Y., DeMayo J. L., Hahn T. M., Finegold M. J., Konecki D. S., Lichter-Konecki U., Woo S. L. (1992): Tissue- and development-specific expression of the human phenylalanine hydroxylase/chloramphenicol acetyltransferase fusion gene in transgenic mice. J. Biol. Chem. **267,** 15105–15110

Waters P. J., Parniak M. A., Nowacki P., Scriver C. R. (1998): In vitro expression analysis of mutations in phenylalanine hydroxylase: linking genotype to phenotype and structure to function. Hum. Mutat. **11,** 4–17
https://doi.org/10.1002/(SICI)1098-1004(1998)11:1<4::AID--HUMU2>3.0.CO;2-L

Waters P. J., Scriver C. R., Parniak M. A. (2001): Homomeric and heteromeric interactions between wild-type and mutant phenylalanine hydroxylase subunits: evaluation of two-hybrid approaches for functional analysis of mutations causing hyper-phenylalaninemia. Mol. Genet. Metab. **73,** 230–238
https://doi.org/10.1006/mgme.2001.3198

Waters P. J. (2003): How PAH gene mutations cause hyper-phenylalaninemia and why mechanism matters: insights from in vitro expression. Hum. Mutat. **21,** 357–369
https://doi.org/10.1002/humu.10197

Zhu T., Qin S., Ye J., Qiu W., Han L., Zhang Y., Gu X. (2010): Mutational spectrum of phenylketonuria in the Chinese Han population: a novel insight into the geographic distribution of the common mutations. Pediatr. Res. **67,** 280–285
https://doi.org/10.1203/pdr.0b013e3181c9fb85