Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: www.tandfonline.com/journals/tbsd20

Identifying and elucidating the roles of Y198N and Y204F mutations in the PAH enzyme through molecular dynamic simulations

Tolga Aslan, Aslı Yenenler-Kutlu, Umut Gerlevik, Ayşe Çiğdem Aktuğlu Zeybek, Ertuğrul Kıykım, Osman Uğur Sezerman & Necla Birgul Iyison

To cite this article: Tolga Aslan, Aslı Yenenler-Kutlu, Umut Gerlevik, Ayşe Çiğdem Aktuğlu Zeybek, Ertuğrul Kıykım, Osman Uğur Sezerman & Necla Birgul Iyison (2022) Identifying and elucidating the roles of Y198N and Y204F mutations in the PAH enzyme through molecular dynamic simulations, Journal of Biomolecular Structure and Dynamics, 40:19, 9018-9029, DOI: 10.1080/07391102.2021.1921619

To link to this article: https://doi.org/10.1080/07391102.2021.1921619

View supplementary material

Published online: 10 May 2021.

Submit your article to this journal

Article views: 416

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tbsd20

JOURNAL OF BIOMOLECULAR STRUCTURE AND DYNAMICS
2022, VOL. 40, NO. 19, 9018–9029
https://doi.org/10.1080/07391102.2021.1921619

Taylor & Francis
Taylor & Francis Group

Check for updates

Identifying and elucidating the roles of Y198N and Y204F mutations in the PAH enzyme through molecular dynamic simulations

Tolga Aslan[a], Aslı Yenenler-Kutlu[b,c], Umut Gerlevik[b], Ayşe Çiğdem Aktuğlu Zeybek[d], Ertuğrul Kıykım[d], Osman Uğur Sezerman[b] and Necla Birgul Iyison[a]

[a]Faculty of Arts and Sciences, Department of Molecular Biology and Genetics, Bogazici University, Istanbul, Turkey; [b]Department of Biostatistics and Medical Informatics, School of Medicine, Acibadem Mehmet Ali Aydinlar University, Istanbul, Turkey; [c]Faculty of Science and Literature, Department of Molecular Biology & Genetics, Istinye University, Istanbul, Turkey; [d]Cerrahpasa Faculty of Medicine, Divisions of Nutrition and Metabolism, Department of Pediatrics, Istanbul University, İstanbul, Turkey

Communicated by Ramaswamy H. Sarma

ABSTRACT

Phenylketonuria is an autosomal recessive disorder caused by mutations in the phenylalanine hydroxylase gene. In phenylketonuria causes various symptoms including severe mental retardation. PAH gene of a classical Phenylketonuria patient was sequenced, and two novel heterozygous mutations, p.Y198N and p.Y204F, were found. This study aimed to reveal the impacts of these variants on the structural stability of the PAH enzyme. *In-silico* analyses using prediction tools and molecular dynamics simulations were performed. Mutations were introduced to the wild type catalytic monomer and full length tetramer crystal structures. Variant pathogenicity analyses predicted p.Y198N to be damaging, and p.Y204F to be benign by some prediction tools and damaging by others. Simulations suggested p.Y198N mutation cause significant fluctuations in the spatial organization of two catalytic residues in the temperature accelerated MD simulations with the monomer and increased root-mean-square deviations in the tetramer structure. p.Y204F causes noticeable changes in the spatial positioning of T278 suggesting a possible segregation from the catalytic site in temperature accelerated MD simulations with the monomer. This mutation also leads to increased root-mean-square fluctuations in the regulatory domain which may lead to conformational change resulting in inhibition of dimerization and enzyme activation. Our study reports two novel mutations in the PAH gene and gives insight to their effects on the PAH activity. MD simulations did not yield conclusive results that explains the phenotype but gave plausible insight to possible effects which should be investigated further with *in-silico* and *in-vitro* studies to assess the roles of these mutations in etiology of PKU.

Abbreviations: BH_4: Tetrahydrobiopterin; ExAC: Exome Aggregation Consortium; gDNA: genomic DNA; HRM: high resolution melting analysis; MD: molecular dynamics; NAMD: Nanoscale Molecular Dynamics; PAH: Phenylalanine hydroxylase; Phe: phenylalanine; PKU: Phenylketonuria; PME: particle-mesh Ewald; RMSD: root-mean-square deviation; RMSF: root-mean-square fluctuation; Rg: radius of gyration; SASA: solvent-accessible surface area; TAMD: temperature accelerated molecular dynamics; Tyr: tyrosine; VMD: Visual Molecular Dynamics; wtPAH: the wild-type PAH enzyme; Y198N-PAH: the PAH enzyme containing Y198N mutation; Y204F-PAH: the PAH enzyme containing Y204F mutation.

ARTICLE HISTORY
Received 16 March 2020
Accepted 20 April 2021

KEYWORDS
Phenylketonuria; phenylalanine hydroxylase; PAH; molecular dynamics simulation; hyperphenylalaninemia; NAMD; VMD

Background

Phenylketonuria (PKU, OMIM 261600) is an autosomal recessive disorder caused by the absence or reduced activity of the hepatic enzyme phenylalanine hydroxylase (PAH, EC 1.14.16.1) (Hanley, 2004). In the presence of its cofactor tetrahidrobiopterin (BH_4), iron, and molecular oxygen, PAH catalyzes the oxidative hydroxylation of L-phenylalanine (Phe) to L-tyrosine (Tyr) (Scriver et al., 1994). In PKU patients, due to decreased PAH activity, Phe level in the blood increases drastically and, if not treated, leads to various symptoms including severe mental retardation, impaired postnatal cognitive development, microcephaly, hypopigmentation, aggressive behavior, attention deficiency, and behavioral abnormalities (Committee on Genetics, 2008; Longo et al., 2014; Scriber, 2001; Williams et al., 2008; Zurflüh et al., 2008). Current standard treatment for PKU is lifelong Phe restricted diet (Blau, 2016). Patients with certain genotypes can benefit from pharmacological BH_4 supplementation, and especially patients with milder forms of PKU experience improvements or even complete recovery in blood Phe levels (Blau et al., 2014). Recently, a PEGylated form of the vegetal enzyme phenylalanine ammonia lyase that converts L-Phe to trans-cinnamic acid and ammonia has been approved for use in adult PKU patients (U.S. Food & Drug Administration, 2018).

CONTACT Necla Birgul Iyison birgul@boun.edu.tr Faculty of Arts and Sciences, Department of Molecular Biology and Genetics, Bogazici University, Istanbul, Turkey

Supplemental data for this article can be accessed online at https://doi.org/10.1080/07391102.2021.1921619.

© 2021 Informa UK Limited, trading as Taylor & Francis Group

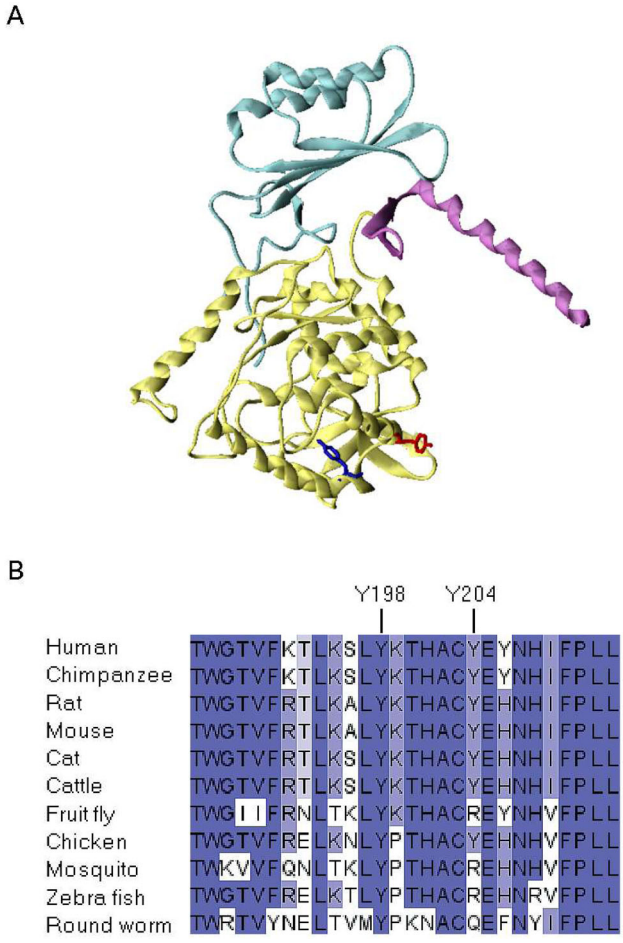

Figure 1. Positional and conformational information for the two tyrosine residues affected in the patient. **(A)** The overall structure of PAH and positions of the mutated tyrosines (Y198 and Y204). Regulatory (cyan), catalytic (yellow) and tetramerization (purple) domains, as well as residues Y198 (blue) and Y204 (red) are shown on human PAH (PDB ID: 6HYC). **(B)** Evolutionary conservation of Tyr 198 and 204 among clades close and distant to human.

The human *PAH* gene is about 90 kb, has 13 exons and produces a 52 kDa PAH monomer. Human PAH enzyme is in equilibrium between homodimer and homotetramer states (Blau et al., 2014). PAH monomer contains three domains: a regulatory domain (residues 1-142), a catalytic domain (residues 143-410), and a multimerization domain (residues 411-542) (Blau et al., 2014) (Figure 1A). PAH monomers form dimers through their multimerization domains, and dimerization of these dimers form the tetramers (Gersting et al., 2010). Disease-causing variations were reported for all exons, with exons 6 and 7 carrying mutations more frequently than others (Blau et al., 2014). More than 1000 mutations in *PAH* have been identified and associated to PKU and other hyperphenylalaninemias (PAHvdb, http://www.biopku.org; HGMD, http://www.hgmd.cf.ac.uk/). Around 98% of PKU cases are caused by mutations in the PAH gene itself, and the remaining 2% is due to mutations in genes responsible for synthesis or turnover of BH_4 (Scriver, 2007). Around 60% of mutations in *PAH* are missense mutations, followed by splice variants (14%) and small or large deletions (14%) (Blau et al., 2014). Most patients (76%) are compound heterozygous for the mutation in PAH (Blau, 2016).

Assessing the correlation between genotype and disease severity and understanding the effects of variants on enzyme confirmation and activity is important to predict clinical phenotype and to enable prognostic approaches. Thanks to technological advancements and open-access databases, it is possible to predict PKU severity and BH_4 responsiveness with acceptable accuracies. Accurate phenotype predictions for 77.9% of compound heterozygous genotypes were reported while BH_4 responsiveness prediction accuracy was 71% (Wettstein et al., 2015). Therefore, enriching these databases with new variant – phenotype correlations and to investigate the effect of these variants on the enzyme's confirmation and stability is still relevant and important.

In a mutation analysis performed on our patient cohort (not published), we have identified four variants in a male classical PKU patient born in June, 2015. The patient was found to carry two novel heterozygous missense variants p.Y198N (c.592T > A) and p.Y204F (c.611A > T), as well as a heterozygous silent variant p.Q232Q (c.696G > A) and a homozygous silent variant p.L385L (c.1155C > G).

This study is aimed to reveal the impacts of Y198N and Y204F variants on the structural stability of the PAH enzyme via molecular dynamics approach. We performed 100 ns MD simulations at physiological temperature (310 K) and at elevated temperature (450 K) with PAH monomers and tetramer carrying novel mutations p.Y198N (Y198N-PAH) and p.Y204F (Y204-PAH). We provide a theoretical basis for the actual reason of PKU disease caused by these mutations by reporting the alterations in the structural integrity and the intramolecular interactions of the PAH enzyme.

Materials and methods

Identification and analysis of variants in PAH

Patients included in the study were diagnosed with PKU via the Turkish National Newborn Metabolic and Endocrine Diseases Screening Program. The study was approved by the Bogazici University Institutional Review Board for Research with Human Subjects (2016/31), and all patients or their guardians gave written informed consent. Patients' blood samples were collected in vacutainer EDTA tubes. Genomic DNA (gDNA) was isolated using Macherey-Nagel Nucleospin Blood kit (catalog no: 740951). High Resolution Melting (HRM) analysis of collected samples was performed using Roche LightCycler® 480 High Resolution Melting Master (catalog no: 04909631001) as described previously (Dobrowolski et al., 2007). Amplicons with an abnormal melting profile were confirmed with Sanger sequencing using same primers. Sequences were then aligned against wild type *PAH* (NM_000277.2).

Novelty of variants were examined in PAHvdb (http://www.biopku.org), HGMD (http://www.hgmd.cf.ac.uk), Exome Aggregation Consortium (ExAC) (Lek et al., 2016), and 1000 Genomes Project (Gibbs et al., 2015) databases. MutationTaster (http://mutationtaster.org) (Schwarz et al., 2014), Polyphen-2 (http://genetics.bwh.harvard.edu/pph2) (Adzhubei et al., 2010), MutationAssessor (http://mutationassessor.org) (Reva et al., 2011), PROVEAN (http://provean.jcvi.

org) (Choi et al., 2012), FATHMM (http://fathmm.biocompute.org.uk) (Shihab et al., 2013), and CADD (https://cadd.gs.washington.edu) (Kircher et al., 2014) were used for pathogenicity analysis of variants.

Amino acid changes corresponding to variants were created manually and analyzed for translation elongation rates and ribosomal densities with RFMapp (Zur & Tuller, 2012).

All variants were analyzed for effects on splicing using Human Splicing Founder (http://www.umd.be/HSF3/) (Desmet et al., 2009).

Evolutionary analysis

For evolutionary analysis, phenylalanine hydroxylase protein sequences for human as well as various organisms in clades close and distant to human were obtained from UniProt (Consortium, 2019). These sequences were then aligned using ClustalX (Larkin et al., 2007).

Structure preparation

Crystal structure of the catalytic domain of the PAH enzyme (PDB ID: 1J8T) and the full tetramer complex of PAH enzymes (PDB ID: 6HYC) (Flydal et al., 2019) were used to create aforementioned mutants. The crystal structures of catalytic domain and full length of PAH enzyme were captured as homo 2-mer together with the iron, and tetramer together with the iron and BH_4 molecule, respectively (Andersen et al., 2001; Flydal et al., 2019). First, the full-length monomer of PAH was modeled via Robetta webserver (Song et al., 2013) by defining the crystal with PDB code of 6HYC (Flydal et al., 2019) as the template. BH_4 and water molecules in the crystal were copied into the model. Fe(II) ion, which is required for PAH activity, was not included in 6HYC, this crystal structure was aligned to 6N1K (Arturo et al., 2019) to dock the Fe(II) ion with its coordinating waters. Cl^- ions in the 6N1K were also kept in the model. Based on the resolved tetrameric state in the 6HYC crystal, we converted our model into tetramer in Visual Molecular Dynamics (VMD) (v1.9.3) (Humphrey et al., 1996) by copying the constructed monomer and aligning with the crystal chains. Then, mutations were introduced to the reference structure using VMD (Humphrey et al., 1996). The mutant model of tetramer was designed as one homodimer carrying Y198N and the other homodimer including Y204F, and these two dimers form the tetramer. The systems were solved in a box of TIP3P model water (Jorgensen et al., 1983). Monomer systems were neutralized with NaCl using the cutoff distance for solute and protein of 5 Å whereas tetramer systems were ionized with 0.15 M KCl solution. The overall charge of all systems was adjusted to zero for mimicking cellular conditions. To determine the protonation states of the amino acids in the protein, we used PROPKA (v3.1) (Olsson et al., 2011; Søndergaard et al., 2011) which calculated E353 and E422 as protonated.

Molecular dynamic simulations

Monomer-PAH and tetramer-PAH systems were simulated with the Nanoscale Molecular Dynamics (NAMD) software using CHARMM 36 m all-atom force fields parameters including the correction maps (Brooks et al., 2009; MacKerell et al., 1998, 2004; Phillips et al., 2005). BH_4 molecule was parametrized as previously described (Floquet et al., 2011) by using GAFF (Wang et al., 2004), Antechamber (Wang et al., 2006) and "amb2chm_par.py" program of Amber2018 (Case et al., 2018). For monomer system, following an 80,000-step minimization with conjugate gradient algorithm, these systems were slowly heated to 298 K from 0 K within 20 ps, and then they were equilibrated at 298 K at 1 atm for 3 ns. For tetramer-PAH system, 5,000-step minimization and 1 ns equilibration were performed by fixing the protein to relax the system. Then, another 5,000-step minimization and 1 ns equilibration were performed without any constraints, except the SHAKE algorithm (Ryckaert et al., 1977) on water molecules, to relax the protein and system. The production simulations were performed along 100 ns trajectory at 310 K collected under NpT ensemble. To sample more conformations related to the impact of mutations in shorter trajectories than obtained from the classical MD simulations, temperature accelerated MD (TAMD) simulations of monomer-PAH systems were also run at 450 K for 100 ns under NvT ensemble. Controls of pressure and temperature were performed with Langevin temperature and pressure coupling. All MD simulations were performed with an integration time-step of 2-fs, and the long-range Coulomb interactions were calculated with particle-mesh Ewald (PME) method (Darden et al., 1993; Essmann et al., 1995). All MD simulations were repeated twice by using different initial points to include any possible energy minima with different conformations. Second initial points were created by increasing the equilibration step at 298 K by 0.2 ns. Averages of both trajectories used for analyses.

Data analysis

VMD was used for both visualization and analysis of MD trajectories (Humphrey et al., 1996). For all systems, backbone root-mean-square deviations (RMSD), average Cα root-mean-square fluctuations (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), dihedral angle measurements, H-bond and salt bridge analyses were performed by using in-house tcl scripts and built-in plugins of VMD. In addition, Cβ distance between catalytic residues and the distance between the Fe(II) ion and coordinated residues highlighted in PDBsum were also measured with VMD (de Beer et al., 2014; Humphrey et al., 1996). Changes in folding free energy due to mutations were calculated by using the last 20 ns in monomer and tetramer trajectories as they were in equilibrium. Configurations of all systems were saved every ns and the effects of mutations on destabilization tendency of the enzyme were calculated by first calculating ΔG (kcal/mole) using FoldX (Schymkowitz et al., 2005)

and then calculating $\Delta\Delta G$ values using the formula: $\Delta\Delta G = \Delta G^{mutant} - \Delta G^{wild-type}$.

Binding free energy changes upon the mutations were predicted between the interfaces A-B, B-D and C-D by using BeAtMuSiC (Dehouck et al., 2013). Of note, the tetramer structure obtained at the end of the 100 ns MD simulations was used as input. In addition, the $\Delta\Delta G$ was calculated by using the formula: $\Delta\Delta G_{bind} = \Delta G_{bind}^{mutant} - \Delta G_{bind}^{wild-type}$

Results

Mutation analysis

High resolution melting analysis and consecutive confirmation with Sanger sequencing in a male classical PKU patient who was diagnosed at birth via the national newborn screening program showed two heterozygous (c.592T > A, p.Y198N and c.611A > T, p.Y204F), a heterozygous silent (c.696G > A, p.Q232Q), and a homozygous silent (c.1155C > G, p.L385L) variant. c.592T > A, c.611A > T, and c.626G > A were in exon 6, c.1155C > G was in exon 11. The patient's blood L-Phe and L-Tyr concentrations at the time of admission were 1464 μM and 54 μM, respectively. A L-Phe-restricted diet was initiated immediately after the diagnosis. He was unresponsive to BH_4 treatment. Neither of the missense mutations were previously reported in PAHvdb, HGMD, ExAC, or 1000 Genomes Project. Silent variants c.696G > A (dbSNP: rs1126758) and c.1155C > G (dbSNP: rs772897) were reported with high allele frequencies of 0.5785 and 0.8582 (Lek et al., 2016), therefore omitted in our analyses.

Tyrosines 198 and 204 reside in α-helices 6 (residues 181-201) and 7 (residues 204-216), respectively (Figure 1A). Y198 is highly conserved (Figure 1B) and variant pathogenicity analysis predicted Y198N mutation to be damaging (Table 1). Y204 is conserved with variations in distant clades (Figure 1A). This mutation replaces polar Tyr with nonpolar Phe and was predicted as damaging by Polyphen-2, PROVEAN and MutationAssessor, and benign by MutationTaster, FATHMM and CADD (Table 1). Translation elongation rates and ribosomal densities of the four mutations were not significantly different from those of wtPAH (data not shown). The patient's parents were unreachable for the paternal and maternal genotyping.

RMSD and RMSF results: monomer

RMSD values were calculated along the MD simulation trajectories (at 310 K) by aligning the conformations observed in the trajectories to the initial model. In the case of monomer structure that represents the catalytic domain of PAH enzyme, RMSD patterns of Y198N and Y204F mutants were similar to the wild-type structure up to 40 ns of all trajectory. After this point, the lowered backbone motion patterns are observed in both mutant enzymes through the end of trajectory.

To observe the further impacts of these mutations on the stability of the enzyme, within a shorter trajectory, we performed TAMD simulations (at 450 K) for 100 ns. Contrary to the observed RMSD values at 310 K, Y198N and Y204F mutations lead to increase in RMSD patterns at elevated temperature (Figure S1A). Y204F mutation caused a higher level of backbone motion, but the protein structures were kept intact in both cases.

In agreement with the backbone RMSD patterns, RMSF patterns of Y198N and Y204F mutants of the PAH enzyme along 100 ns MD trajectories at 310 K show no crucial alterations compared to the wtPAH (Figure 1B). At elevated temperature (450 K), higher fluctuations in Y204F-PAH were observed compared to Y198N-PAH, especially in the C-terminus (Figure S1B).

RMSD and RMSF results: tetramer

Similar to the monomer structure, the wild type and mutant tetramer structures showed similar RMSD patterns (Figure 2A). RMSD values reached ~9 Å towards the end of the 100 ns trajectory. Native structure achieved higher RMSD values compared to the mutant structure between 10 and 20 ns, and mutant structure achieved higher RMSD values compare to the native structure between 40 and 60 ns. Both of these differences, however, were at small scales which indicates no major impacts on the overall structure of the tetramer.

Due to the compound heterozygous nature of detected variants, tetramer structure of PAH is designed as chain A carried N198 and Y204, chain B carried N198 and Y204, chain C carried Y198 and F204, and chain D carried Y198 and F204 to obtain 1:1 ratio between the Y204F and Y198N mutations. RMSD analysis of tetramer structure is also performed for each chain separately (Figure S2). In contrast to the effects of Y198N mutation, presence of Y204F in chains C and D leads to restriction in the motion of backbone Cα atoms (Figure S2B). From 25 ns through the end of MD trajectory, RMSD value of the Y204F mutant is around ~1 Å and lower compared to that in the wild type structure (Figure S2B).

RMSF analysis reveals different levels of fluctuation in N-termini of each chain of the mutant structure compared to the wild type enzyme. Specifically, chain A shows greater fluctuation between residues 1-125, chain B shows less fluctuation between residues 1-25 but higher fluctuation between residues 50-125, chain C shows higher fluctuations between residues 1-25, and chain D shows less fluctuations between residues 1-25 (Figure 2B). Compared to the monomer structure, RMSF patterns of the tetramer complex shows changes in RMSF values between I125 and K150 in chains B and C compared to the wild type tetramer (Figure 2B). Presence of Y198N mutation in chain B and Y204F mutation in chain C leads to a reduction and increment in RMSF values, respectively, between I125 and K150 residues. This particular region is composed of one helix and longer loop structure (Figure S3). I125-K150 region of chain B and C is faced to each other whereas these regions are surface exposed in chain A and D (Figure S3).

Tetramer-PAH complex was analyzed further with SASA and Rg. Presence of mutations lead to increase in surface exposed area (Figure S4). As expected, the presence of

Table 1. Description and *in-silico* analyses results of the novel p.Y198N and p.Y204F mutations.

Genomic location (hg19)	chr12:103249028A > T	chr12:103249009T > A
Codon change	TAT/AAT	TAT/TTT
Transcript change (NM_000277.2)	c.592T > A	c.611T > A
Amino acid change (NP_000268.1)	p.Y198N	p.Y204F
In silico predictions		
Polyphen 2 (HumDiv/HumVar)	Probably damaging / probably damaging	Benign / benign
PROVEAN	Deleterious	Neutral
MutationTaster	Disease causing	Disease causing
MutationAssessor	High functional impact	Low functional impact
FATHMM	Damaging	Damaging
CADD	Among the 1% most deleterious mutations (PHRED 23.1)	Among the 1% most deleterious mutations (PHRED 22.9)
HSF3.1	Probably no impact on splicing	Probably no impact on splicing

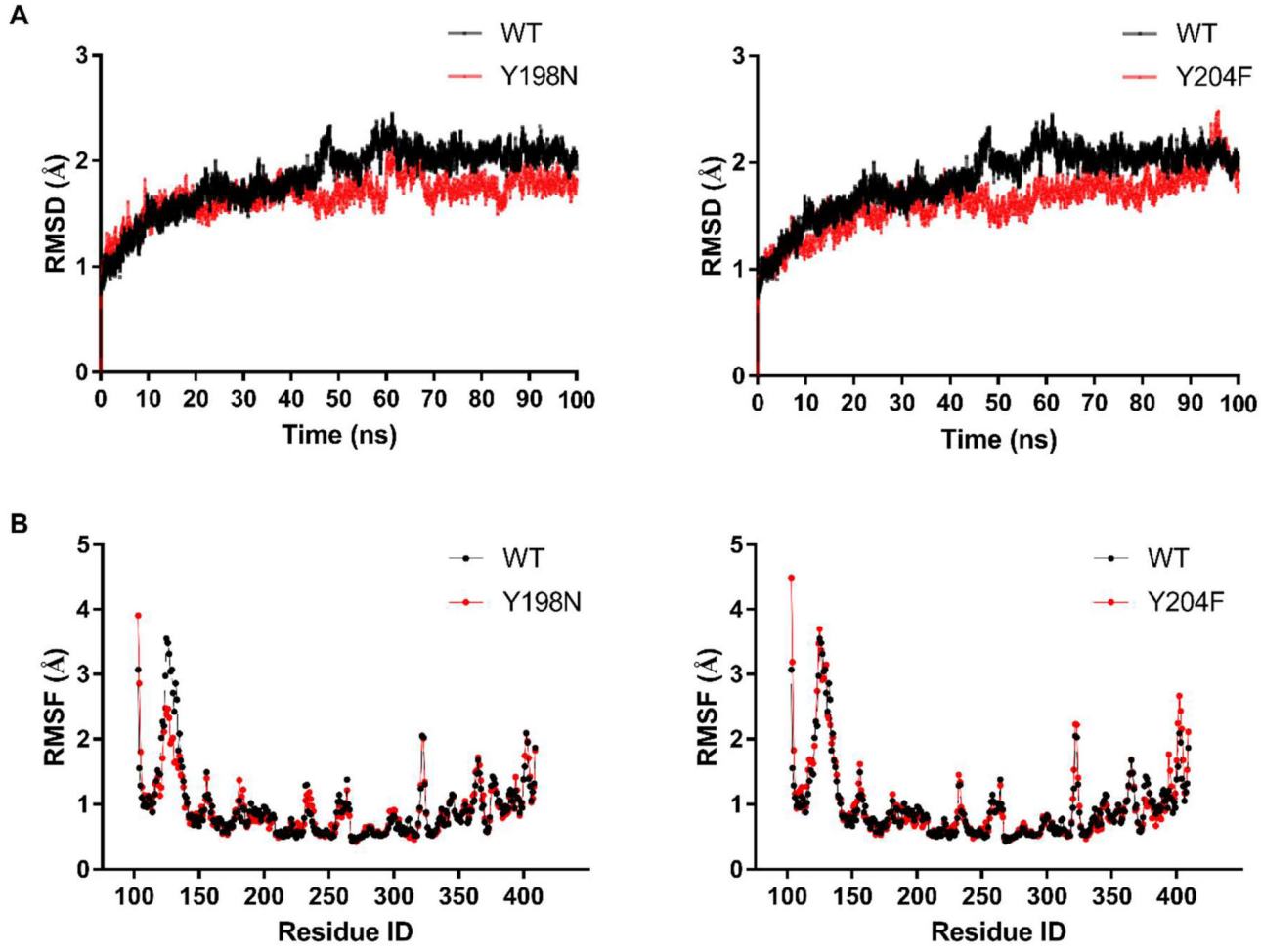

Figure 2. RMSD-backbone (A) and RMSF results (B) of Y198N-PAH and Y204F-PAH monomer structures, respectively, at 310 K along 100 ns MD simulations.

Y198N and Y204F mutations resulted in ~2 Å reduction in Rg of the tetramer structure (Figure S5). Chain based Rg patterns were analyzed in chain-based manner, the slight opening in the compactness has been observed in the group of chain A and B between 40 and 60 ns of MD trajectory (Figure S5) which corresponds to increased tendency in backbone motions of Cα (Figure S2).

Cβ-Cβ measurement in the catalytic site: monomer

To assess the impacts of Y198N and Y204F mutations on the coordination of the catalytic triad, their spatial positions in respect to each other were analyzed by measuring the Cβ-Cβ distances. Our analyses suggest that the presence of Y198N and Y204F mutations do not lead to an unfavorable conformational change or any structural dissociation in the active site of PAH enzyme (Figure 3). When individual effects of Y198N and Y204F mutations on the active site of enzyme are analyzed separately, Y204F mutation results in more fluctuations in R270-S349 and T278-S349 interactions. These fluctuations, however, are marginal and do not change the overall pattern with excess deviations from the preserved conformation.

In TAMD, the conformational stability of the catalytic site is lost upon Y198N and Y204F point mutations (Figure S6). Y204F mutation in particular leads to dramatic increases in Cβ-Cβ distances of up to 14 Å between the catalytic triad

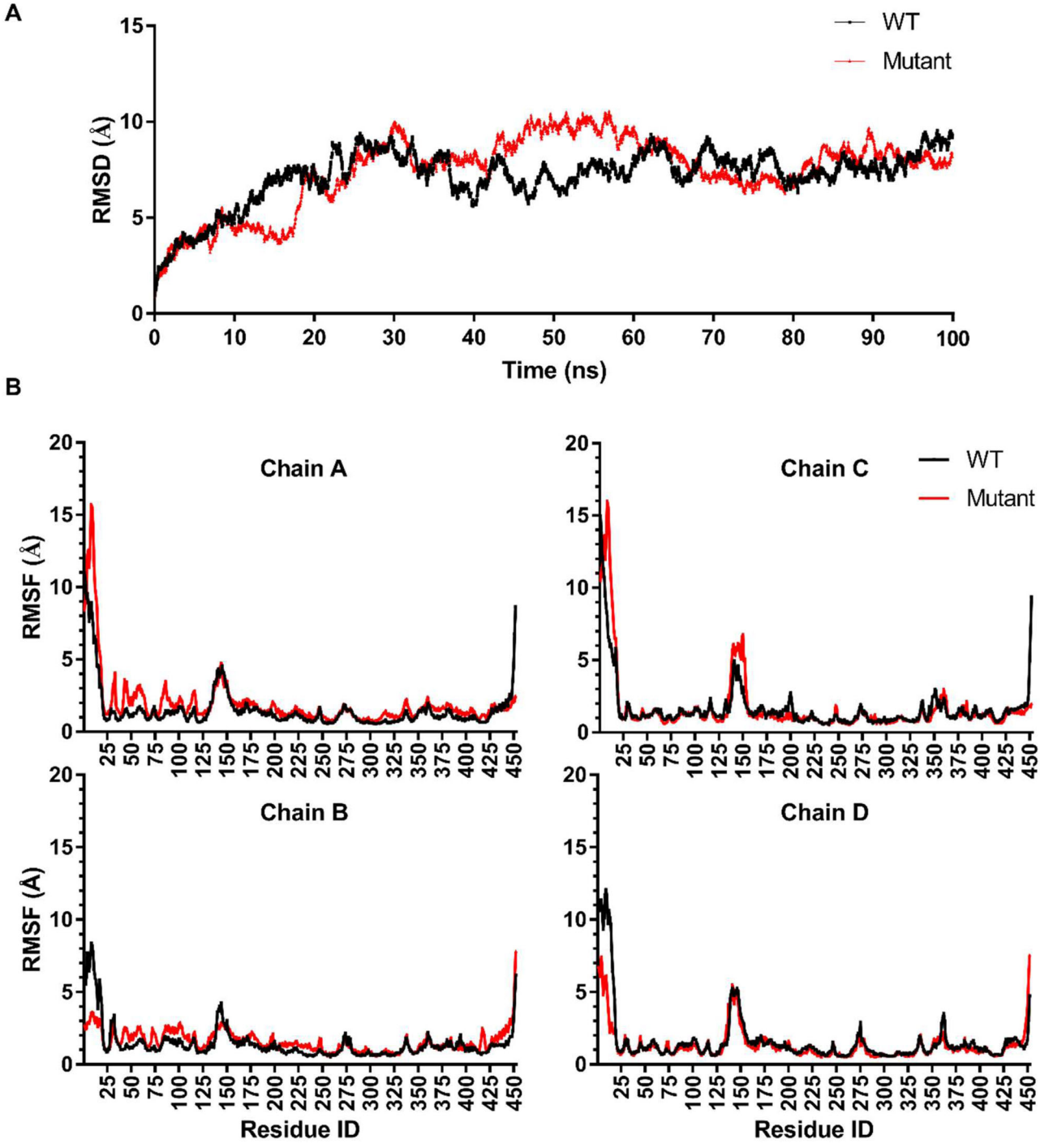

Figure 3. RMSD (A) and RMSF (B) patterns of PAH-tetramer structures along 100 ns MD simulations run at 310 K.

starting from around ~20 ns (Figure S6B). This change is also in line with the backbone RMSD (Figures S1A) and RMSF (Figure S1B) patterns of Y204F-PAH.

Cβ-Cβ measurement in the catalytic site: tetramer

Cβ-Cβ distances of catalytic residues are also measured along 100 ns MD trajectory in the PAH tetramer, and chain-based evaluation was performed. In chain A of the wild type structure, there is a change in R272-T278 distances immediately after 90 ns whereas the distances between the other catalytic residues in this chain remain the same along the trajectory (Figure 4). In chain B of the wild type structure, T278 moves towards S349 between 20 and 50 ns of MD trajectory and then returns back to its initial conformation. When Y198N mutation is introduced in chain B, this particular motion is lost. Moreover, the presence of Y198N mutation in chain B results in loosening of catalytic triad in 1 Å range (Figure 4), but still the overall structure of the catalytic triad is preserved. In chain C, the most noticeable change caused by Y204F mutation is in the Cβ-Cβ distance between T278-S349 leading to an increase in around 2 Å This noticeable increase

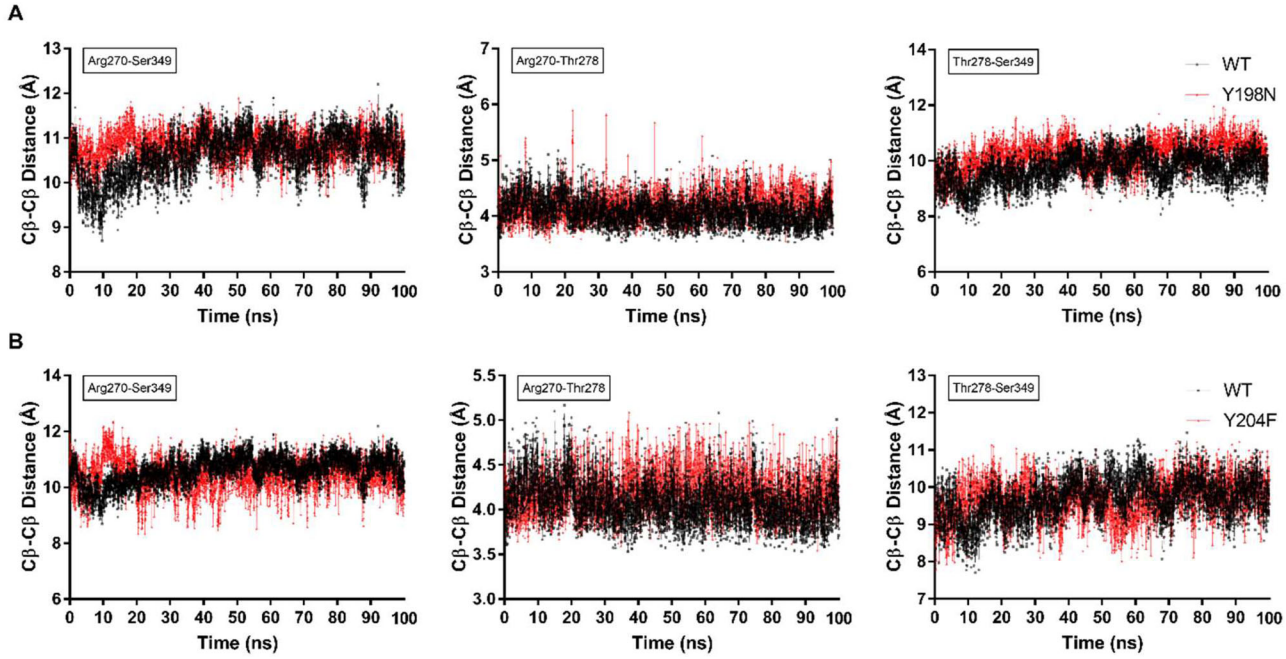

Figure 4. The evolution of Cβ-Cβ distances in the catalytic triad of the (A) Y198N-PAH and (B) Y204F-PAH with respect to native enzyme at 310 K along 100 ns MD simulations.

Table 2. BeAtMuSiC predictions for the binding free energy changes upon the Y198N and Y204F mutations.

Interface	Mutation	ΔΔGbind (kcal/mol)	Solvent accessibility
A-B	Y198N	0.86	8.24%
	Y204F	0.29	67.78%
C-D	Y198N	1.24	1.48%
	Y204F	0.21	69.47%
B-C	Y198N	0.95	7.18%
	Y204F	0.19	74.75%

Solvent accessibilities (in partners) of the mutation sites were also calculated by BeAtMuSiC.

is monitored up to 60 ns and then decreases to values seen in the wild type structure. Although both chains C and D are carrying the same mutation, Cβ-Cβ distance of R270-T278 is almost doubled between 10 and 30 ns in chain D but well-preserved in chain C along the whole MD trajectory. As discussed earlier, similar differences are reported for R270-T278 Cβ-Cβ distances in chain A and B. This finding has demonstrated that the crosstalk between chain B and C together with the presence of Y198N and Y204F mutations, respectively, has brought an impact on the packing of catalytic triad but this particular impact did not cause to disruption of catalytic triad.

Binding free energy change predictions by BeAtMuSiC

Both Y198N and Y204F mutations decreased the binding affinity in all of given interfaces (i.e. A-B, C-D and B-C). Based on BeAtMuSiC calculations, the Y198 residue was not solvent accessible whereas the Y204 amino acid was highly accessible. Moreover, the Y198N variant led to a higher decrease in the binding affinity compared to the Y204F mutation. The predicted ΔΔG values and solvent accessibilities were shown in Table 2.

Salt bridge interactions: monomer & tetramer

Salt bridge interactions with 5 Å maximum distance were identified as D282:R270, E280:R158, K195:E221, and K195:E214 for Y198N mutant, and as E390:K335 and D338:K335 for Y204F mutant. In monomer-PAH enzyme, we report that Y198N mutation does not lead to disruption of these listed salt bridge interactions (Figure S7). Specifically, D282:R270 interaction could be crucial for the wtPAH enzyme since R270 is a member of the catalytic triad. Neither Y198N nor Y204F mutations affect the conformational integrity in the active site of PAH along 100 ns of MD trajectory at 310 K (Figure 3). In support of this observation, D282:R270 salt bridge interaction is conserved along the whole MD trajectory at 310 K. Even during TAMD simulations, salt bridge interactions were not affected (Figure S8). Similar to Y198N mutant, Tyr to Phe change at the 204th position does not lead to drastic alterations in salt bridge interactions, E390: K335 and D338: K335, in a proximity of 5 Å (Figure S9). The dynamic properties of these interactions are preserved in the Y204-PAH and wtPAH enzymes, and they are eventually convergent to similar distance values (Figure S9). Moreover, all these listed interactions are also monitored for tetramer-PAH complex over 100 ns MD trajectory to reveal any impacts of mutations in chains A, B, C and D. Among the listed salt bridge interactions, E280:R158 is both closely positioned to the interaction region between chain B and C, and is situated in close proximity to Y198 residue. For both chain B and C, we observe a common pattern that this particular interaction has strengthened in the presence of mutations (Figure S12). Also, a similar tendency was observed in E280:R158 interaction in chains A and D but the effects seem to be negligible compared to those observed in chains B and C (Figure S12). K195:E221 salt bridge shows higher level of fluctuations in the presence of Y198N mutation in chains

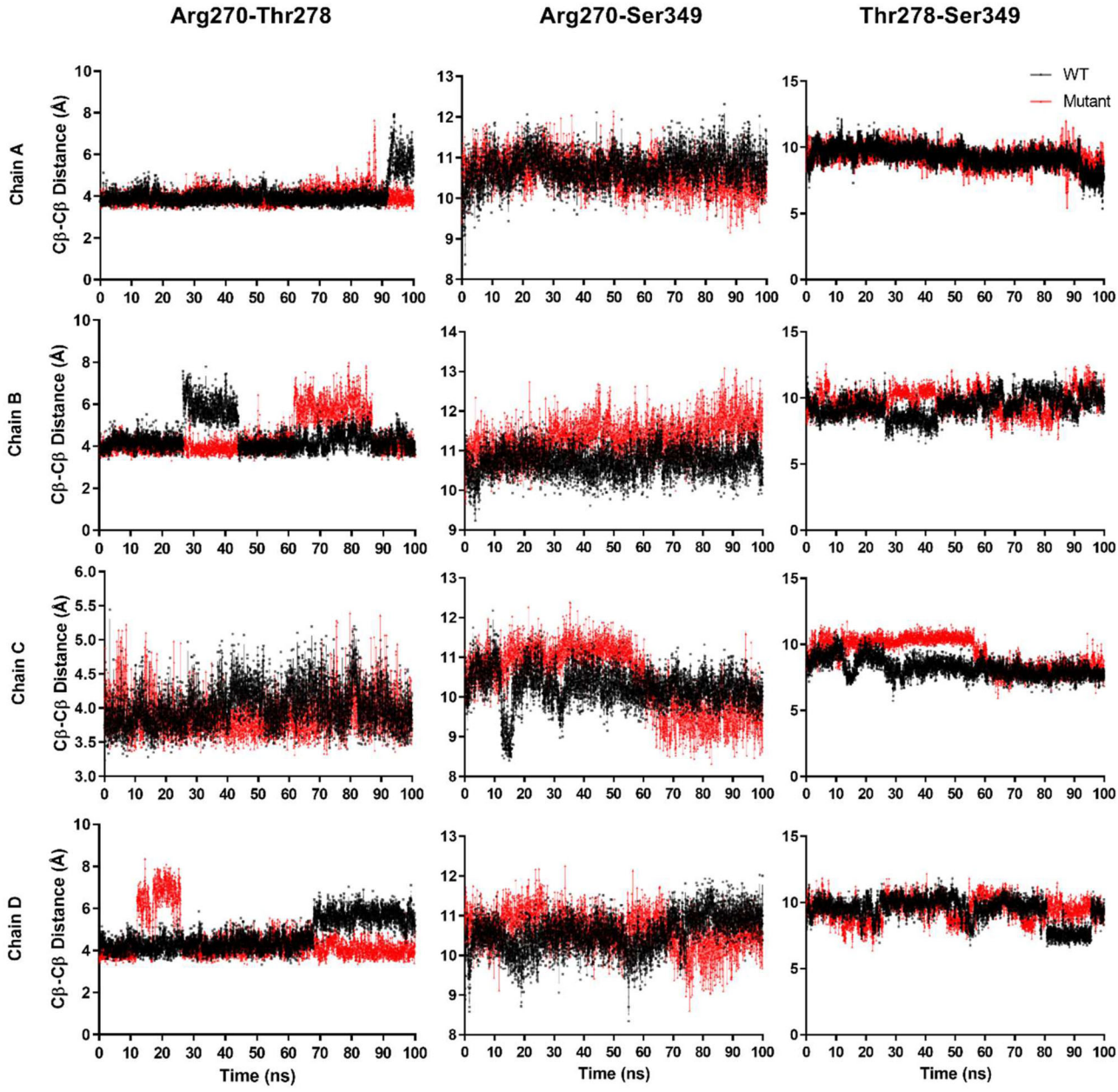

Figure 5. The evolution of Cβ-Cβ distances in the catalytic triad of tetramer-PAH structures in chains A, B, C and D along 100 ns MD simulations at 310 K.

A and B (Figure S13). Contrary to our expectations, this particular interaction follows a different pattern in chains C and D even though both chains carry the same mutation. This interaction is stronger and more stable compared to that in chain D (Figure S13). This difference could be the effect of the crosstalk between chains B and C. The restrictions caused by this crosstalk could impose additional restrictions and changes in intramolecular interactions may not be easily tolerated. Salt bridge interactions of K195:E214 are strengthened in the presence of Y198N mutation in chain A and B. These interactions, however, converge to same strength as the wild type structure at the end of MD trajectory (Figure S14). Even though the E390:K335 salt bridge interaction is within 5 Å proximity of the Y204 residue, the presence of relatively distant Y198N mutation in chain B results in the disruption of stability in this interaction (Figure S13). For the rest of interactions in all chains, there are no significant findings to report.

Fe(II) ion coordination & FoldX results: monomer & tetramer

Presence of neither Y198N nor Y204F mutation in monomer-PAH enzyme seems to affect the Fe(II) coordination. The coordination of Fe(II) and BH_4 is well preserved along MD trajectories of the tetramer-PAH enzyme as well (data not shown).

To reveal the impacts of Y198N and Y204F mutations on the destabilization tendency of PAH, the enzyme configurations were saved every 10 ns through 100 ns of MD trajectories at 310 K and the destabilization tendencies were documented by calculating ΔΔG (kcal/mole) with FoldX

(Schymkowitz et al., 2005). For Y198N and Y204F mutations, we recorded only small fluctuations in ΔΔG values ($\Delta\Delta G_{Y198N}$: 3.18 kcal/mole, $\Delta\Delta G_{Y204F}$: 6.52 kcal/mole) corresponding to 2.7% and 5.45% change compared to ΔG_{wt}, respectively. This is also supported by RMSD and RMSF results of Y198N-PAH and Y204F-PAH cases along 100 ns trajectory at 310 K (Figure 5). Similarly, there are no discernable differences between the ΔΔG values of the wild type tetramer and the mutant tetramer ($\Delta\Delta G_{tetramer}$: 36.24 kcal/mole, 4.2% change from ΔG_{wt}) suggesting that the introduction of Y198N and Y204F mutations does not pose any significant effects on the overall stability and structure of the protein, which is also supported by RMSD and RMSF results (Figure 2).

Discussion

In a mutational screen, we identified four sequence variants in a male classical PKU patient. Two novel heterozygous missense variants p.Y198N (c.592T > A) and p.Y204F (c.611A > T) were discovered, as well as previously reported heterozygous silent variant p.Q232Q (c.696G > A) and a homozygous silent variant p.L385L (c.1155C > G). To find out if these mutations contribute to PKU disease, we performed *in-silico* analyses using online prediction tools and molecular dynamics simulations.

Effects of two novel variants of the *PAH* gene, c.592T > A (p.Y198N) and c.611A > T (p.Y204F), on the conformation and stability of the human PAH enzyme were analyzed via molecular dynamic simulations on human PAH catalytic domain (PDB ID: 1J8T) and the full-length human PAH in tetramer-complex with its cofactor BH_4 (PDB ID: 6HYC). Both mutations are heterozygous and parental genotypes were not available. Therefore, these mutations were handled separately.

Although the tyrosine at position 198 in PAH is a highly conserved residue, Y198N mutation showed no significant effect on the overall structure and stability of the catalytically active monomer probably due to the polar-to-polar change. This mutation affected neither formation or disruption of salt bridges, nor spatial organization of the catalytic residues and the iron ion at physiological temperature. At elevated temperatures, however, the distance between two residues of the catalytic triad (T278-R270) showed significant fluctuations at around 15 ns as an implication of possible loss of the catalytic activity. This fluctuation is recovered towards the 50 ns but reemerged around ~90 ns. Since this mutation is present in a clinically diagnosed PKU patient with no response to the biopterin treatment, we initially speculated that this mutation may have an effect on multimerization, recruitment of BH_4 and/or Phe, mRNA decay and/or protein levels. The crystal structure used in this analysis (PDB ID: 1J8T) lacks the tetramerization and regulatory domains, as well as co-crystalized BH_4. We took a similar approach similar approach using a full-length human PAH in complex with BH_4 which has been published during the this work (PDB ID: 6HYC) (Flydal et al., 2019). The analyses performed with the full-length PAH tetramer revealed minor effects on RMSF values of individual chains. RMSD values in individual chains carrying Y198N mutation showed around 2-fold increase in deviations in the last half of the trajectory. These changes in individual chains do not affect the overall RMSD values of the tetramer, indicating no significant impact on the enzyme's structure. These analyses also did not show dissociation of the tetramer structure or changes in BH_4 or Fe(II) interactions.

Y204F mutation in the PAH monomer results in higher RMSD and RMSF values towards the C-terminus compared to the wild-type enzyme at elevated temperatures. This pattern suggested a tendency for dissociation of the C-terminus of the protein and eventual disruption of multimerization leading to the enzyme inactivation. Analyses performed with full-length PAH tetramer did not show any tendency for dissociation in the C-terminus. This mutation also causes a significant increase in $C\beta$-$C\beta$ distances between T278-R270 and S349-T278 but not between R270-S349 in the monomer structure, suggesting segregation of T278 from the catalytic site. Similar effects are also observed in the tetramer in chains A, C and D for T278-R270, and in chains B, C and D for S349-T278. T278 resides at the entrance of the active site and forms a hydrogen bond with E280 (Guldberg et al., 1993). Various mutations, including T278N are speculated to disrupt this hydrogen bond and, in effect, the entrance of Phe to the active site (Erlandsen et al., 2003). The observed segregation of T278 causes its distance to E280 to fluctuate (Figure S17 and S18) which may affect the hydrogen bond with E280 and may disrupt the catalytic activity of the PAH enzyme.

SASA analyses show that Y204F mutation leads to a dramatic increase in solvent accessible surface area compared to Y198N. Even though the presence of Y204F mutation leads to restriction in backbone motion, this striking increase in SASA value suggests a higher level of reorganization in intramolecular interactions within chain C and D compared to the rest of the structure.

PAH regulatory domain (RD) resides at the N-terminal of the enzyme. Although the exact mechanisms through which the RD modulate the activity of the enzyme, it shows an auto-inhibitory effect as evidenced by constitutive activity of the enzyme upon its deletion (Knappskog et al., 1996). Recent publications suggest a possible binding of Phe to the RD (Gjetting et al., 2001; Li et al., 2011; Roberts et al., 2014; Zhang et al., 2014). This binding is reported to cause dimerization (Patel et al., 2016; Zhang et al., 2014) and activation of the PAH enzyme (Roberts et al., 2014). The proposed activation model suggests a major conformational change which results in removal of the steric hindrance on the catalytic domain on each chain caused by their respective RDs and providing access to the catalytic sites (Patel et al., 2016). In our study, RMSF patterns of the mutant enzyme show differences in all chains compared to the wild type enzyme. These changes in the RMSF patterns may pose a negative effect on the binding of Phe to the RD and/or on the proposed conformational change leading to the inhibition of dimerization and enzyme activation. The approach taken by Patel et al. (2016) may be applied to the novel mutations reported in this study to reveal their effects on the regulatory domain.

According to the *in-silico* results, individual novel mutations reported in this study does not seem to show immediate devastating effects on the stability and overall conformation of the PAH enzyme. However, since the mutation data are obtained from a clinically diagnosed classical PKU patient, it is evident that these two novel mutations cause severe malfunction in the enzyme leading to the observed phenotype. Since most of the PKU patients are compound heterozygous, the effect of compound heterozygosity on the function of the PAH tetramer is investigated by Shen et al. Shen et al., Shen et al., (2016). Dubbed "interallelic complementation", this phenomenon is shown to cause milder (PAH hetero-tetramer activity is up to 51.1% higher than that of either homo-tetramers) or more severe (PAH hetero-tetramer activity is up to 55.2% lower than that of either homo-tetramers) phenotypes than expected solely based on the individual mutations on either allele. The compound heterozygous nature of this patient may have led to a negative complementation resulting in more severe effects than either mutation could cause alone.

Conclusions

Two novel mutations, p.Y198N and p.Y204F, leading to classical PKU are reported in this study. *In-silico* analyses performed in this work shows that these mutations do not pose any discernable impact on the catalytic site of the PAH enzyme. However, their effects on the regulatory domain suggest a possible disruption in allosteric regulation leading to disruption of the enzyme activation. The results also suggest that the hydrogen bond between T278 and E280 may be affected, which may restrict the accessibility of the catalytic domain to the substrate. Further *in-vitro* experiments are required to investigate the effects of these individual mutations on PAH activity as well as possible negative interallelic complementation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was funded by Bogazici University Scientific Research Projects (11480). Funding body had no role in the design of the study, data collection, analysis, interpretation of data, writing the manuscript or decision to publish.

ORCID

Tolga Aslan http://orcid.org/0000-0002-0728-0880
Aslı Yenenler-Kutlu http://orcid.org/0000-0002-9169-388X
Umut Gerlevik http://orcid.org/0000-0003-4288-6340
Ayşe Çiğdem Aktuğlu Zeybek http://orcid.org/0000-0001-7256-0750
Ertuğrul Kıykım http://orcid.org/0000-0001-5170-5594
Osman Uğur Sezerman http://orcid.org/0000-0003-0905-6783
Necla Birgul Iyison http://orcid.org/0000-0003-3511-3139

Availability of data and materials

All of the data and materials used in this study are included in this article and its supplementary files.

Authors' contributions

TA performed DNA isolation from patient sample, analysis of Sanger sequencing. TA, AYK and UG performed all remaining experiments, prepared figures and tables, and drafted the manuscript. AÇAZ and EK diagnosed the patient, measured blood Phe and Tyr levels and provided clinical data. TA, AYK, OUS and NBI designed the study and contributed to data analysis. All authors read and approved the final version of the manuscript.

Consent for publication

Not applicable.

Ethics approval and consent to participate

The study was approved by the Bogazici University Institutional Review Board for Research with Human Subjects (2016/31), and all patients or their guardians gave written informed consent.

References

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*, 248–249. https://doi.org/10.1038/nmeth0410-248

Andersen, O. A., Flatmark, T., & Hough, E. (2001). High resolution crystal structures of the catalytic domain of human phenylalanine hydroxylase in its catalytically active Fe(II) form and binary complex with tetrahydrobiopterin. *Journal of Molecular Biology*, *314*, 279–291. https://doi.org/10.1006/jmbi.2001.5061

Arturo, E. C., Gupta, K., Hansen, M. R., Borne, E., & Jaffe, E. K. (2019). Biophysical characterization of full-length human phenylalanine hydroxylase provides a deeper understanding of its quaternary structure equilibrium. *Journal of Biological Chemistry*, *294*, 10131–10145. https://doi.org/10.1074/jbc.RA119.008294

Blau, N. (2016). Genetics of phenylketonuria: Then and now. *Human Mutation*, *37*(6), 508–515. https://doi.org/10.1002/humu.22980

Blau, N., Shen, N., & Carducci, C. (2014). Molecular genetics and diagnosis of phenylketonuria: State of the art. *Expert Review of Molecular Diagnostics*, *14*, 655–671. https://doi.org/10.1586/14737159.2014.923760

Brooks, B. R., Brooks, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., & Caflisch, A. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, *30*, 1545–1614. https://doi.org/10.1002/jcc.21287

Case, D. A., Ben-Shalom, I., Brozell, S. R., Cerutti, D. S., Cheatham, III, T. E., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Ghoreishi, D., Gilson, M. K., Gohlke, H., Goetz, A. W., Greene, D., Harris, R., Homeyer, N., Huang, Y., Izadi, S., Kovalenko, A., Kurtzman, T., … Kollman, P. A. (2018). *AMBER 2018*. University of California.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, *7*, e46688. https://doi.org/10.1371/journal.pone.0046688

Committee on Genetics. (2008). Maternal phenylketonuria. *Pediatrics*, *122*, 445–449. https://doi.org/10.1542/peds.2008-1485

Consortium, T. U. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, *47*, D506–D515. https://doi.org/10.1093/nar/gky1049

Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems. *Journal of Chemical Physics*, *98*(12), 10089–10092. https://doi.org/10.1063/1.464397

de Beer, T. A. P., Berka, K., Thornton, J. M., & Laskowski, R. A. (2014). PDBsum additions. *Nucleic Acids Research*, *42*, D292–6. https://doi.org/10.1093/nar/gkt940

Dehouck, Y., Kwasigroch, J. M., Rooman, M., & Gilis, D. (2013). BeAtMuSiC: Prediction of changes in protein–protein binding affinity on mutations. *Nucleic Acids Research*, *41*, W333–9. https://doi.org/10.1093/nar/gkt450

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., & Béroud, C. (2009). Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Research*, *37*, e67–e67. https://doi.org/10.1093/nar/gkp215

Dobrowolski, S. F., Ellingson, C., Coyne, T., Grey, J., Martin, R., Naylor, E. W., Koch, R., & Levy, H. L. (2007). Mutations in the phenylalanine hydroxylase gene identified in 95 patients with phenylketonuria using novel systems of mutation scanning and specific genotyping based upon thermal melt profiles. *Molecular Genetics and Metabolism*, *91*, 218–227. https://doi.org/10.1016/j.ymgme.2007.03.010

Erlandsen, H., Patch, M. G., Gamez, A., Straub, M., & Stevens, R. C. (2003). Structural studies on phenylalanine hydroxylase and implications toward understanding and treating phenylketonuria. *Pediatrics*, *112*, 1557–1565.

Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *Journal of Chemical Physics*, *103*(19), 8577–8593. https://doi.org/10.1063/1.470117

Floquet, N., Hernandez, J.-F., Boucher, J.-L., & Martinez, J. (2011). L-arginine binding to human inducible nitric oxide synthase: An antisymmetric funnel route toward isoform-specific inhibitors? *Journal of Chemical Information and Modeling*, *51*, 1325–1335. https://doi.org/10.1021/ci100422v

Flydal, M. I., Alcorlo-Pagés, M., Johannessen, F. G., Martínez-Caballero, S., Skjaerven, L., Fernandez-Leiro, R., Martinez, A., & Hermoso, J. A. (2019). Structure of full-length human phenylalanine hydroxylase in complex with tetrahydrobiopterin. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 11229–11234. https://doi.org/10.1073/pnas.1902639116

Gersting, S. W., Staudigl, M., Truger, M. S., Messing, D. D., Danecka, M. K., Sommerhoff, C. P., Kemter, K. F., & Muntau, A. C. (2010). Activation of phenylalanine hydroxylase induces positive cooperativity toward the natural cofactor. *Journal of Biological Chemistry*, *285*, 30686–30697. https://doi.org/10.1074/jbc.M110.124016

Gibbs, R. A., Boerwinkle, E., Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J. G., Zhu, Y., Wang, J., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin, X., Lan, T., et al. (2015). A global reference for human genetic variation. *Nature*, *526*, 68–74. https://doi.org/10.1038/nature15393

Gjetting, T., Petersen, M., Guldberg, P., & Güttler, F. (2001). Missense mutations in the N-terminal domain of human phenylalanine hydroxylase interfere with binding of regulatory phenylalanine. *American Journal of Human Genetics*, *68*, 1353–1360. https://doi.org/10.1086/320604

Guldberg, P., Romano, V., Ceratto, N., Bosco, P., Cluna, M., Indellcato, A., Mollica, F., Mell, C., Glovannini, M., Riva, E., & Biasuccl, G. (1993). Mutational spectrum of phenylalanine hydroxylase deficiency in Sicily: Implications for diagnosis of hyperphenyl-alaninemia in Southern Europe. *Human Molecular Genetics*, *2*, 1703–1707. https://doi.org/10.1093/hmg/2.10.1703

Hanley, W. B. (2004). Adult phenylketonuria. *The American Journal of Medicine*, *117*, 590–595. https://doi.org/10.1016/j.amjmed.2004.03.042

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, *79*(2), 926–935. https://doi.org/10.1063/1.445869

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*, 310–315. https://doi.org/10.1038/ng.2892

Knappskog, P. M., Flatmark, T., Aarden, J. M., Haavik, J., & Martínez, A. (1996). Structure/function relationships in human phenylalanine hydroxylase. Effect of terminal deletions on the oligomerization, activation and cooperativity of substrate binding to the enzyme. *European Journal of Biochemistry*, *242*, 813–821. https://doi.org/10.1111/j.1432-1033.1996.0813r.x

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., & Thompson, J. D. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., & Tukiainen, T. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*, 285–291. https://doi.org/10.1038/nature19057

Li, J., Ilangovan, U., Daubner, S. C., Hinck, A. P., & Fitzpatrick, P. F. (2011). Direct evidence for a phenylalanine site in the regulatory domain of phenylalanine hydroxylase. *Archives of Biochemistry and Biophysics*, *505*, 250–255. https://doi.org/10.1016/j.abb.2010.10.009

Longo, N., Harding, C. O., Burton, B. K., Grange, D. K., Vockley, J., Wasserstein, M., Rice, G. M., Dorenbaum, A., Neuenburg, J. K., Musson, D. G., Gu, Z., & Sile, S. (2014). Single-dose, subcutaneous recombinant phenylalanine ammonia lyase conjugated with polyethylene glycol in adult patients with phenylketonuria: An open-label, multicentre, phase 1 dose-escalation trial. *The Lancet*, *384*(9937), 37–44. https://doi.org/10.1016/S0140-6736(13)61841-3

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., & Joseph-McCarthy, D. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins[†]. *Journal of Physical Chemistry B*, *102*, 3586–3616. https://doi.org/10.1021/jp973084f

MacKerell, A. D., Feig, M., & Brooks, C. L. (2004). Improved treatment of the protein backbone in empirical force fields. *Journal of the American Chemical Society*, *126*, 698–699. https://doi.org/10.1021/ja036959e

Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M., & Jensen, J. H. (2011). PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. *Journal of Chemical Theory and Computation*, *7*, 525–537. https://doi.org/10.1021/ct100578z

Patel, D., Kopec, J., Fitzpatrick, F., McCorvie, T. J., & Yue, W. W. (2016). Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Scientific Reports*, *6*, 23748. https://doi.org/10.1038/srep23748

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., & Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*, 1781–1802. https://doi.org/10.1002/jcc.20289

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: Application to cancer genomics. *Nucleic Acids Research*, *39*, e118–e118. https://doi.org/10.1093/nar/gkr407

Roberts, K. M., Khan, C. A., Hinck, C. S., & Fitzpatrick, P. F. (2014). Activation of phenylalanine hydroxylase by phenylalanine does not require binding in the active site. *Biochemistry*, *53*, 7846–7853. https://doi.org/10.1021/bi501183x

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341. https://doi.org/10.1016/0021-9991(77)90098-5

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, *11*, 361–362. https://doi.org/10.1038/nmeth.2890

Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., & Serrano, L. (2005). The FoldX web server: An online force field. *Nucleic Acids Research*, *33*, W382–W388. https://doi.org/10.1093/nar/gki387

Scriber, C. R. (2001). Hyperphenylalaninemia: Phenylalanine hydroxylase deficiency. In C. R. Scriver, A. Beaudet, W. Sly, D. Valle, B. Childs, & K. Kinzler (eds.), *Metabolic and molecular bases of inherited disease* (pp. 1667–1724). McGraw-Hill. https://doi.org/10.1036/ommbid.97.

Scriver, C. R. (2007). The PAH gene, phenylketonuria, and a paradigm shift. *Human Mutation*, *28*, 831–845. https://doi.org/10.1002/humu.20526

Scriver, C. R., Eisensmith, R. C., Woo, S. L., & Kaufman, S. (1994). The hyperphenylalaninemias of man and mouse. *Annual Review of Genetics*, *28*, 141–165. https://doi.org/10.1146/annurev.ge.28.120194.001041

Shen, N., Heintz, C., Thiel, C., Okun, J. G., Hoffmann, G. F., & Blau, N. (2016). Co-expression of phenylalanine hydroxylase variants and effects of interallelic complementation on in vitro enzyme activity and genotype-phenotype correlation. *Molecular Genetics and Metabolism*, *117*, 328–335. https://doi.org/10.1016/j.ymgme.2016.01.004

Shihab, H. A., Gough, J., Cooper, D. N., Stenson, P. D., Barker, G. L. A., Edwards, K. J., Day, I. N., & Gaunt, T. R. (2013). Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. *Human Mutation*, *34*, 57–65. https://doi.org/10.1002/humu.22225

Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M., & Jensen, J. H. (2011). Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. *Journal of Chemical Theory and Computation*, *7*, 2284–2295. https://doi.org/10.1021/ct200133y

Song, Y., DiMaio, F., Wang, R. Y.-R., Kim, D., Miles, C., Brunette, T., Thompson, J., & Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure*, *21*, 1735–1742. https://doi.org/10.1016/j.str.2013.08.005

U.S. Food and Drug Administration. (2018). *FDA approves a new treatment for PKU, a rare and serious genetic disease*. https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm608835.htm

Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, *25*, 247–260. https://doi.org/10.1016/j.jmgm.2005.12.005

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general amber force field. *Journal of Computational Chemistry*, *25*, 1157–1174. https://doi.org/10.1002/jcc.20035

Wettstein, S., Underhaug, J., Perez, B., Marsden, B. D., Yue, W. W., Martinez, A., & Blau, N. (2015). Linking genotypes database with locus-specific database and genotype–phenotype correlation in phenylketonuria. *European Journal of Human Genetics*, *23*, 302–309. https://doi.org/10.1038/ejhg.2014.114

Williams, R. A., Mamotte, C. D. S., & Burnett, J. R. (2008). Phenylketonuria: An inborn error of phenylalanine metabolism. *The Clinical Biochemist Reviews*, *29*, 31–41.

Zhang, S., Roberts, K. M., & Fitzpatrick, P. F. (2014). Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase. *Biochemistry*, *53*, 6625–6627. https://doi.org/10.1021/bi501109s

Zur, H., & Tuller, T. (2012). RFMapp: Ribosome flow model application. *Bioinformatics*, *28*, 1663–1664. https://doi.org/10.1093/bioinformatics/bts185

Zurflüh, M. R., Zschocke, J., Lindner, M., Feillet, F., Chery, C., & Burlina, A. (2008). Molecular genetics of tetrahydrobiopterin-responsive phenylalanine hydroxylase deficiency. *Human Mutation*, *29*, 167–175. https://doi.org/10.1002/humu.20637