Bioinformatical analysis of influence of human tyrosyl-tRNA synthetase mutations associated with neuropathy Charcot-Marie-Tooth type C, on its local spatial structure properties

K. A. Odynets, A. I. Kornelyuk

Institute of molecular biology and genetics NAS of Ukraine
Academicain Zabolotnog str., 150, Kyiv, 03680 Ukraine

odynets@imbg.org.ua, kornelyuk@imbg.org.ua

A causal relationship between three independent mutant forms of the human cytoplasmic tyrosyl-tRNA synthetase (TyrRS) and dominant Charcot-Marie-Tooth neuropathy type C (DI-CMTC) has been recently reported. In this article we carried out the bioinformatical investigation of the effect of amino acid residues change in the cases of two point mutations Gly41Arg, Glu196Lys and deletion 153-156delVKQV on structural properties of catalytic module dimer of TyrRS (2x39 kDa or, so called, mini-TyrRS), especially on its local environment and electrostatic potential of molecular surface of the protein.

Key words: tyrosyl-tRNA synthetase, neuropathy Charcot-Marie-Tooth type C, spatial structure, computer mutagenesis.

Introduction. In the recent years the application of novel methods of investigation in molecular genetics of hereditary diseases has unexpectedly revealed connections between some neurodegenerative diseases and mutations in housekeeping genes, which encode cytoplasmic aminoacyl-tRNA synthetases (aaRS). This enzymatic family catalyses highly-specific binding of L-amino acids to corresponding tRNA. aaRS are the key components of cell system of protein biosynthesis as they implement the rules of genetic code correspondence between amino acids and tRNA anticodons. Beside this canonical function, aaRS are specific for a number of various other functions (reviews [1, 2]).

Recently we have shown that around 10 independent mutations in glycyl-tRNA synthetase gene (GlyRS) are associated with sensomotor axonal neuropathy of Charcot-Marie-Tooth (CMT) type 2D [3, 4]. One more synthetase, cytoplasmic tyrosyl-tRNA (TyrRS), has been discovered to have the connection with neurodegeneration. Three independent mutations in the TyrRS gene result in the development of CMT type C [5, 6]. Mammalian TyrRS have been the subject of our investigation [7, 8], thus, current publication is dedicated to the analysis of these mutations. The connection of alanine-specific synthetase and the death of Purkinje cells and ataxia [9, 10] has been revealed for aaRS of different specificity, while tryptophan-specific aaRS were revealed to be connected with Alzheimer's

disease [11, 12]. The aforementioned facts show that some tRNA aminoacylation enzymes, in addition to their canonical function, also play a significant role in the functioning of structure of peripheral neurons.

CMT neuropathy. Charcot-Marie-Tooth is one of the most widespread groups of hereditary peripheral neuropathies (disease frequency 1:2500, or 2–4 cases per 10000 people). CMT neuropathies is a clinically and genetically heterogenic group of hereditary monogenic neuron-muscular diseases of peripheral nervous system, which are specific for progressing weakening and atrophy of peroneal muscles and distal arm muscles. CMT can be classified according to electrophysical and neuropathological criteria, conditioned by the demyelination or degeneration of axons, or by their combination (reviews [13–16]). Both basic CMT forms – demyelinating and axonal – may be inherited by autosomal-dominant, autosomal-recessive, and X-attached inheritance modes. There are *de novo* mutations as well.

The most frequent dominant mutation during CMT (68–90% CMT1) is the duplication of $1.4 \cdot 10^3$ bp in 17p11.2-12 locus of 17 chromosome, occurring during unequal crossing over of homologous chromosomes in meiosis and resulting in the development of demyelination CMT type 1A (CMT1A) [17]. Clinical presentations in the majority of cases were revealed to be determined by overexpression of *PMP22* gene of myelin peripheral protein and to occur at the presence of three or four copies of this gene. Another 10–20% of cases are due to X-attached mutations in connexin 32 gene (Cx32, Gjb1) in Xq13.1 locus [18]. Recessive mutations, corresponding to the development of the disease, were also revealed in protein-0 of myelin (P0, MPZ, locus 1q22) [19], periaxin (locus 19q13) [20], myotubularin-related protein-2 (locus 11q22) [21]. Therefore, the investigation of primary genetic defects of demyelinating form of CMT neuropathy was primarily focused on the mentioned genes.

Over 10 different mutant alleles, which cause dominant axonal form of CMT type 2D, were revealed in glycyl-tRNA synthetase (gene *GARS*). Functional analysis of mutant GlyRS forms did not reveal any significant changes in the levels of protein expression. It was found that in the majority of cases mutant GlyRS forms are specific for changed localization in neurons.

Four of five investigated mutations were shown to be specific for the absence of GlyRS function using at least one method, confirming at the same time that the deficiency in tRNA aminoacylation plays the important role in pathogenesis of this disease [3]. Endogenic GlyRS-associated granules were also detected in the cultivated neurons and axons of peripheral nerves of normal human tissue.

Mutation mapping for recently discovered crystallographic structure of human GlyRS revealed their band-like localization, surrounding dimerization interface on both sides [6]. CMT phenotype has no correlation with aminoacylation activity, however the majority of mutations prevent the formation of dimer by increasing or decreasing dimerization. Seven CMT-causing variants and wild type protein were expressed in transfected neuroblastoma cells, which develop into primitive neurons. Wild type GlyRS is distributed in newly-formed neurons and is associated with their normal development, and vice versa all GlyRS mutant forms were shown to be defected by distribution specificities.

Thus, GlyRS localization in cases of all CMT-causing mutations is disordered, which may be related to the change in the surface of its dimerization interface.

DI-CMTC neuropathy and TyrRS. Intermediate autosomal dominant form of neuropathy CMT type C (DI-CMTC, MIM 608323) is a genetic and phenotypical variant of CMT, specific for intermediate nerve conduction velocity and histological features of both degeneration and demyelination of axons. DI-CMT is peripheral neuropathy, specific for the features of both CMT forms (demyelinating and axonal) as well as for intermediate nerve conduction velocities (NCV) in the range of 25–45 m/sec. Identified locus 1p34-p35 of chromosome 1 is connected with DI-CMTC. This locus corresponds to *YARS* gene (GeneID:8565), which encodes human cytoplasmic TyrRS [5, 6].

Being typical housekeeping gene, *YARS* is expressed in any tissue, including brain and spinal marrow [5]. TyrRS is localized in axonal tips in the cultures of differential primary motor neurons and neuroblastomas. Such specific distribution decreases significantly in the cells, which express the mentioned mutant TyrRS forms [5]. The application of biochemi-

cal experiments and genetic complementation in yeasts revealed partial loss of aminoacylation activity of these mutant proteins and the decrease in yeast growth speed under the influence of mutations in orthological yeast gene *TYS1*.

Earlier *YARS* cDNA was cloned from several different human cDNA-libraries [22]. This cDNA encodes TyrRS subunit, consisting of 528 amino acid residues (aa) (GenPept code NP 003671.1). TyrRS (EC 6.1.1.1) is $_2$ dimer, molecular weight 2 59 kDa, capable of catalyzing L-tyrosine activation and its specific attachment to tRNA^{Tyr}. Analysis of protein amino acids revealed NH$_2$-terminal module (residues Met1-Pro342) to contain catalytic and anticodon binding domains (Fig.1, *a–c*). N-module dimer (2 39 kDa), which preserves catalytic features of full length TyrRS, is also known as *mini-TyrRS* [23]. COOH-terminal part of TyrRS or its C-module belong to the family of Endothelial Monocyte-Activating Polypeptide II (EMAP II)-like domains and can be found in TyrRS of higher eukaryotes only [22]. Human TyrRS were expressed *in vitro* and their biochemical properties have been studied [24].

Although native human TyrRS is an inactive molecule towards the transfer of intercellular signal, it may be split into two fragments both with cytokine-like activities [23]. TyrRS is secreted at the apoptotic conditions in cell culture, where it is split into N- and C-terminal fragments, corresponding to N- and C-modules. N-module is an interleukin-8 (IL8)-like cytokine, whereas released C-terminal fragment is EMAP II like cytokine. It has been shown that cytokine activities of split human TyrRS are determined by highly-differentiated motifs, specific for mammalian fragment [25, 26]. Proangiogenic and antiangiogenic activity of both TyrRS fragments, which stimulate immune cells and thus, link translation with important cell signaling pathway, have been revealed [2, 27, 28].

Defects in *YARS* occur due to the dominant intermediate form of CMT type C (DI-CMT C). Currently, there are two heterozygous missense mutations and one *de novo* deletion in TyrRS in different families of patients with DI-CMTC [5]. Thus, heterogeneous transition 121G>A in exon 2 *YARS*, resulting in missense mutation of Gly41Arg (mutation CMT-160), has been identified in one family from North America. In the

Bulgarian family, heterozygous transition 586G>A in exon 5 of the mentioned gene has been revealed to result in mutations of Glu196Lys (mutation CMT-176). Besides, deletion of 458–469 12, 12 bp long, in translation frame of exon 4 was identified in a Belgian patient. This deletion results in another deletion of 153–156delVKQV (deletion PN-765). Analysis of this mutation and genetic typing of asymptomatic parents showed that this mutation occurred *de novo*.

Spatial structure of N-module of human TyrRS (region of Met1–Pro342) was determined using the method of X-ray diffraction (resolution 0.12 nm and 0.16 nm) in 2003 (Protein Data Bank (PDB) codes 1N3L and 1Q11) [29, 30]. Spatial structures of catalytic modules of two TyrRS, structural homologues of mammalian N-module, in complexes with corresponding tRNA^{Tyr}, have been determined. These are TyrRS of *Methanocaldococcus jannaschii* archaebacteria with non-modified transcript of its tRNA^{Tyr} (PDB code 1J1U, residues Met1–Leu306) [31] and *Saccharomyces cerevisiae* with native tRNA^{Tyr} (PDB code 2DLC, Asp8–Pro356) [32]. All these structural data allow performing bioinformatical analysis of the influence of changing aforementioned amino acid residues on structural specificities of human TyrRS in norm and in three mutant forms.

Materials and Methods. Information from Online Mendelian Inheritance in Man (OMIM) database (http://www.ncbi.nlm.nih.gov/omim) was used. Coordinate files of spatial structures of proteins were obtained from PDB [33], additional information was obtained from PDBsum [34]. Spatial structure was visualized and analysis of local surrounding of residues was performed using Swiss-PdbViewer (DeepView) 3.7 [35]. Web-server ModLoop [36] was used for filling in for absent catalytic KMSKS loop 7 bp long. To construct the models of mutant forms of TyrRS N-module monomer using the method of computer mutagenesis web-server Swiss-Model 3.5 [37] in First Approach Mode and program for homological modeling Modeller 9.2 [38] were used independently [38]. The quality of each of 10 obtained models for each of three mutations was evaluated by the value of Modeller Objective Function and with Biotech Validation Suite web-server (http://biotech.ebi.ac.uk:/8400/), and then according to the general rating, the best models were selected.

Structures of N-module dimers were formed using superpositions of both monomers with Biological Unit dimer from Protein Quaternary Structure (PQS) web-server.

To analyze possible conformation of substituted side radicals we used Swiss-PdbViewer 3.7 software with its rotamer library and Mutate tool in order to select rotamers with the lowest rating. Optimal conformation for a new side chain was chosen as the best rotamer, criterion selected – the lowest conflicts of Van der Waals radii of its atoms and the atoms of neighboring residues.

N-module spatial structure was analyzed using its coordinates from 1N3L file in the form of symmetrical dimer with incorporated "absent" interdomain catalytic KMSKS loop of 7 bp (222–228-KMSSSEE) using ModLoop web-server. Mutant forms of protein were independently modeled, using Swiss-PdbViewer 3.7, Modeller 9.2, and Swiss-Model 3.5 web-server (after substitution of corresponding residues, completion of loops and addition of absent atoms, the structures of models were optimized according to ProModII algorithm in GROMOS96 force field).

Corresponding regions of amino acid residues of homologous TyrRS of eukaryotes and archaebacteria were investigated using ready-made alignments from tyrosyl-tRNA synthetase catalytic core domain (TyrRS_core, codes cd00805 and CDD:28813), corresponding to the region of 38–320 of human TyrRS. The degree of exposure of amino acid residues was analyzed using GetArea web-server 1.1 [39] and the degree of evolutionary conservatism was analyzed using ConSurf 3.0 web-server [40].

Results and Discussion. Bioinformatical assay of local features of TyrRS spatial structure of normal and mutant forms has been carried out for two point mutations – Gly41Arg and Glu196Lys – and one deletion 153–156VKQV of human TyrRS, which are associated with CMT neuropathy type C. N-module dimer in normal condition and its 3 mutant forms were analyzed on the following issues: location of these residues in regards to the elements of secondary structure of protein, their evolutionary conservatism in amino acid sequences of homologous TyrRS, closest environment features, capability to form hydrogen bonds with neighboring residues, exposure degree on the surface of

macromolecule and motility degree of neighboring regions of polypeptide backbone by B-factors of C -atoms of amino acid residues in crystal structure. The changes in isoelectric point were predicted and the influence on the distribution of electric potential on the surface of macromolecule and possible effect on binding of three TyrRS substrates, including tRNA^{Tyr}, were analyzed.

General characteristics of mutation-related sites in norm. All investigated sites are located in TyrRS N-module to mobile KMSKS loop, within the catalytic domain with Rossmann fold topology. Location of these residues against TyrRS amino acid subunit and the elements of its secondary structure in norm is presented in figure 1, *a–c*, and their location against polypeptide backbone of N-module dimer in norm is presented in figure 1, *d*. All three sites are localized on one side of N-module dimer, along with TyrRS catalytic center, and occupy a significant part of the surface, exposed to the solvent, the view from the top on mini-TyrRS dimer is presented in figure 1, *e*. Multiple alignment of 44 TyrRS sequences of higher eukaryotes, from human down to nematode, demonstrates evolutionary invariance of all six residues in mutation sites (data not presented). The calculations of theoretical isoelectric point (pI) values of TyrRS N-module demonstrate expected and rather significant changes as follows: 6.49 in norm; 6.74 – Gly41Arg; 6.29 – delVKQV, and 7.10 for Glu196Lys. Detailed analysis of the structural possibilities of all three mutations is presented below.

Gly41Arg mutation. Regarding the elements of secondary structure, Gly41 is localized on the end of -bundle A1 and is known to belong to the group of the most evolutionary conservative TyrRS residues (Fig.1, *c*, Fig.2). Noteworthy is the significant structural difference of amino acid residue in norm and in mutant protein – arginine retains higher volume than glycine, moreover, guanidine group of the former is positively charged. Thus, rather unexpected was the fact that Gly41Arg mutation is compatible with stable structure of TyrRS. Within 0.5 nm radius from Gly41, there are seven residues (39-YW, 42-TA, L72, and 182-QF), with which hydrogen bonds are not formed, yet Leu72 forms hydrogen bonds with frame of neighboring residues in layer (Fig.3, *a*). Within 0.7 nm radius from
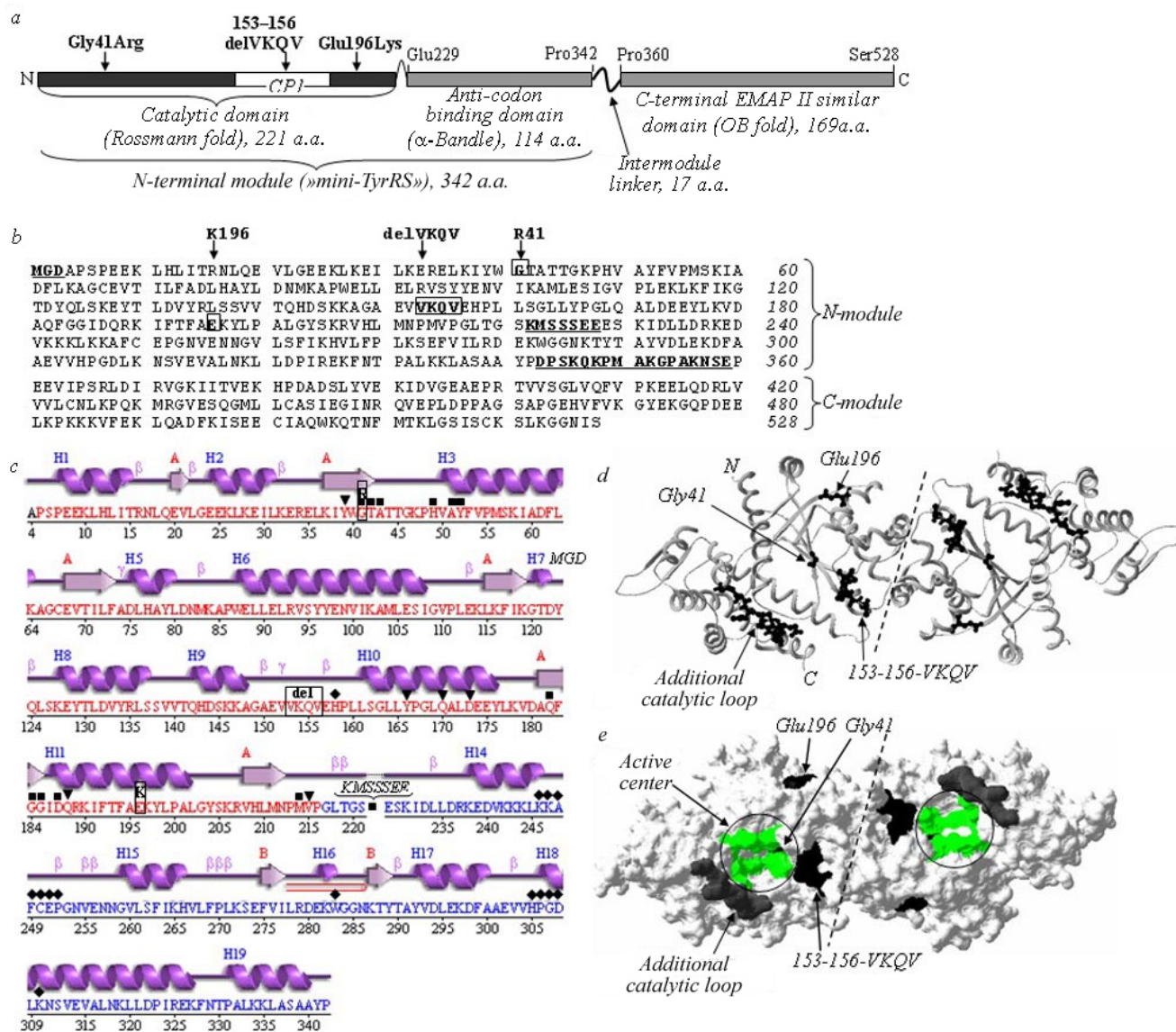
Fig.1 Location of two point heterozygous mutations Gly41Arg and Glu196Lys and deletion 153-156delVKQV of human cytoplasmic TyrRS, associated with Charcot-Marie-Tooth neuropathy type C: *a* – scheme of domain organization of TyrRS subunit (N-module (mini-TyrRS) is specific for aminoacylation and interleukin-8-like activities, while C-module is specific for EMAP II cytokine activity); catalytic and anticodon binding domains (Met1–Ser221 and Glu229–Pro342) are divided by catalytic KMSKS loop); *b* – amino acid sequence of TyrRS subunit in norm (GenPept code NP_003671; DI-CMTC-related residues are marked in bold, the regions with unidentified atomic coordinates in crystallographic structure of N- and C-modules (PDB codes 1N3L, 1Q11, and 1NTG) are italicized); *c* – secondary structure of N-module according to the data of X-ray structure analysis (PDB code 1N3L; corresponding residues are marked by frames; amino acid residues, forming hydrogen bonds with L-tyrosine and AMP-adenine, are marked by triangles; other residues of active center are marked by rectangles, and 14 amino acid residues, capable of interacting with tRNA^Tyr are marked by diamonds); *d* – location of residues in norm regarding the spatial structure of dimer of its N-module (PDB code 1N3L; polypeptide protein framework is presented as band-like model; side chains for residues in norm are colored in black); *e* – molecular surface of N-module dimer in norm, which has the markings of all DCI-CMTC residues (black), completed catalytic loop (grey), and 16 key residues of active center, marked by ovals

Type of organism	Code	Gly41Arg	153-156delVKQV	Glu196Lys
Homo sapiens	4507947	LKI-YW**G**TATTGK	KAGAEV**VKQV**EHPLLS	KIFTFA**E**KYLPAL
Other eukaryotes				
Canis familiaris	73949939	LKV-YW**G**TATTGK	KAGAEV**VKQV**EHPLLS	KIFTFA**E**KYLPAL
Gallus gallus	57530465	VKI-YW**G**TATTGK	KAGAEV**VKQV**EHPLLS	KIFTFA**E**KYLPSL
Danio rerio	41151996	LKV-YW**G**TATTGK	KAGAEV**VKQV**EHPLLS	KIFTLA**E**KYLPSL
Anopheles gambiae	118781822	LKI-YW**G**TATTGK	KAGAEV**VKQV**EHPLMS	KIFTFA**E**KYLPQL
Apis mellifera	110762892	LKI-YW**G**TATTGK	KAGAEV**VKQV**VNPLLS	KIFTFS**E**KYLPLL
Caenorhabditis elegans	71992775	PHV-YW**G**TATTGK	KAGAEV**VKQV**ESPLLS	KIFILA**E**EQLPKL
Yarrowia lipolytica	50548779	LKI-YW**G**TAPTGK	RAGADV**VKQV**DNPLLS	KIFVLA**E**ENLTSI
Ashbya gossypii	45184676	LRL-YW**G**TAPTGK	RAGADV**VKQV**SNPLLS	KIFVLA**E**ENMGSL
Kluyveromyces lactis	50311493	LKL-YW**G**TAPTGR	RAGADV**VKQV**ANPLLS	KIFVLA**E**ENLEKL
Debaryomyces hansenii	50427467	LKI-YW**G**TAPTGK	RAGADV**VKQV**ANPLLS	KIFVLA**E**ENLQSI
Candida albicans	68470880	VKI-YW**G**TAPTGK	RAGADV**VKQV**ANPLLS	KIFVLA**E**ENLPSI
Neurospora crassa	85090711	PRI-YW**G**TATTGR	KAGAEV**VKQS**DNAPLS	KLFAAA**K**DWLPRL
Aspergillus nidulans	40746079	PRI-YW**G**TATTGR	KAGAEI**VKQS**ANAPLS	KLFAAA**V**EWLPKI
Plasmodium falciparum	23612935	LIC-YD**G**FEPSGR	CLKIM**GRSE**GEENYCS	KVNMLA**R**EYCDIK
Cryptosporidium parvum	32398875	HPICYD**G**FEPSGR	CCQIM**GRQEN**DEQPCA	KVNMLA**R**EYCDAA
Plasmodium berghei	56498114	KLICYD**G**FEPSGR	CLTIM**GRTEG**EDNYCS	KVNMLA**R**EYCDIK
Archaebacteria				
Archaeoglobus fulgidus	11498382	PRA-YV**G**YEPSGE	RSMDEV**SRRK**EDPMVS	KIHMLA**R**ENLPRL
Methanopyrus kandleri	20093954	EPVAYV**G**FEPSGK	RSMDMI**ARSE**ENPPVS	KIHMLA**R**DVLPKL
Methanosarcina barkeri	73669252	PRA-YV**G**YEPSGK	RSMDEV**GRAM**DDPTVS	KIHMLA**R**ENLKNL
Methanosarcina mazei	21228065	PRA-YV**G**YEPSGK	RSMDEV**GRAM**DDPTVS	KIHMLA**R**ENLKSL
Methanothermobact.ther.	15679755	PVV-YT**G**YEPSGR	RSMAQI**TRES**RDHKVA	KIHMLA**R**ENLPRL
Methanocaldococcus jan.	15668565	KSA-YI**G**FEPSGK	RSMELI**ARED**ENPKVA	KIHMLA**R**ELLPK-
Pyrococcus abyssi	14521173	LQH-YI**G**FEISGY	RSITIM**GRQM**GESIDF	KAHVIA**I**EVAQKL

Fig.2 Multiple alignment of 24 out of 74 amino acid residues of eukaryotic and archaebacterial TyrRS (conservative domain COG0162.2, according to CDD database) for the surrounding of three mentioned mutations; The most conservative positions are written in bold

Gly41, there are 15 mostly hydrophobic residues, five of which are the aromatic ones. Gly41 residue is located in a small hollow, formed by residues W40, T42, A43, and Q188, and has a significant surface, available for solvent (Fig.4, *b*, table). The values of B-factor for Gly41 C -atom in crystal structures of 1N3L and 1Q11 are 1.73 and 1.29 nm respectively, which points at its insignificant mobility as a part of core -layer of protein (Fig.4, *c*).

The substitution of Gly41 for arginine using Swiss-PdbViewer software suggested 18 rotamers, rating 26 to –3, only four of which did not have conflicts with surrounding (rating 0 – –3). The possibility of formation of two hydrogen bonds between its side group and side groups of Tyr166 and Gln170 was revealed for rotamer Arg41 No18 with the lowest rating (–3) (Fig.3, *b*). At the same time possible conformation of long and flexible side chain Arg41 is limited by neighboring amino acid residues Y39, L72, H77, Y166, Q170, D173, and Q188. Superposition of 9 models, obtained using Modeller 9.2, demonstrated possible significant conformational mobility of side chain Arg41 in the mentioned pocket region.

It has to be noted also that close proximity of one of four tryptophan residues, namely Trp40, which are known to be natural fluorophores in the structure of human TyrRS, may *a priori* influence the change of fluorescent properties at the presence of Gly41Arg mutations.

Glu196Lys mutation. In regard to the secondary structure, Glu196 is localized at the end of -helix H12 and is specific for significant evolutionary conservatism. One more residue occurring in this position is arginine (Fig.1, *c*, Fig.2, Fig.4, *a*). Glu196 is partially exposed on TyrRS surface and there are 13 aa within 0.5 nm radius (L22, 192-FTFA, 197-KYLP, and 206-KRVH), four of which possess the positive charge. Glu196 is almost parallel to TyrRS molecular surface and it forms two hydrogen bonds with atoms of Phe192 and Thr193 polypeptide framework (Fig.3, *b*). The values of B-factor for C -atom of Glu196 are 1.99 and 1.41 nm for structures of 1N3L and 1Q11 respectively (Fig.4, *c*), which testifies to its insignificant mobility.

The substitution of Glu196 for close size-wise lysine using Swiss-PdbViewer suggested 16 different rotamers with ratings from 15 to –2, among which we
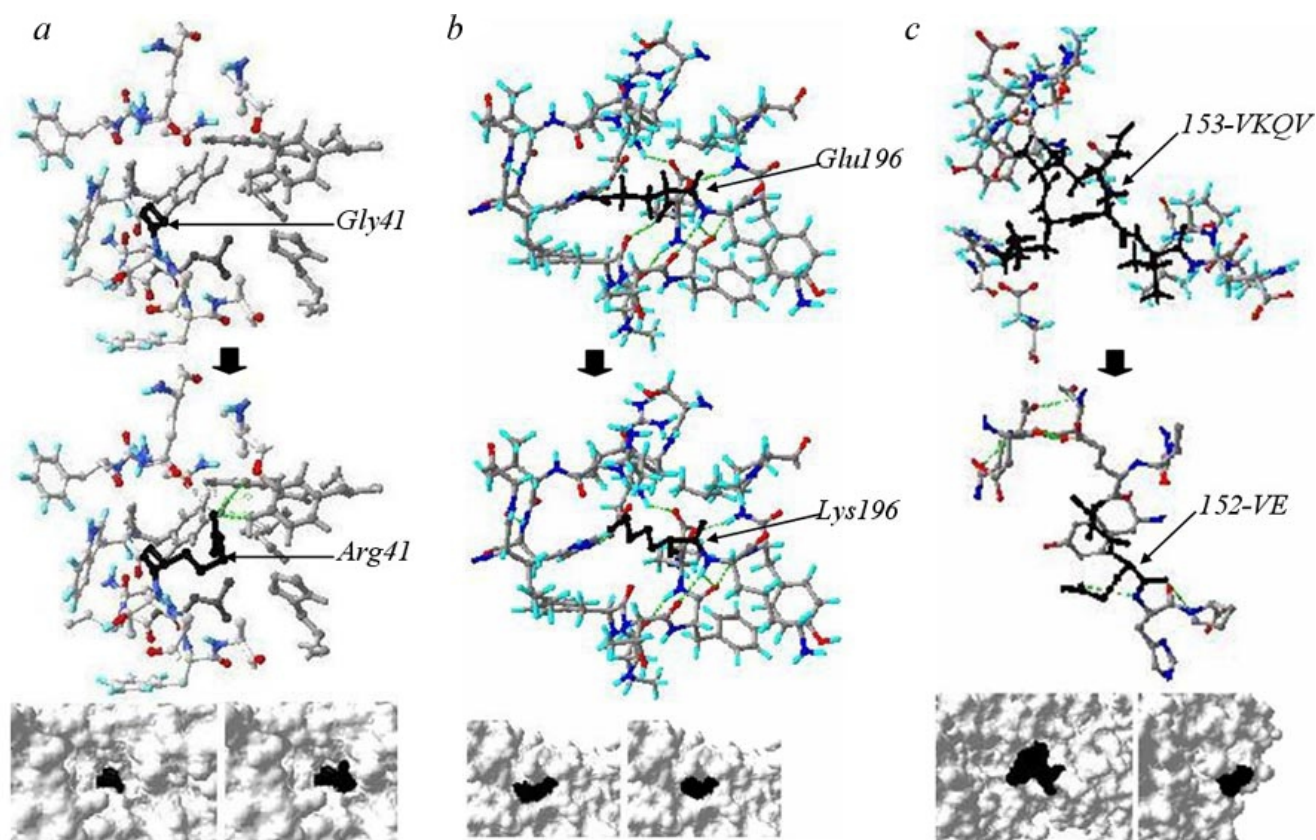
Fig.3 Local environment of amino acid residues, related to mutations, in norm and after computer mutagenesis (marked in bold), the pictures of corresponding equally oriented regions of molecular surface: *a* – Gly41 and 12 residues within 0.6 nm radius (Leu72 forms two hydrogen bonds; for mutant form Gly41Arg, the optimum is the compact conformation of side chain, which allows possible formation of two hydrogen bonds); *b* – Glu196 and 13 residues within 5.0 A (side chain Glu196 does not form hydrogen bonds; for mutant Glu196Lys form there are no significant changes in the location of the side chain and its local environment); *c* – 153–156-VKQV residues and 11 residues within 0.5 nm radius (surface loop Gly149–Leu161 forms the insignificant number of contacts with other protein regions; in case of deletion 153–156, the loop is shortened, from 13 to 9 aa)

selected rotamer No.2, rating –2. Four other rotamers were specific for the same rating and close conformation of side chain. Glu196Lys mutation results in the change of negative charge for the positive one, while no significant changes are observed in the conformation of neighboring residues. Hydrogen bonds between 196 framework remained (Phe192 and Thr193) while new side chain Lys196 lost its hydrogen bond with Arg207 framework.

153–156delVKQV mutation. These amino acid residues are located in the region of so called Connecting Polypeptide 1, CP1, of the domain, which joins two halves of Rossmann fold. The region of Val153–His158 loop between -helices H9 and H10 is one of two partially unregulated regions in mini-TyrRS domain, located on its surface [29]. This segment is a part of specific element for recognition of tRNA^{Tyr} acceptor stem which possibly acquires ordered conformation at binding of tRNA^{Tyr} [29, 41]. Mutation 153–156delVKQV almost matches this region and the deletion of the mentioned residues results in the shortening of the loop – from 13 to 9 aa (Fig.1, *c*). It has to be noted that this region is specific for certain conservatism in amino acid residues of homologous TyrRS of eukaryotes, but not archaebacteria (Fig.2), however, the structure of loop has insignificant influence on the internal structure of protein core.
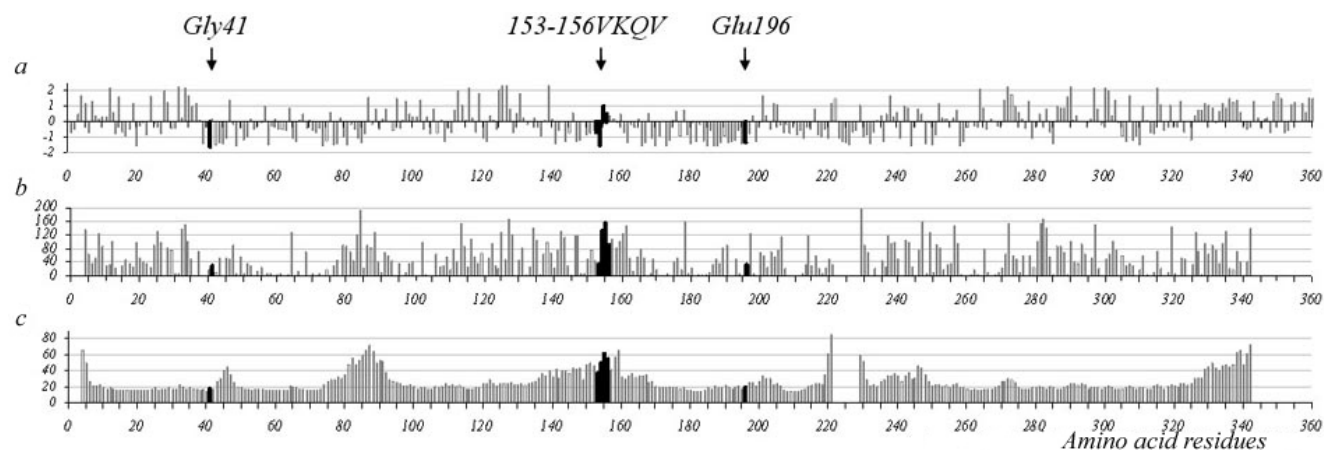
455

Fig.4 Parameters of amino acid residues for human TyrRS N-module (Met1–Pro342, residues with mutations described are marked black; structure of three most mobile residues from N-terminal end of macromolecule and seven residues of catalytic KMSKS loop were not determined in the crystallographic structure, PDB code 1N3L): *a* – degree of evolutionary conservatism of residues as the value of normalized rating provided by ConSurf 3.0 web-server; *b* – degree of exposure as solvent accessible surface area, SASA according to GetArea 1.1 web-server; *c* – degree of mobility of aa, as the value of B-factor for C atoms in crystallographic structure (PDB code 1N3L)

Characteristics of evolutionary conservatism, exposure, and mobility of amino acid residues, related to DI-CMTC mutations

Amino acid residue	Normalized conservatism rating (ConSurf 3.0) (ConSurf 3.0)	Multiple alignment data*	Residue variant**	Solvent accessible surface area, nm² (GetArea 1.1)	B-factor of C -atom	
					1N3L	1Q11
Gly41	−1,554	50/50	G	0,31	17,28	12,91
Val153	−0,684	50/50	A, G, K, Q, S, T, V	0,36	37,75	25,18
Lys154	−1,469	50/50	K, Q, R	1,30	49,09	28,19
Gln155	1,025	50/50	A, E, F, K, N, Q, R, S, T, Y	1,56	60,51	31,45
Val156	0,514	50/50	A, D, E, G, K, M, N, Q, R, S, T, V	0,91	55,18	35,93
Glu196	−1,338	50/50	E, I, R	0,34	19,85	14,08

* number of aligned non-gapped amino acid residues of TyrRS, containing this residue, out of total 50 least homologous sequences for each position;
** amino acid residues, occurring in the current position of multiple alignment of sequences

Within 0.5 nm radius from these four loop residues there are nine amino acid residues, neighboring in the chain (149-GAEV, 157-EHP, G163, and Y166), with which they form only one hydrogen bond (Fig.3, *c*). The values of B-factor for C atoms of 153–156VKQV region are within the range of 3.78–6.05 nm for 1N3L, which corresponds to the peak mobility value of polypeptide framework and is expectable for surface loop residues (Fig.4, *c,* Table).

Possible influence of mutations on TyrRS-substrates interaction. Stating on the crystallographic structures of archaebacteria and yeast TyrRS, it is possible to foresee the existence of evolutionary conservative residues, forming hydrogen bonds with L-tyrosine
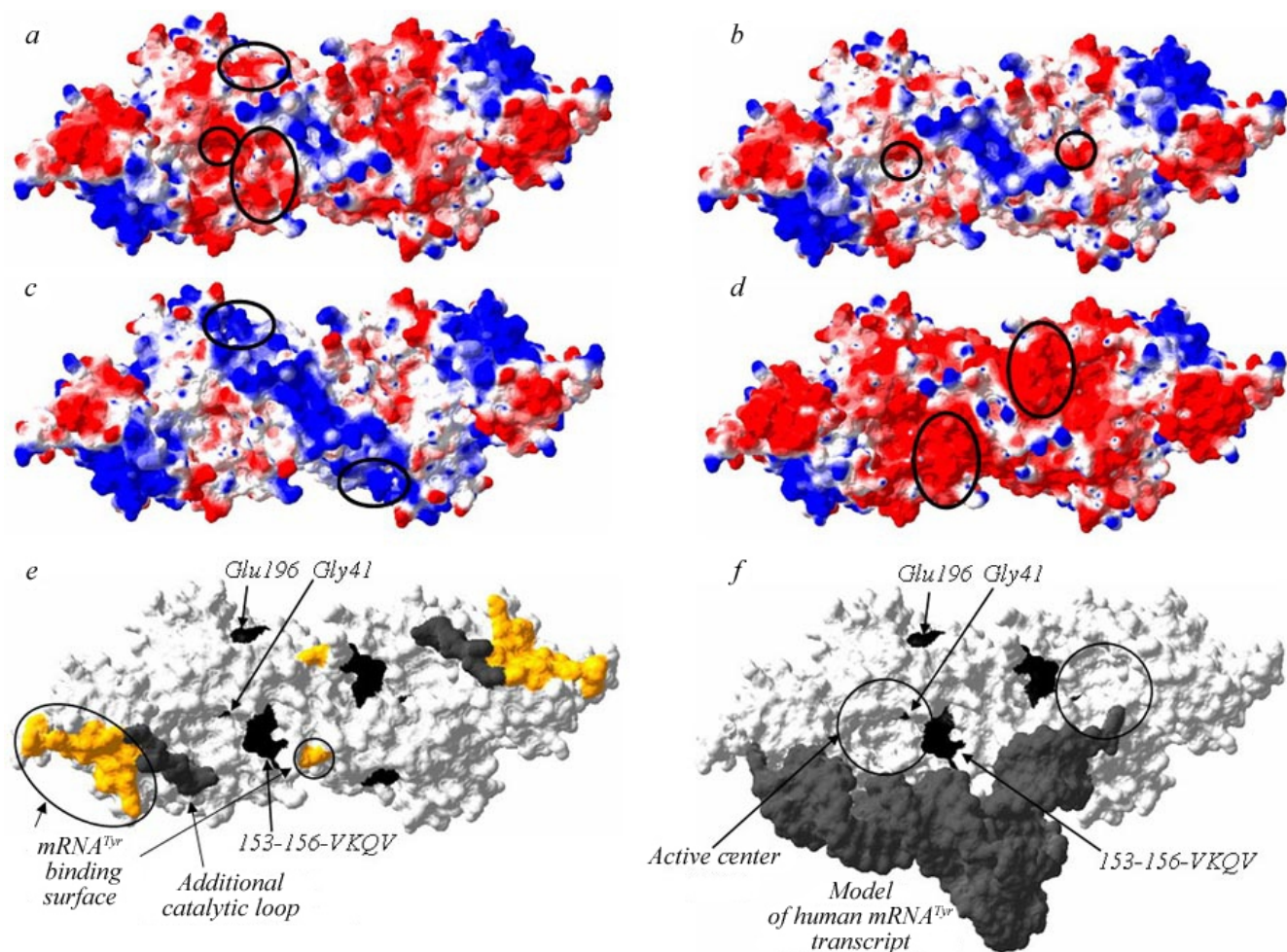
Fig.5 Location of three amino acid residues of human cytoplasmic TyrRS, corresponding to Gly41Arg, Glu196Lys, and 153–156delVKQV mutations, regarding the spatial structure of its N-module dimer or mini-TyrRS (PDB code 1N3L) in norm and in mutant protein forms (interdomain catalytic KMSKS loop was completed using ModLoop web server; the picture, obtained using Swiss-PdbViewer 3.7, is presented in top-view only): *a* – molecular surface of N-module dimer in norm, stained according to electrostatic potential (electrostatic potential was calculated using *Coulomb* method, taking into account partial atomic charges); *b* – the same for Gly41Arg mutation in both subunits; *c* – the same for Glu196Lys in both subunits; *d* – the same for 153–156delVKQV deletion in both subunits; *e* – location of DI-CMTC sites of mutations regarding aa, which form tRNA^{Tyr}-binding surface, except its 3'-CCA-terminal end (grey color, two ovals); *f* – model of N-module dimer complex with non-modified human tRNA^{Tyr} transcript without 3'-CCA-terminal end (stoichiometry 2:1)

and adenine AMP base. In case of human TyrRS they are as follows: Y39, Y166, Q170, D173, Q188, and Val215. Residues of catalytic center are as follows: 41-GTA, H49, 51-AY, Q182, 184-GG, D-187, M214, and 222-KMSSS (Fig.1, *c*). It is expectable that structural changes due to the occurrence of DI-CMTC mutations may have indirect (distal) influence on the binding

of low-molecular substrates. First of all, it concerns Gly41Arg mutation, which is the neighboring one to the majority of L-tyrosine binding residues.

Comparative analysis of N-module dimer molecular surfaces, stained by electrostatic potential, in norm and in cases of three mutations reveals that the distribution of electrostatic potential on the surface of molecule

is subject to significant changes and very sensitive to certain mutations (Fig.5, *a–d*).

Possible effect of mutations on TyrRS-tRNA^Tyr interaction. To study the possible role of residues, involved in mutation, in binding of tRNA^Tyr macromolecule, we designed the complex of human TyrRS N-module dimer with non-modified transcript of its tRNA^Tyr, not including the mobile 3'-CCA-end of the latter (stoichiometry 2:1), on the basis of homologous complexes of archaebacteria *M. jannaschii* and yeast *S. cerevisiae* (PDB codes 1J1U and 2DLC). The foreseen contacts of this model between human TyrRS with tRNA^Tyr in norm revealed 14 aa, namely, H158, 246-KKAFCEP, W283, 305-HPGD, and K310, which form the contact surface. Consequently, we determined close location of Glu196 and tRNA molecule in designed TyrRS-tRNA^Tyr complex and defined possible effect of deletions of 153–156-VKQV residues on the location of neighboring His158, which interacts directly with acceptor stem of tRNA^Tyr (Fig.5, *e*, *f*).

It has to be mentioned that all three mutations, Gly41Arg in the first place, are the potential factors of incorrect folding of mutant TyrRS forms, which results in the decrease in their solubility and increase in the capacity to form macromolecular aggregates.

The other possibility is in the fact that the changes in molecular surface of these mutant TyrRS forms may influence the properties of protein-protein interactions of TyrRS and its partners in norm or in case of any new non-typical interactions with other components of the cell.

Currently, non-synonymous single nucleotide polymorphisms, altering amino acid TyrRS sequence, were determined for human *YARS* gene. In particular, these mutations are as follows: Arg34Gln, Met56Ile, Lys114Thr, Glu128Asp, and Gln170His in exons 2, 2, 3, 4, and 4, respectively. These mutations are capable of altering TyrRS properties and may be connected with currently undiscovered hereditary pathologies.

In the course of further investigation we plan to model molecular dynamics of human mini-TyrRS dimer in norm and in three obtained mutant forms in aquatic ionic environment using GROMACS 3.3.2 software pack to carry out the comparative analysis and to study possible changes in dynamic features of mutant TyrRS forms.

К. А. Одынец, А. И. Корнелюк

Биоинформационный анализ влияния мутаций тирозил-тРНК синтетазы человека, ассоциированных с нейропатией Шарко-Мари-Тус типа С, на локальные особенности ее пространственной структуры

Резюме

Недавно установлена причинная связь между тремя определенными мутантными формами цитоплазматической тирозил-тРНК синтетазы (TyrRS) человека и доминантной нейропатией Шарко-Мари-Тус типа С (DI-CMTC). В данной работе для двух точечных мутаций Gly41Arg, Glu196Lys и делеции 153–156VKQV проведено биоинформационное исследование влияния замены аминокислотных остатков на структурные особенности димера каталитического модуля TyrRS («мини-TyrRS», 2 39 кДа), в частности, на ее ближайшее окружение и электростатический потенциал молекулярной поверхности.

Ключевые слова: тирозил-тРНК синтетаза, нейропатия Шарко-Мари-Тус типа С, пространственная структура, компьютерный мутагенез.

REFERENCES

1. *Park S. G., Ewalt K. L., Kim S.* Functional expansion of aminoacyl-tRNA synthetases and their interacting factors: new perspectives on housekeepers // Trends Biochem. Sci.–2005.–**30**.–P. 569–574.

2. *Ivakhno S. S., Kornelyuk A. I.* Cytokine-like activities of some aminoacyl-tRNA synthetases and auxiliary p43 cofactor of aminoacylation reaction and their role in oncogenesis // Exp. Oncol.–2004.–**26**.–P. 250–255.

3. *Antonellis A., Lee-Lin S. Q., Wasterlain A., Leo P., Quezado M., Goldfarb L. G., Myung K., Burgess S., Fischbeck K. H., Green E. D.* Functional analyses of glycyl-tRNA synthetase mutations suggest a key role for tRNA-charging enzymes in peripheral axons // J. Neurosci.–2006.–**26**.–P. 10397–10406.

4. *Nangle L. A., Zhang W., Xie W., Yang X. L., Schimmel P.* Charcot-Marie-Tooth disease-associated mutant tRNA synthetases linked to altered dimer interface and neurite distribution defect // Proc. Nat. Acad. Sci. USA.–2007.–**104**.–P. 11239–11244.

5. *Jordanova A., Irobi J., Thomas F. P., Van Dijck P., Meerschaert K., Dewil M., Dierick I., Jacobs A., De Vriendt E., Guergueltcheva V., Rao C. V., Tournev I., Gondim F. A., D'Hooghe M., Van Gerwen V., Callaerts P., Van Den Bosch L., Timmermans J. P., Robberecht W., Gettemans J., Thevelein J. M., De Jonghe P., Kremensky I., Timmerman V.* Disrupted function and axonal distribution of mutant tyrosyl-tRNA synthetase in dominant intermediate Charcot-Marie-Tooth neuropathy // Nat. Genet.–2006.–**38**.–P. 197–202.

6. *Jordanova A., Thomas F. P., Guergueltcheva V., Tournev I., Gondim F. A. A., Ishpekova B., De Vriendt E., Jacobs A., Litvinenko I., Ivanova N., Buzhov B., De Jonghe P., Kremensky I., Timmerman V.* Dominant intermediate Charcot-Marie-Tooth type C maps to chromosome 1p34-p35 // Amer. J. Hum. Genet.–2003.–**73**.–P. 1423–1430.

7. *Корнелюк А. И.* Структурно-функциональное исследование тирозил-тРНК синтетазы млекопитающих // Биополимеры и клетка.–1998.–**14**.–С. 349–359.

8. Kornelyuk A. I., Tas M. P. R., Dubrovsky A., Murray J. C. Cytokine activity of the non-catalytic EMAP 2-like domain of mammalian tyrosyl-tRNA synthetase // Биополимеры и клетка.–1999.–**15**.–P. 168–172.

9. *Lee J. W., Beebe K., Nangle L. A., Jang J., Longo-Guess C. M., Cook S. A., Davisson M. T., Sundberg J. P., Schimmel P., Ackerman S. L.* Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration // Nature.–2006.–**443.**–P. 50–55.

10. *Roy H., Ibba M.* Molecular biology: sticky end in protein synthesis // Nature.–2006.–**443**.–P. 41–42.

11. *Paley E. L., Denisova G., Sokolova O., Posternak N., Wang X., Brownell A. L.* Tryptamine induces tryptophanyl-tRNA synthetase-mediated neurodegeneration with neurofibrillary tangles in human cell and mouse models // Neuromol. Med.–2007.–**9**.–P. 55–82.

12. *Paley E. L., Smelyanski L., Malinovskii V., Subbarayan P. R., Berdichevsky Y., Posternak N., Gershoni J. M., Sokolova O., Denisova G.* Mapping and molecular characterization of novel monoclonal antibodies to conformational epitopes on NH(2) and COOH termini of mammalian tryptophanyl-tRNA synthetase reveal link of the epitopes to aggregation and Alzheimer's disease // Mol. Immunol.–2007.–**44**.–P. 541–557.

13. *Berger P., Niemann A., Suter U.* Schwann cells and the pathogenesis of inherited motor and sensory neuropathies (Charcot-Marie-Tooth disease) // Glia.–2006.–**54**.–P. 243–257.

14. *Meyer zu Horste G., Prukop T., Nave K. A., Sereda M. W.* Myelin disorders: Causes and perspectives of Charcot-Marie-Tooth neuropathy // J. Mol. Neurosci.–2006.–**28**.–P. 77–88.

15. *Vallat J. M., Tazir M., Magdelaine C., Sturtz F., Grid D.* Autosomal-recessive Charcot-Marie-Tooth diseases // J. Neuropathol. Exp. Neurol.–2005.–**64**.–P. 363–370.

16. *Shy M. E.* Charcot-Marie-Tooth disease: an update // Curr. Opin. Neurol.–2004.–**17**.–P. 579–585.

17. *Roa B. B., Greenberg F., Gunaratne P., Sauer C. M., Lubinsky M. S., Kozma C., Meck J. M., Magenis R. E., Shaffer L. G., Lupski J. R.* Duplication of the PMP22 gene in 17p partial trisomy patients with Charcot-Marie-Tooth type-1 neuropathy // Hum. Genet.–1996.–**97**.–P. 642–649.

18. *Bruzzone R., White T. W., Scherer S. S., Fischbeck K. H., Paul D. L.* Null mutations of connexin 32 in patients with X-linked Charcot-Marie-Tooth disease // Neuron.–1994.–**13**.–P. 1253–1260.

19. *Ionasescu V. V.* Charcot-Marie-Tooth neuropathies: from clinical description to molecular genetics // Muscle Nerve.–1995.–**18**.–P. 267–275.

20. *Guilbot A., Williams A., Ravise N., Verny C., Brice A., Sherman D. L., Brophy P. J., LeGuern E., Delague V., Bareil C., Megarbane A., Claustres M.* A mutation in periaxin is responsible for CMT4F, an autosomal recessive form of Charcot-Marie-Tooth disease // Hum. Mol. Genet.– 2001.–**10**.–P. 415–421.

21. *Bolino A., Muglia M., Conforti F. L., LeGuern E., Salih M. A., Georgiou D. M., Christodoulou K., Hausmanowa-Petrusewicz I., Mandich P., Schenone A., Gambardella A., Bono F., Quattrone A., Devoto M., Monaco A. P.* Charcot-Marie-Tooth type 4B is caused by mutations in the gene encoding myotubularin-related protein-2 // Nat. Genet.–2000.–**25**.–P. 17–19.

22. *Kleeman T. A., Wei D., Simpson K. L., First E. A.* Human tyrosyl-tRNA synthetase shares amino acid sequence homology with a putative cytokine // J. Biol. Chem.–1997.–**272**.–P. 14420–14425.

23. *Wakasugi K., Slike B. M., Hood J., Ewalt K. L., Cheresh D. A., Schimmel P.* Induction of angiogenesis by a fragment of human tyrosyl-tRNA synthetase // J. Biol. Chem.–2002.–**277**.–P. 20124–20126.

24. *Jia J., Li B., Jin Y., Wang D.* Expression, purification, and characterization of human tyrosyl-tRNA synthetase // Protein Exp. Purif.–2003.–**27**.–P. 104–108.

25. *Wakasugi K., Schimmel P.* Highly differentiated motifs responsible for two cytokine activities of a split human tRNA synthetase // J. Biol. Chem.–1999.–**274**.–P. 23155–23159.

26. *Wakasugi K., Schimmel P.* Two distinct cytokines released from a human aminoacyl-tRNA synthetase // Science.–1999.–**284**.–P. 147–151.

27. *Ewalt K. L., Schimmel P.* Activation of angiogenic signaling pathways by two human tRNA synthetases // Biochemistry.–2002.–**41**.–P. 13344–13349.

28. *Tzima E., Schimmel P.* Inhibition of tumour angiogenesis by a natural fragment of a tRNA synthetase // Trends Biochem. Sci.–2006.–**31**.–P. 7–10.

29. *Yang X. L., Skene R. J., McRee D. E., Schimmel P.* Crystal structure of a human aminoacyl-tRNA synthetase cytokine // Proc. Nat. Acad. Sci. USA.–2002.–**99**.–P. 15369–15374.

30. *Yang X. L., Otero F. J., Skene R. J., McRee D. E., Schimmel P., Ribas de Pouplana L.* Crystal structures that suggest late development of genetic code components for differentiating aromatic side chains // Proc. Nat. Acad. Sci. USA.–2003.–**100**.–P. 15376–15380.

31. *Kobayashi T., Nureki O., Ishitani R., Yaremchuk A., Tukalo M., Cusack S., Sakamoto K., Yokoyama S.* Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion // Nat. Struct. Biol.–2003.–**10**.–P. 425–432.

32. *Tsunoda M., Kusakabe Y., Tanaka N., Ohno S., Nakamura M., Senda T., Moriguchi T., Asai N., Sekine M., Yokogawa T., Nishikawa K., Nakamura K. T.* Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms // Nucl. Acids Res.–2007.–**35**.–P. 4289–4300.

33. *Berman H. M., Westbrook J., Feng Z., Gilliland G., Bhat T. N., Weissig H., Shindyalov I. N., Bourne P. E.* The Protein Data Bank // Nucl. Acids Res.–2000.–**28**, N 1.–P. 235–242.

34. *Laskowski R. A., Chistyakov V. V., Thornton J. M.* PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids // Nucl. Acids Res.–2005.–**33**, Database issue.–P. D266–D268.

35. *Guex N., Peitsch M. C.* SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling // Electrophoresis.–1997.–**18**.–P. 2714–2723.

36. *Fiser A., Sali A.* ModLoop: automated modeling of loops in protein structures // Bioinformatics.–2003.–**18**.–P. 2500–2501.

37. *Schwede T., Kopp J., Guex N., Peitsch M. C.* Swiss-Model: an automated protein homology-modeling server // Nucl. Acids Res.–2003.–**31**.–P. 3381–3385.

38. *Marti-Renom M. A., Stuart A., Fiser A., Sanchez R., Melo F., Sali A.* Comparative protein structure modeling of genes and genomes // Annu. Rev. Biophys. Biomol. Struct.–2000.–**29**.–P. 291–325.

39. *Fraczkiewicz R., Braun W.* Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules // J. Comp. Chem.–1998.–**19**.–P. 319–333.

40. *Landau M., Mayrose I., Rosenberg Y., Glaser F., Martz E., Pupko T., Ben-Tal N.* ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information // Nucl. Acids Res.–2005.–**33**, Web Server issue.–P. W299–W302.

41. *Wakasugi K., Quinn C. L., Tao N., Schimmel P.* Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant // EMBO J.–1998.–**17**.–P. 297–305.