Journal of **A**pplied **B**iotechnology **R**eports

Original Article

Effects of missense R84Q mutation on human Pyrroline-5-carboxylate synthase enzyme properties, a*n in-silico* analysis

Maryam Zare[1], Faranak Hadi[2], Zarrin Minuchehr[3], Jafar Amani[4], Ali Hatef Salmanian[*5]

Abstract

Mammalian Δ-(1)-Pyrroline-5-carboxylate synthase (P5CS) enzyme catalyzes the coupled phosphorylation and reduction-conversion of glutamate to Δ-(1)-pyrroline-5-carboxylate (P5C), a critical step in the proline, ornithine, citrulline and arginine biosynthesis. In plants and mammals, P5CS consists of two separate enzymatic domains: N-terminal γ-glutamyl kinase (γ-GK) and C-terminal γ-glutamyl phosphate reductase (γ–GPR). Hyperammonemia has been reported as a new inborn disorder, with a range of clinical symptoms which is associated with a reduced synthesis of proline, ornithine, citruline and arginine. A missense mutation, R84Q, which alters the conserved residue in γ-GK domain, is responsible for this disorder. In this study using *in-silico* approaches as a new bioinformatics method, sequence analysis was performed and the tertiary structure of γ-GK domain of human P5CS, which includes the R84Q missense mutation, was predicted and the mutation effects on structural and functional features of P5CS enzyme were analyzed. Our analysis showed that this substitution has an affect on the molecular surface accessibility and total energy of the modeled structure. We conclude that this mutation results in a reduced activity of P5CS enzyme and an impaired synthesis of these amino acids.

1. Department of Biology, Payame Noor University, Avaj, Iran
2. Department of Biology, Faculty of Science, Lorestan University, Khoramabad, Iran.
3. Industrial and Environmental Biotechnology Department, National Institute of Genetic Engineering & Biotechnology (NIGEB), Tehran, Iran.
4. Applied Microbiology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran.
5. Department of Plant Biotechnology, National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran

* **Corresponding Author**
Ali Hatef Salmanian
Department of plant biotechnology, National Institute of Genetic Engineering and Biotechnology, (NIGEB), Shahrak-e-Pajoohesh, 15thKm, Tehran-Karaj Highway, Tehran, Iran.
E.mail: salman@nigeb.ac.ir

Keywords: Pyrroline-5-Carboxylate Synthase, Amino Acid Substitutions, Tertiary Structure Modeling

Submission Date: 2/15/2014
Accepted Date: 3/12/2014

Introduction

P5CS is a bifunctional, mitochondrial inner membrane, ATP- and NADPH-dependent enzyme, which catalyzes the reduction of glutamate to P5C. This reaction is critical for *de novo* biosynthesis of proline, ornithine and arginine [1-3]. In this metabolic pathway, glutamate is converted by γ-GK to γ-glutamyl phosphate, which is reduced by γ-GPR to γ-glutamyl semialdehyde. This product is converted spontaneously to P5C, which is reduced by NADPH to proline by P5C reductase. In addition, P5C is metabolized to ornithine and ornithine to citrulline and arginine in urea cycle [4]. In prokaryotes and lower eukaryotes like *Saccharomyces cerevisiae*, the combined functions of separate γ-GK and γ-GPR enzymes activity is required [3, 5]. In plants and mammals, P5CS has a γ-GK domain (EC 2.7.2.11) at the N terminus and a γ-GPR domain (EC 1.2.1.41) at the C terminus which catalyzes the first two steps of the pathway [6, 7]. In proline biosynthesis, γ-GK enzyme transfers the terminal phosphoryl group of ATP to the gamma-carboxyl group of glutamate and is subjected to prohibit allosteric inhibition by ornithine or proline [7, 8]. In plants, proline plays an important role as an osmoprotectant and, in mammals, ornithine biosynthesis is crucial for proper ammonia detoxification [9, 10]. Mammalian P5CS has two isoforms which are produced by exon sliding. The short isoform contains 793 amino acid residues and the long isoform, with two additional amino acids in the N-terminal region of γ-GK active site contains 795 residues [1, 9]. In the gut, the short isoform with high activity participates in arginine biosynthesis and is inhibited by ornithine. Furthermore, the expression of long isoform is seen in different tissues, which catabolizes glutamate to proline, without any sensitivity to ornithine [1, 11]. Disfunction of both long and short isoforms of P5CS is caused by a G→A transition at the position 251, leading to generate a new restriction site for the endonuclease *styI* (CCGAGG→CCAAGG) and hence substitutes glutamine for arginine at position 84 (R84Q), this missense mutation reduces the production of ornithine, citrulline, argenine and proline. The metabolic phenotype is represented as hyperammonemia, hypoornithinemia, hypocitrullinemia, hypoargininemia and hypoprolinemia which associated with a range of clinical abnormalities including progressive neurodegeneration, joint laxity, skin hyperelasticity and bilateral subcapsular cataracts [9, 12]. It was shown that R84 in γ–GK domain is conserved in the orthologous enzymes of mouse, *Drosophila*, *Caenorhabditis elegans* and *vigna aconitifolia* [9]. The availability of determined protein structures in Protein Data Bank (PDB) could provide helpful tools for analysis of substitutions at the pro-

tein structure level. In addition, tertiary structure modeling tools using sequence homology can be applied when an experimentally determined structure does not exist [13]. According to PDB, the tertiary structure of human P5CS was determined (2h5j), but it is only homolog to the C-terminal part of the human P5CS, which has a γ -GPR activity. In this study, the secondary and tertiary structure of the N-terminal domain of the wild type P5CS from human were predicted and compared with the mutated (R84Q) one. Several computational tools were applied to analyze how this mutation has affected on the structural and functional features of P5CS enzyme which results in impairing of proline biosynthesis pathway. These *In-silico* techniques are not time-labor- and cost-demanding [14].

Material and Methods

Amino acid sequence of human P5CS

The Amino acid sequence of the long isoform of the human P5CS Enzyme, accession number: P54886 was retrieved from EMBL- EBI database (www.ebi.ac.uk).

Secondary structure prediction of human P5CS

Secondary structure of N-terminal part of human P5CS, with γ-GK activity was predicted, for both wild type and mutated (R84Q) sequence. For this purpose; several secondary structure prediction server such as GORIV [15], Porter [16], nn predict [17] and SABLE [18-20] were applied.

Modeling of tertiary structure for γ-GK domain of human P5CS and Analysis of the R84Qe mutation on this structure

The only determined structure of P5CS existing in PDB was related to C-terminal domain of human enzyme, which starts from residue 361 and has the γ-GPR activity. Furthermore, three different determined structures for γ-GK from bacteria exist in PDB (2j5v, 2j5t, 2ako). Based on these data, the tertiary structure of the N-terminal part of human P5CS with γ-GK activity was predicted using CPH models [21] as a comparative modeling server. The output predicted pdb file for N-terminal domain of human P5CS, was visualized using swiss-pdb viewer (http://www.expasy.org/spdbv/) and Accelry Discovery Studio visulizer softwares (http://www.accelrys.com/services/support.html). In addition, the arginine residue at position 84 was substituted by glutamine on predicted structure and it's probably effects on protein structure and properties were analyzed.

Identification of functionally and structurally important residues

To achieve the conservation of arginine residue at position 84 of P5CS, the enzyme sequence was analyzed, applying ConSeq server [22], which represents the amino acid conservation scores. ClustalW [23] software as a multiple sequence alignment tools was used to compare the human P5CS sequence with the sequences of this enzyme in other organisms.

Prediction for amino acid substitutions which affected on the protein function

The SIFT (sorting intolerant from tolerant) server [24, 25] based on sequence homology and the physical properties of amino acids and PolyPhen (Polymorphism Phenotyping) server [26] based on a straightforward physical and comparative considerations were applied.

Analysis of human P5CS conserved domains and patterns

Conserved domains of the N-terminal part of human P5CS with γ-GK activity were retrieved using CD-Search database [27]. The conserved pattern existing in N-terminal γ-GK domain of human P5CS, PS codes: PS00902 was obtained, using PROSITE database (http://expasy.org/prosite/).

Results

As previously described, there are two isoforms for human P5CS; the long and the short form, with 795 and 793 amino acids, respectively. According to published data [1], the first 361 residues is the γ-GK domain. Secondary structure of this domain for both wild type and mutated (R84Q), were predicted applying several available tools, including GORIV, Porter, SABLE and nn predict. These secondary structure prediction tools have the lease overlap in their designed algorithm; therefore a combination of them would give us the best comparable result. Results from all servers showed that arginine 84 is located in a coil structure. Furthermore, all softwares except GORIV supposed an alpha helix structure as a subsequent fold. The predicted structures of wild type and mutated one were compared, using Porter, SABLE and nn predict. No significant differences were observed. According to GORIV prediction, although the R84Q missense mutation does not change the coiled position of residue 84 also, it makes the next alpha helix longer for one residue. In fact, the last residue in this coil structure inters in the adjacent alpha helix which may be due to the lower confidentiality seen at the secondary structure ends in different prediction methods (Fig. 1).

The crystal structure of the C-terminal part with γ-GPR activity of human P5CS was determined (2H5G). The determined structure for γ-GK enzyme from bacteria also existed in PDB: 2J5V and 2J5T both for *Escherichia coli*, and 2AKO for *Campylobacter jejuni*. According to these homologous, CPH models as a comparative modeling strategy was used to predict the tertiary structure of γ-GK domain. CPH models server definitely used 2J5T, chain D as a template for γ-GK tertiary structure prediction and residue 70 to 353 was modeled with score 174.0 bits and 33.8% identity and RMSD between them was 10.7. The model file was visualized by swiss pdb viewer and Accelry Discovery Studio visualizer. At this model, the arginine residue and mutated glutamine at position 84 were located in the coil region (Fig. 2).

Computing the energy (force field) for wild type and mutated structure of γ-GK domain of P5CS by swiss pdb viewer force field revealed that substitution of R84Q alters the total energy of structure. It was observed that the total and electrostatic energy were -10400.396 KJ/mol and -6925.29 KJ/mol for wild type, while there were-10179.23029 KJ/mol and -6800.0229 KJ/mol for mutated structure. It is clear that the wild type structure had a lower energy, leads to a more stable enzyme structure.

In addition, based on computed accessible surface for residue 84 by this software, R has 29.4% accessible surface in

wild type but when Q substitutes for R in mutated molecule, the accessibility is decreased to 20.5% level. This alteration in accessible surface also influence in adjacent residues, such as glycine 85 which its accessible surface reduces from 40% in wild type structure to 38% in mutant protein. Actually, the measured area and volume are 11473 squared angstroms and 35220 cubic angstroms for wild type modeled structure, while there are 11446 squared angstroms and 35174 cubic angstroms for mutated one. It means R84Q mutation reduces the molecular surface of the protein, which leads to less accessibility of the major functional or structural sites of enzyme.

Furthermore, the conservation status of arginine residue at position 84 in human P5CS was analyzed. As expected, the results from Clustal W indicated that position 84 is highly conserved among 50 various organisms, particularly mammals and is occupied often by amino acid R or sometimes G (Data not shown). ConSeq server, which identifies the important amino acids in protein function and structure, revealed that the conservation score for R84 is 7 (score 9 and 1 are conserved and variable amino acid, respectively) and also this residue has an exposed position in protein structure (Table 1).

Therefore, any substitution for this highly conserved amino acid could affect its properties and impair the function of P5CS enzyme. In addition, according to this data, several amino acids which are located in neighboring of R84 such as A80, V81, A91 and R94 are structurally important residues and also, the results of SABLE server showed that the buried or exposed position of residue at position A80, V81, A91 and R94 as well as 84 alter when glutamine replaces for arginine. Therefore, the missense mutation at position 84 may affect on these key amino acids and cause structural and functional changes in mutated form of the human P5CS enzyme (Table 2).

These findings were in accompany with the results of SIFT server which, showed that substitution at position 84 from R to Q could affect the protein function. On the other hand, other amino acids at position 84 instead of R could not be tolerable. The experimentally accuracy of SIFT server prediction was reported for nsSNPs which associated with DNA repair genes [13, 28-31]. Furthermore, the results from PolyPhen server which predicts possible impact of an amino acid substitution on the structure and function of a human protein using straight forward physical and comparative considerations, confirmed that this substitution is potentially damaging protein function. Therefore, replacement of highly conserved arginine residue by glutamine could alter the enzyme conformation and structure, influence the function and reduce the enzyme activity, which results in impaired synthesis of proline, citrulline, ornithine and arginine amino acids.

Analyses within CD-Search database, showed that the N-terminal part of human P5CS has three Conserved domains: putative nucleotide binding site based on similarity to *Campylobacter jejuni* glutamate 5-kinase which is included A80, S265, D266, V267, L270, F271, M305, G307 and K330, putativephosphatebinding site based on similarity to region identified in tomato glutamate 5-kinase located in region 118-125 with GAVAFGKQ sequence and finally putative allosteric binding site based on similarity to mutational studies in tomato glutamate 5-kinase which are included A120, D223 and D248.It is clear that R84 does not involve in these sites but as previously mentioned, P5CS is an ATP-dependent enzyme, any changes which affect on amino acids involved in putative nucleotide (ATP) binding site such as the residue at position 80 could alter its ATP binding capacity and impair its function (before the effects of this mutation on amino acid 84 were described).

Table 1. Amino acid conservation score for γ-GK dmain of human P5CS retrieved from ConSeq server. The position 84 is indicated by arrow.

POS	SEQ	SCORE	COLOR	B/E	FUNCTION	MSA DATA	RESIDUE VARIETY
76	K	-0.918	8	e	f	10/50	K
77	L	-0.459	7	b		10/50	L,V
78	G	-0.844	8	b		10/50	G
79	S	-0.664	7	b		10/50	S,T
80	A	-0.990	9	b	s	10/50	A
81	V	-0.984	9	b	s	10/50	V
82	V	-0.597	7	b		10/50	I,V
83	T	-0.680	7	b		10/50	S,T
84	R	-0.539	7	e		10/50	G,R
85	G	1.894	1	e		10/50	A,E,G,K,Q
86	D	-0.082	5	e		10/50	D,E,G
87	E	-0.336	6	e		10/50	E,G
88	C	-0.421	7	e		10/50	C,R
89	G	-0.584	7	b		4/50	G
90	L	-0.862	8	b		11/50	L
91	A	-1.004	9	b	s	11/50	A

- POS: The position of the AA.
- SEQ: AA code.
- COLOR: The color scale representing the conservation score (9- conserved, 1- variable)
- B/E: Buried (b) or Exposed (e) residue.
- FUNCTION: functional (f) or structural (s) residue (f- highly conserved and exposed, S- highly conserved and buried).
- MSA DATA: The number of aligned sequences having an amino acid (non-gaped) from the overall number of sequences at each.position.
- RESIDUE VARIETY: The amino acids variety in multiple sequence alignment at each position.

Table 2. Relative solvent accessibility prediction by SABLE secondary structure prediction server.
R and Q at position 84 and the amino acids at position 79- 96 are indicated (0- fully buried, 100- fully exposed).

Amino acid	S	A	V	V	T	R/Q	G	D	E	C	G	L	A	L	G	R	L	A
Wild type sequence	11	9	9	5	14	37	49	53	48	41	41	12	24	25	53	16	4	31
Mutatedsequence	10	8	8	6	14	33	47	56	45	46	38	11	29	24	55	18	3	31

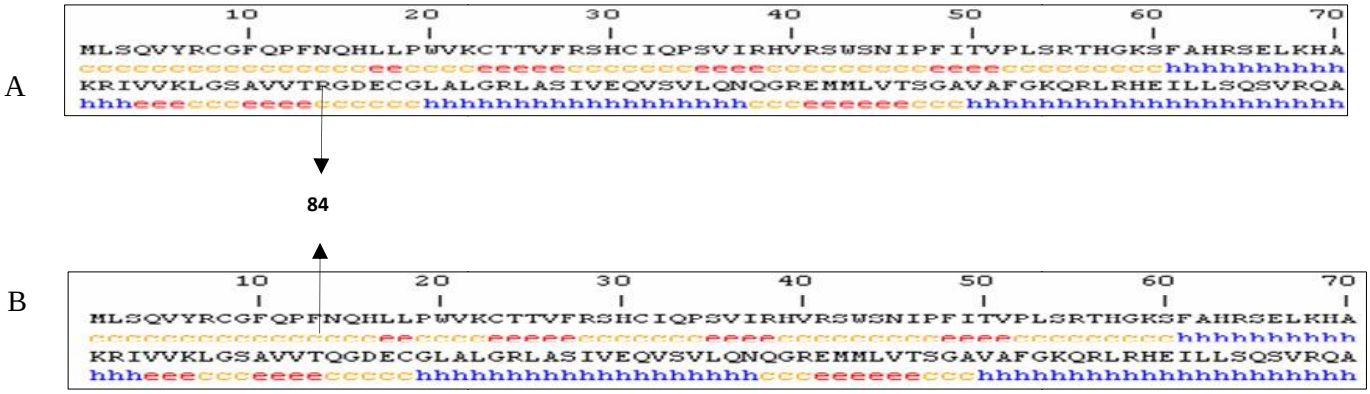

Figure 1. GOR IV secondary structure prediction results for A. wild type and B. mutated γ-GK domain of human P5CS (results for amino acids at position1 to 140 are shown).

Using PROSITE database showed that the conserved pattern is located far from the residue 84 and contains the residues at position 301 to 318 (Data not shown). In addition to the overall similarity of the N-terminal half ofhumanP5CS to γ-GKs from a variety of organisms (*e.g.* 36% identity with *S. cerevisiae* γ-GK), BLASTP and Blocks database [32] searches revealed a 32-amino acid motif (from N245 to P276) in the N-terminal half of human P5CS with similarity to a conserved sequence for members of the aspartokinase family [33-35]. This sequence similarity could be observed between *E. coli* and *Serratia marcescens* aspartokinases and γ-GKs, also. So it was suggested that it is the active site of these related enzymes [34]. Three residues in this motif corresponding to N245, N260, and P276 in long isoform of human P5CS are conserved in all γ-GKs and nearly all members of the aspartokinase family. When residues involved in enzyme active site was highlighted in the 3-D structure, it was found that although R84 is not involved in the active site, but it is approximately adjacent to this region in 3-D conformation (Fig. 2).

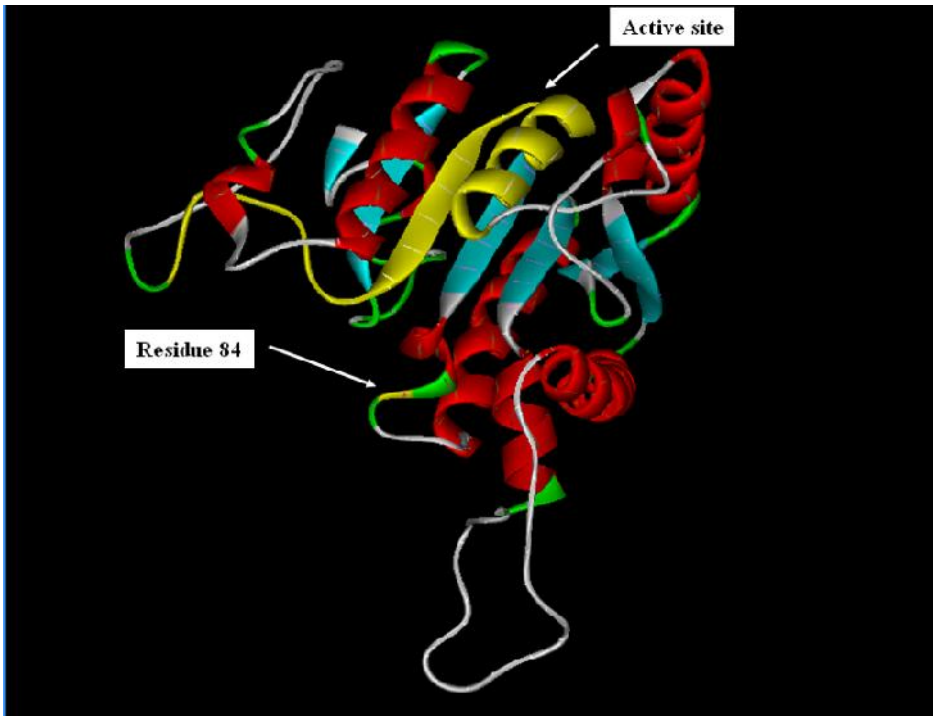

Figure 2. Schematic representation of proposed tertiary structure for γ-GK domain of human P5CS.The amino acid at the position 84 in a coil structure and the amino acids 245 to 276 which involved in active site are indicted by arrow.

Discussion

R84Q missense mutation in Mammalian P5CS enzyme, which catalyzes a critical step in the proline, ornithine, citrulline and arginine biosynthesis, alters the conserved residue in γ-GK domain and is responsible for Hyperammonemia disorder in human. According to our results, substitution of R84Q alters the total energy of structure; in which the higher energy level of mutant type structure could lead to a less stable enzyme structure. In addition, the reduced molecular surface of the protein in mutant structure, results in less accessibility of the major functional or structural sites of enzyme. On the other hand, any replacement of arginine 84, which is a highly conserved residue and also important in protein function and structure, could not be tolerable and may affect its properties and impair the function and activity of P5CS enzyme; since it alters the buried or exposed position of R84 as well as some other functionally important neighboring residues.

Arginine is a charged polar amino acid which has the strongest basic power among other amino acids within its group and its positive charge is stabilized by resonance. At the neutral pH (pH: 7) the pK of Arginine side chain is about 12.5, furthermore, glutamine has an uncharged polar side chain with pK~4.2 at the same pH. Amino acids with uncharged polar side chains are relatively hydrophilic and are usually exposed on the outside of proteins. So replacement of glutamine with arginine residue in a protein makes an uncharged medium size amino acid substitution (R) for a charged large amino acid (Q).

This amino acid substitution could neutralize or reverse the charge on the protein surface and may impair the electrical environment with the adjacent residues and thereby may affect on the structure and function of the enzyme. Of course, this charge alteration on protein surface could interfere with other functional sites, such as allosteric, nucleotide and phosphate binding sites. It is also possible that the active site properties alter, due to the reduced electrostatic energy and charged density.

References

1. Hu, C.A. Lin, W.W. Obie, C., Valle, D., Molecular enzymology of mammalian Delta1-pyrroline-5-carboxylate synthase. Alternative splice donor utilization generates isoforms with different sensitivity to ornithine inhibition.*J Biol Chem*, 1999, vol. 274, pp. 6754-62.

2. Yoshiba, Y. Kiyosue, T. Katagiri, T. Ueda, H. Mizoguchi, T. Yamaguchi-Shinozaki, K. Wada, K. Harada, Y., Shinozaki, K., Correlation between the induction of a gene for delta 1-pyrroline-5-carboxylate synthetase and the accumulation of proline in Arabidopsis thaliana under osmotic stress.*Plant J*, 1995, vol. 7, pp. 751-60.

3. Sekine, T. Kawaguchi, A. Hamano, Y., Takagi, H., Desensitization of feedback inhibition of the Saccharomyces cerevisiae gamma-glutamyl kinase enhances proline accumulation and freezing tolerance.*Appl. Environ. Microbiol.*, 2007, vol. 73, pp. 4011-9.

4. Phang, J.M., Yeh, G.C, and Scriver, C.R, Disorders of proline and hydroxyproline metabolism.in Scriver, C.R., Beaudet, A.L., sly, W.S. and Valle, D. (eds),The metabolic and molecular bases of inherited disease. McGraw Hill, New York, NY. pp 1125-1146. 1995.

5. Wang, S.S. Brandriss, M.C., Proline utilization in Saccharomyces cerevisiae: analysis of the cloned PUT1 gene.*Mol. Cell. Biol.*, 1986, vol. 6, pp. 2638-45.

6. Hu, C.A. Delauney, A.J., Verma, D.P., A bifunctional enzyme (delta 1-pyrroline-5-carboxylate synthetase) catalyzes the first two steps in proline biosynthesis in plants.*Proc. Natl. Acad. Sci. U. S. A.*, 1992, vol. 89, pp. 9354-8.

7. Fujita, T. Maggio, A. Garcia-Rios, M. Stauffacher, C. Bressan, R.A., Csonka, L.N., Identification of regions of the tomato gamma-glutamyl kinase that are involved in allosteric regulation by proline.*J Biol Chem*, 2003, vol. 278, pp. 14203-10.

8. Hong, Z. Lakkineni, K. Zhang, Z., Verma, D.P., Removal of feedback inhibition of delta(1)-pyrroline-5-carboxylate synthetase results in increased proline accumulation and protection of plants from osmotic stress.*Plant Physiol*, 2000, vol. 122, pp. 1129-36.

9. Baumgartner, M.R. Hu, C.A. Almashanu, S. Steel, G. Obie, C. Aral, B. Rabier, D. Kamoun, P. Saudubray, J.M., Valle, D., Hyperammonemia with reduced ornithine, citrulline, arginine and proline: a new inborn error caused by a mutation in the gene encoding delta(1)-pyrroline-5-carboxylate synthase.*Hum Mol Genet*, 2000, vol. 9, pp. 2853-8.

10. Kamoun, P. Aral, B., Saudubray, J.M., [A new inherited metabolic disease: delta1-pyrroline 5-carboxylate synthetase deficiency].*Bull. Acad. Natl. Med.*, 1998, vol. 182, pp. 131-7; discussion 138-9.

11. Lodato, R.F. Smith, R.J. Valle, D. Phang, J.M., Aoki, T.T., Regulation of proline biosynthesis: the inhibition of pyrroline-5-carboxylate synthase activity by ornithine.*Metabolism*, 1981, vol. 30, pp. 908-13.

12. Hu, C.A. Khalil, S. Zhaorigetu, S. Liu, Z. Tyler, M. Wan, G., Valle, D., Human Delta(1)-pyrroline-5-carboxylate synthase: function and regulation.*Amino Acids*, 2008.

13. Nakken, S. Alseth, I., Rognes, T., Computational prediction of the effects of non-synonymous single nucleotide polymorphisms in human DNA repair genes.*Neuroscience*, 2007, vol. 145, pp. 1273-9.

14. Wainreb, G. Ashkenazy, H. Bromberg, Y. Starovolsky-Shitrit, A. Haliloglu, T. Ruppin, E. Avraham, K.B. Rost, B., Ben-Tal, N., MuD: an interactive web server for the prediction of non-neutral substitutions using protein structural data.*Nucleic acids research*, vol. 38 Suppl, pp. W523-8.

15. Garnier, J. Gibrat, J.F., Robson, B., GOR method for predicting protein secondary structure from amino acid sequence.*Methods Enzymol.*, 1996, vol. 266, pp. 540-53.

16. Pollastri, G. McLysaght, A., Porter: a new, accurate server for protein secondary structure prediction.*Bioinformatics*, 2005, vol. 21, pp. 1719-20.

17. Kneller, D.G. Cohen, F.E., Langridge, R., Improvements in protein secondary structure prediction by an enhanced neural network.*J. Mol. Biol.*, 1990, vol. 214, pp. 171-82.

18. Ludwikowski, G. Szymanski, W. Szymanski, M. Adamczak, R. Kazdepka-Zieminska, A., Pasinska, M., [Influence of cigarette smoking on some sperm parameters in males with decreased fertility].*Przegl Lek*, 2004, vol. 61, pp. 1031-2.

19. Adamczak, R. Porollo, A., Meller, J., Combining prediction of secondary structure and solvent accessibility in proteins.*Proteins*, 2005, vol. 59, pp. 467-75.

20. Wagner, M. Adamczak, R. Porollo, A., Meller, J., Linear regression models for solvent accessibility prediction in proteins.*J. Comput. Biol.*, 2005, vol. 12, pp. 355-69.

21. Nielsen, M. Lundegaard, C. Lund, O., Petersen, T.N., CPHmodels-3.0--remote homology modeling using structure-guided sequence profiles.*Nucleic Acids Res*, 2010, vol. 38, pp. W576-81

22. Berezin, C. Glaser, F. Rosenberg, J. Paz, I. Pupko, T. Fariselli, P. Casadio, R., Ben-Tal, N., ConSeq: the identification

of functionally and structurally important residues in protein sequences.*Bioinformatics*, 2004, vol. 20, pp. 1322-4.
23. Chenna, R. Sugawara, H. Koike, T. Lopez, R. Gibson, T.J. Higgins, D.G., Thompson, J.D., Multiple sequence alignment with the Clustal series of programs.*Nucleic Acids Res*, 2003, vol. 31, pp. 3497-500.
24. Ng, P.C. Henikoff, S., Predicting deleterious amino acid substitutions.*Genome Res.*, 2001, vol. 11, pp. 863-74.
25. Ng, P.C. Henikoff, S., Accounting for human polymorphisms predicted to affect protein function.*Genome Res.*, 2002, vol. 12, pp. 436-46.
26. Ramensky, V. Bork, P., Sunyaev, S., Human non-synonymous SNPs: server and survey.*Nucleic Acids Res*, 2002, vol. 30, pp. 3894-900.
27. Marchler-Bauer, A. Bryant, S.H., CD-Search: protein domain annotations on the fly.*Nucleic Acids Res*, 2004, vol. 32, pp. W327-31.
28. Savas, S. Kim, D.Y. Ahmad, M.F. Shariff, M., Ozcelik, H., Identifying functional genetic variants in DNA repair pathway using protein conservation analysis.*Cancer Epidemiol. Biomarkers Prev.*, 2004, vol. 13, pp. 801-7.
29. Zhu, Y. Spitz, M.R. Amos, C.I. Lin, J. Schabath, M.B., Wu, X., An evolutionary perspective on single-nucleotide polymorphism screening in molecular cancer epidemiology.*Cancer Res.*, 2004, vol. 64, pp. 2251-7.
30. Xi, T. Jones, I.M., Mohrenweiser, H.W., Many amino acid substitution variants identified in DNA repair genes during human population screenings are predicted to impact protein function.*Genomics*, 2004, vol. 83, pp. 970-9.
31. Rudd, M.F. Williams, R.D. Webb, E.L. Schmidt, S. Sellick, G.S., Houlston, R.S., The predicted impact of coding single nucleotide polymorphisms database.*Cancer Epidemiol. Biomarkers Prev.*, 2005, vol. 14, pp. 2598-604.
32. Pietrokovski, S. Henikoff, J.G., Henikoff, S., The Blocks database--a system for protein classification.*Nucleic Acids Res*, 1996, vol. 24, pp. 197-200.
33. Omori, K. Imai, Y. Suzuki, S., Komatsubara, S., Nucleotide sequence of the Serratia marcescens threonine operon and analysis of the threonine operon mutations which alter feedback inhibition of both aspartokinase I and homoserine dehydrogenase I.*J. Bacteriol.*, 1993, vol. 175, pp. 785-94.
34. Omori, K. Suzuki, S. Imai, Y., Komatsubara, S., Analysis of the Serratia marcescens proBA operon and feedback control of proline biosynthesis.*J Gen Microbiol*, 1991, vol. 137, pp. 509-17.
35. Serina, L. Blondin, C. Krin, E. Sismeiro, O. Danchin, A. Sakamoto, H. Gilles, A.M., Barzu, O., Escherichia coli UMP-kinase, a member of the aspartokinase family, is a hexamer regulated by guanine nucleotides and UTP.*Biochemistry (Mosc.)*, 1995, vol. 34, pp. 5066-74.