Research Article

Journal of Molecular Recognition

Received: 10 October 2012, Revised: 20 November 2012, Accepted: 10 December 2012, Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jmr.2259

Asymmetric structure and domain binding interfaces of human tyrosyl-tRNA synthetase studied by molecular dynamics simulations

Oleksandr V. Savytskyi[a]*, Semen O. Yesylevskyy[b] and Alexander I. Kornelyuk[a]

Human tyrosyl-tRNA synthetase (*Hs*TyrRS) is composed of two structural modules: N-terminal catalytic core and an EMAP II-like C-terminal domain. The structures of these modules are known, but no crystal structure of the full-length *Hs*TyrRS is currently available. An all-atom model of the full-length *Hs*TyrRS was developed in this work. The structure, dynamics, and domain binding interfaces of *Hs*TyrRS were investigated by extensive molecular dynamics (MD) simulations. Our data suggest that *Hs*TyrRS in solution consists of a number of compact asymmetric conformations, which differ significantly by their rigidity, internal mobility, orientation of C-terminal modules, and the strength of interdomain binding. Interfaces of domain binding obtained in MD simulations are in perfect agreement with our previous coarse-grained hierarchical rotations technique simulations. Formation of the hydrogen bonds between R93 residue of the ELR cytokine motif and the residues A340 and E479 in the C-module was observed. This observation supports the idea that the lack of cytokine activity in the full-length *Hs*TyrRS is explained by interactions between N-modules and C-modules, which block the ELR motif.

Supporting information may be found in the online version of this paper.

Keywords: tyrosyl-tRNA synthetase; cytokine; ELR motif; molecular dynamics; interdomain interface; domain motion; HIEROT; GROMACS; Pteros

INTRODUCTION

Aminoacyl-tRNA synthetases (aaRSs) are key enzymes of protein biosynthesis that specifically attach each amino acid to its cognate tRNA (Woese *et al*., 2000; Mirande, 1991). Usually aaRSs are modular proteins, with the domains that either play distinct roles in tRNA aminoacylation reaction or possess non-canonical functions, which are not related to translation (Mirande, 1991; Guo *et al*., 2010, 2010). There are functional communications between the domains of aaRSs, which facilitate their coordinated functioning (Alexander and Schimmel, 2001; Zhang and Hou, 2005; Weimer *et al*., 2009).

Tyrosyl-tRNA synthetases (TyrRSs) are homodimeric class I aaRSs, which comprise catalytic domain containing the Rossmann fold. Truncated N-terminal catalytic modules of mammalian TyrRS (mini-TyrRS) retain full enzymatic activity *in vitro* (Kornelyuk *et al*., 1988; Gnatenko *et al*., 1991). The non-catalytic C-terminal domain of mammalian TyrRSs is homologous to EMAP II cytokine (endothelial monocyte-activating polypeptide II) (Kleeman *et al*., 1997; Levanets *et al*., 1997), which stimulates endothelial-dependent coagulation *in vitro* and apparently plays an important role in inflammation, apoptosis, and angiogenesis in tumor tissues (Berger, 2000; Reznikov *et al*., 2007). The full-length TyrRS does not have cytokine activity, but its proteolytic cleavage reveals IL8-like activity of the N-terminal catalytic module and EMAP II-like activity of non-catalytic C-terminal domain (Wakasugi and Schimmel, 1999a, 1999b; Kornelyuk *et al*., 1999). It was established that the ELR-motif (E91, L92, R93) is responsible for IL8-like cytokine activity of mini-TyrRS (Wakasugi and Schimmel, 1999a, 1999b).

Catalytic N-terminal module of human tyrosyl-tRNA synthetase (*Hs*TyrRS) (residues 1–342) and its non-catalytic EMAP II-like C-domain (residues 360–528) are linked by long disordered linker (residues 343–359). Human mini-TyrRS forms rigid dimeric structure, whereas its C-domains possess significant mobility due to unstructured linkers. The crystal structures of the isolated N-terminal mini-TyrRS and C-terminal module of *Hs*TyrRS were resolved (Yang *et al*., 2002; Yang *et al*., 2003), but crystallization of the full-length protein is difficult because of its extreme conformational flexibility. As a result, the compact structure of the full-size *Hs*TyrRS remains unknown. Recently, we investigated compactization of the *Hs*TyrRS and domain binding interfaces between N-modules and C-modules using the coarse-grained hierarchical rotations technique (HIEROT; Yesylevskyy *et al*., 2011). Present work is a logical continuation of this research.

In this work, we developed an all-atom model of the full length *Hs*TyrRS dimer and performed extensive molecular dynamics (MD) simulations of this enzyme. Formation of the compact conformations of *Hs*TyrRS was studied. Interfaces of

* *Correspondence to: O. V. Savytskyi, Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Akademika Zabolotnogo Str., 150, Kyiv-03680, Ukraine.*
E-mail: savytskyi@moldyngrid.org

a *O. V. Savytskyi, A. I. Kornelyuk*
Institute of Molecular Biology and Genetics, National Academy of Sciences of Ukraine, Akademika Zabolotnogo Str., 150, Kyiv-03680, Ukraine

b *S. O. Yesylevskyy*
Institute of Physics, National Academy of Sciences of Ukraine, Prospect Nauki, 46, Kyiv-03039, Ukraine

interdomain binding, domain binding energies, and internal mobility of the domains were investigated. Systematic comparison with the data of our previous coarse-grained HIEROT simulations is performed. Hydrogen bonding patterns of the ELR-motif, which is related to cytokine activity of mini-*Hs*TyrRS, were investigated and compared with recent experimental results.

METHODS

Structural modeling

The structure of the full length *Hs*TyrRS monomer was constructed from the crystal structures of its N-terminal and C-terminal modules (PDB codes 1N3L:A, residues A3–P342 and 1NTG:A residues P360–S528, respectively) by using Modeller 9.7 software (Marti-Renom *et al*., 2000; Eswar *et al*., 2006). Missing N-terminal residues M1–D3, the residues of the catalytic loop K222–E228 and the linker residues D343–E359 were added using loops reconstruction option in Modeller 9.7. Homology modeling with multiple templates protocol was used to generate five ensembles of 100 structures each (see Supplementary Material). Ten best structures were selected from each ensemble using the Modeller Objective Function, the Discrete Optimized Protein Energy score and the normalized Discrete Optimized Protein Energy score (Shen and Sali, 2006). Selected structures were verified using the MolProbity web-server (Chen *et al*., 2010) to ensure the absence of sterical clashes, and so forth. Single structure was used as a starting structure in all subsequent MD simulations.

Molecular dynamics

All simulations were performed using GROMACS 4.0.5 MD package with the GROMOS 43a1 force field. Single Point Charge (SPC) water model was used with the bond lengths constrained by SETTLE algorithm (Miyamoto and Kollman, 1992). All other bonds were constrained using LINCS (Hess *et al*., 1997) algorithm. The temperature of 310 K was maintained by coupling the protein and the solvent to separate v-rescale thermostats with the relaxation times of 0.1 ps. The constant pressure was maintained by the isotropic Parrinello–Rahman barostat with the relaxation constant of 1 ps. The usage of virtual sites and heavy hydrogen atoms (Feenstra *et al*., 1999) allowed the time step of 4 fs in all simulations. Lennard-Jones and Coulombic interactions were computed explicitly within 1 nm cut-off range. Long-range electrostatic interactions were computed using particle mesh Ewald method (Van der Spoel *et al*., 2005) with a grid spacing of 0.12 nm.

The protein was solvated in a rhombic octahedron box with minimal distance between the protein and the box of 1 nm. The system contained ~127 000 water molecules, 355 Na^+ ions, and 351 Cl^- ions, which simulates salt concentration of 0.15 mol/l and counterbalances the charge of protein. MD simulations with position restraints applied to heavy protein atoms were performed for 10 ps. After that, six independent production runs were performed for 100 ns each. Coordinates were saved every 4 ps for further analysis. The equilibration of the protein was monitored by the root mean square deviation (RMSD) of C_α atoms.

Analysis

The analysis of MD trajectories was performed using GROMACS 4.5.5 suite of programs (Hess *et al*., 2008) and visual molecular dynamics (VMD; Humphrey *et al*., 1996). The interfaces of C-terminal and N-terminal modules for each trajectory were determined by the custom Contacts_finder program based on Pteros molecular modeling library (https://sourceforge.net/projects/pteros/) (Yesylevskyy, 2012) and custom Tcl scripts for VMD molecular analysis program. The clustering analysis of domain interfaces was performed by the custom Contacts Analyzer Script written in Tcl. Analysis programs and scripts were launched using Distributed Analyzer Script, which allowed their parallel execution and proper load balancing in the computational cluster (Savytskyi *et al*., 2011).

Classification of domain interfaces

In order to compare the resulting compact structures of *Hs*TyrRS, obtained in MD simulations, the following procedure was used. The interfaces of both C-terminal modules of *Hs*TyrRS with the dimer of N-terminal modules were determined at the residue level. The atoms of C and N-modules were considered to be in contact if the distance between their centers was smaller than the cut-off distance. The residues were considered to be in contact if at least one pair of atoms from them was in contact in any trajectory frame. The lists of contacting residues were constructed for all studied trajectories. The contacts formed by both C-modules were symmetrized. This allows comparing interfaces created by both C-modules directly using the same residue numbering. These lists were subject to clustering analysis using the algorithm developed in our previous work (Yesylevskyy *et al*., 2011):

(i) Each studied trajectory is considered to be a cluster of size 1.
(ii) The similarity matrix is computed as $s_{ij} = a_{ij}/b_{ij}$, where a_{ij} is the number of contacts shared by all $n_i + n_j$ trajectories in the clusters i and j (n_i is the number of trajectories in the cluster i); b_{ij} is the number of contacts, shared by any two or more trajectories in the clusters i and j.
(iii) The largest matrix element s_{ij} is found, and the clusters i and j are merged.
(iv) The steps 3 and 4 are repeated while $s_{ij} < c$, where c is a critical ratio of the shared contacts ($c = 40\%$ in this study).

As a result, all trajectories are subdivided into clusters, which share more than c residue–residue contacts between one of the C-modules and the dimer of N-modules. The trajectories inside the same cluster could be classified as having the domain interface of the same type, whereas the domain interfaces of the trajectories from different clusters do not show any significant similarity. It is necessary to emphasize that this analysis is performed using the pairs of contacting residues, thus the same individual residue may be present in different interface types.

Two different values of the cut-off were used. The cut-off of 2.5 Å corresponds to the closest and strongest non-bound contacts between the domains, whereas the cut-off of 3.5 Å covers all contacts, including weak interactions.

Analysis of the hydrogen bonding was performed by g_hbond program from GROMACS 4.5.5 package using simple geometric criterion (the donor–acceptor distance smaller than 3.5 Å and the hydrogen bond angle smaller than 30°).

Time-resolved RMSF calculations

In order to study the changes of mobility of different residues during MD simulations, the time-resolved root mean square fluctuation (tRMSF) protocol was used (Bennion and Daggett, 2004; Verli and Guimarães, 2005; Goette *et al*., 2009). In this

Journal of Molecular Recognition

protocol, the RMSF of each atom in the protein are computed inside small time windows along the trajectory

$$tRMSF(i,w) = \sqrt{\frac{1}{N\Delta t}\sum_{j=1}^{N}\left(x_i(t_{w+j}) - x_i(t_w)\right)^2},$$

where i is the atom index, w is the number of time window, N is the number of trajectory frames in each window, t_w is the time of the first frame in window w, and Δt is the time interval between the frames. Within each window, the protein was aligned to the structure from the first frame of that window using all heavy atoms. This presumes that the starting structure of each window is considered to be in local equilibrium, and the fluctuations around this structure are computed. This approximation is valid if the windows are sufficiently small in comparison with characteristic times of large conformational changes in the protein. The windows of 200 ps are used in this work.

tRMSF of the residues are computed as mass-weighted averages of tRMSFs of individual atoms

$$tRMSF_r(r,w) = \sum_{i\in ind_r}(tRMSF(i,w)\cdot m_i)/\sum_{i\in ind_r} m_i,$$

where r is the residue index, ind_r is the list of indexes of atoms in the residue r, m_i is the mass of atom i.

Technical details

Simulations and analysis were partially performed in the Grid environment in Ukrainian National Grid infrastructure by using the MolDynGrid virtual laboratory (http://moldyngrid.org) (Salnikov *et al*., 2009; Salnikov *et al*., 2010). VMD program was used for structure visualization. All custom software used in this work is available by request. The Pteros molecular modeling library is freely distributed under Artistic License at https://sourceforge.net/projects/pteros/.

RESULTS

The RMSD of C_α atoms from the starting structure in six production runs is shown in Figure 1. RMSD stabilizes after 40 ns in all simulations; however, the amplitude of its equilibrium fluctuations is very different in different trajectories. Because RMSD does not grow significantly after 40 ns in all simulations, the structures were considered equilibrated after this time, and last 60 ns of trajectories were used for further analysis.

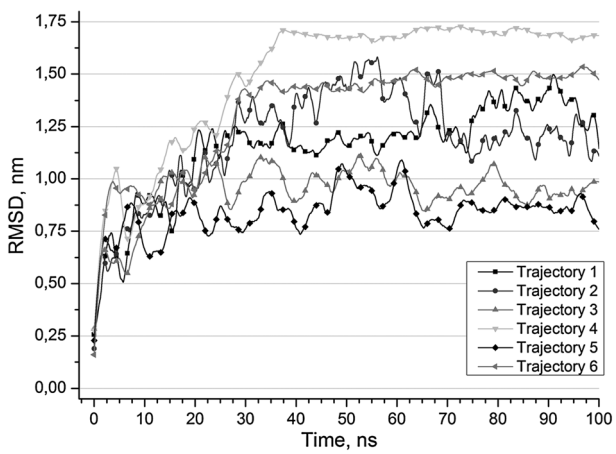

Figure 1. Root mean square deviation (RMSD) of C_α atoms for six production MD trajectories. The curves are smoothed for clarity.

All six MD simulations produced compact final structures, where both C-modules are bound to the dimer of N-modules. There is a pronounced asymmetry in the binding of C-modules, which is explained by their almost independent motions (Figure 2).

The tRMSF maps for three selected MD simulations are shown in Figure 3. Vertical lines in these maps correspond to the evolution of RMSF of individual residues in time, and the changes of color along these lines indicate the changes of mobility of the corresponding residue.

We computed the binding energies between each of C-modules and the dimer of N-modules in all simulations (Figure 4). This energy contains short-range Coulomb and Lennard-Jones interaction terms. It is clearly seen that the evolution of the binding energy (defined as a sum of non-bonded interactions) differs dramatically in different simulations. The strongest binding energy of ~1000 kJ/mol is observed for the second C-module in simulation 1 and for the first C-module in simulation 6. In contrast, the binding energies in simulations 2, 3, and 5 are much smaller (300–400 kJ/mol).

Although there is no unique binding interface between C-modules and N-modules in our simulations, some residues forms contact more often than the others. We computed the probability of each residue from both C-modules and N-modules to be in contact with any residue from another module as $P_c = N_c/N_{sim}$, where N_c is the number of simulations so that a given residue forms at least one contact, N_{sim} is the total number of simulations. P_c was computed for each monomer separately, and then averaged for both of them in order to obtain symmetric picture for identical monomers. The same procedure of finding P_c was used

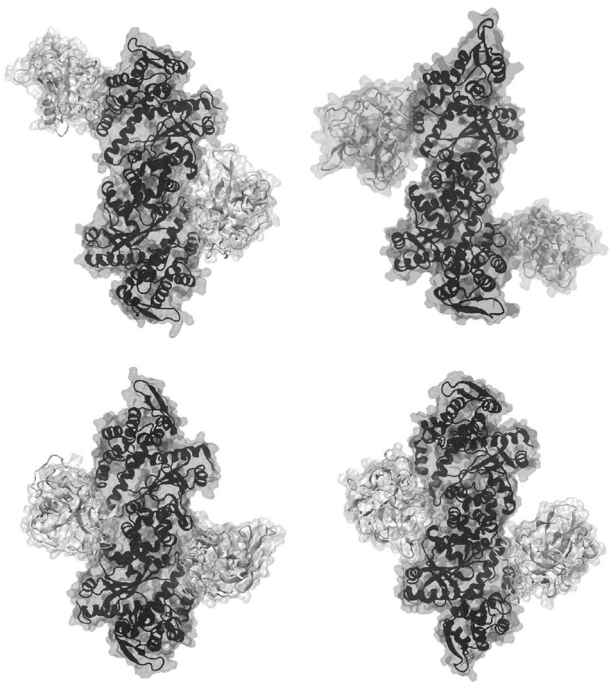

Figure 2. Final compact structures *Hs*TyrRS in MD trajectories 1, 3, 4, and 6. The dimer of N-modules is black, the C-modules are white. The protein is shown in the secondary structure representation with transparent molecular surface around. Different arrangements of C-modules are clearly visible.

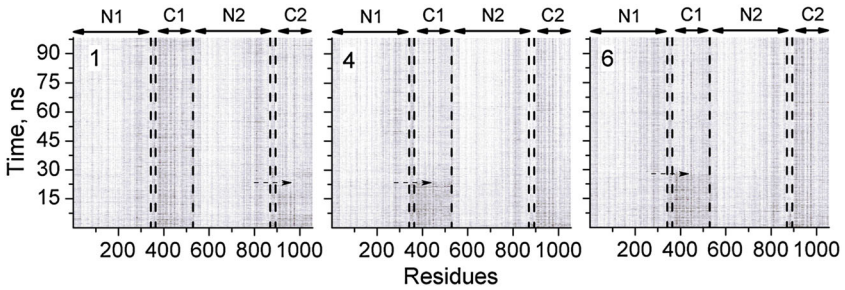

Figure 3. The tRMSF maps for six production MD trajectories. Black corresponds to the highest mobility, white to the lowest. Dashed lines show the boundaries of N-modules and C-modules. Arrows show abrupt decreases of mobility of the first C-module in trajectories 4, 6 and of the second C-module in trajectory 1.

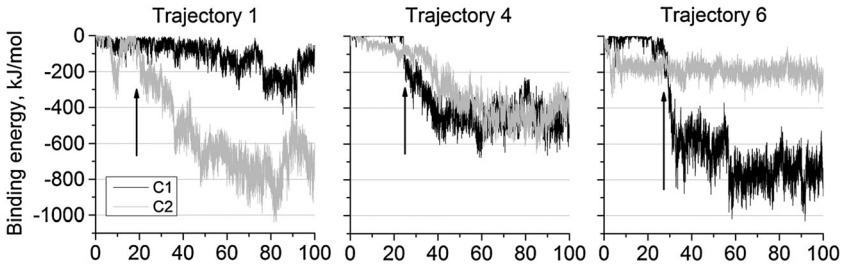

Figure 4. Binding energies of the C-modules with the dimer of N-modules for three selected production MD trajectories.

in our previous paper, where domain compactization in *Hs*TyrRS was studied using the HIEROT technique (Yesylevskyy *et al.*, 2011); thus, the data obtained by both methods are directly comparable. It is necessary to note, however, that the force fields used in MD and HIEROT simulations are different, and the comparison of the binding energies and other properties dependent on the force field is not possible.

Clustering analysis of 12 domain interfaces (two interfaces of C-modules in each trajectory) revealed four distinct types of interfaces, which share more than 40% of residue–residue

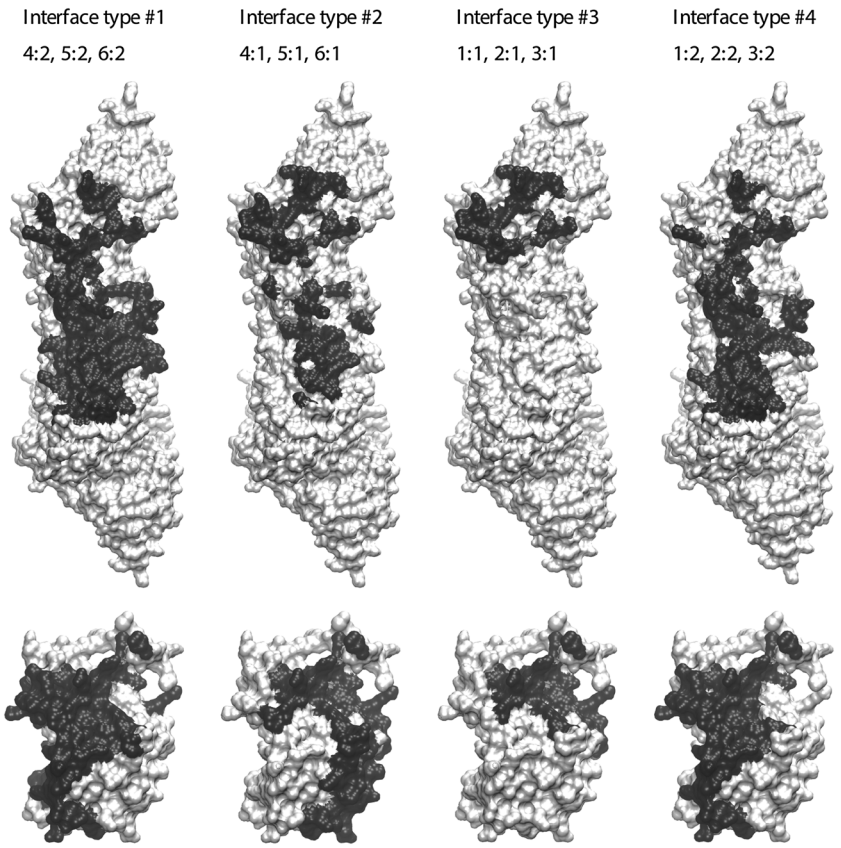

Figure 5. Four distinct types of interfaces revealed by the clustering analysis. For each type the residues, which form the contacts, are shown in black in the dimer of N-modules (top) and the C-module (bottom). C-modules are shown in different scale for clarity.

Journal of Molecular Recognition

contacts. Figure 5 shows all residues, which are shared in each interface type for N-modules and C-modules. Figures 6 and 7 show the comparison of P_c values obtained in these MD simulations and our previous coarse-grained HIEROT simulations (Yesylevskyy *et al*., 2011). The data of HIEROT simulations were taken from Yesylevskyy *et al*. (2011) and the reader if referred to this paper for detailed description of HIEROT technique and the simulation setup.

The hydrogen bonds between the ELR motif (residues E91-R93 of N-module) and the linker and C-module (residues D343 to S528) were computed for all trajectories. The ELR motif formed hydrogen bonds in three out of six MD trajectories. The hydrogen bonds between the residues R93 and Q476 of the second monomer were present for 32% of the simulation time in trajectory 1. In the trajectory 4, the hydrogen bonds were formed between the residues R93 and E473 of the second monomer for 34% of time. In the trajectory 6, the hydrogen bonds were present for up to 46% of time (27% in the first monomer and 19% in the second one) between the residues R93 and Q476, E479. Evolution of the number of hydrogen bonds in three selected trajectories is shown in Figure 8, whereas the typical hydrogen-bonded structure of the ELR motif is visualized in Figure 9.

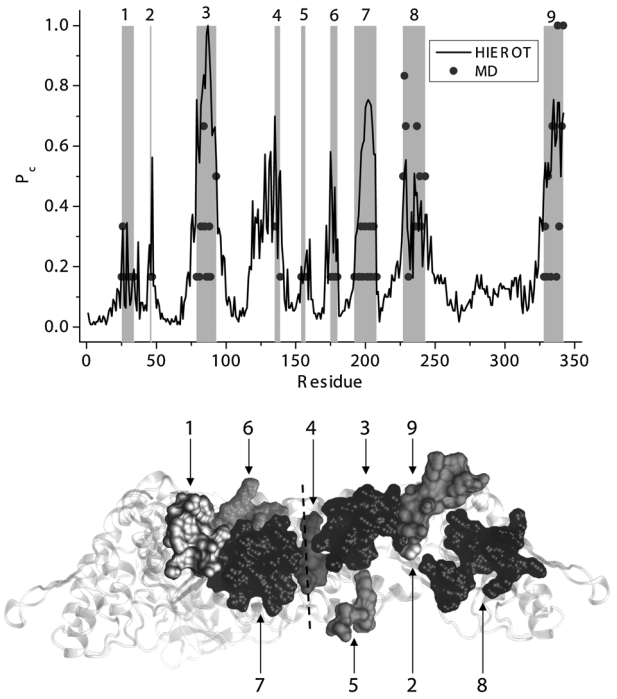

Figure 6. Probabilities of contacts formation P_c obtained by MD simulations in N-module (top) and by the HIEROT technique (Yesylevskyy *et al*., 2011). Continuous regions of preferred binding are shaded and marked by numbers. Bottom panel shows positions of the binding regions on the surface of the dimer of N-modules. Dashed line shows approximate position of the interface of the monomer interfaces.

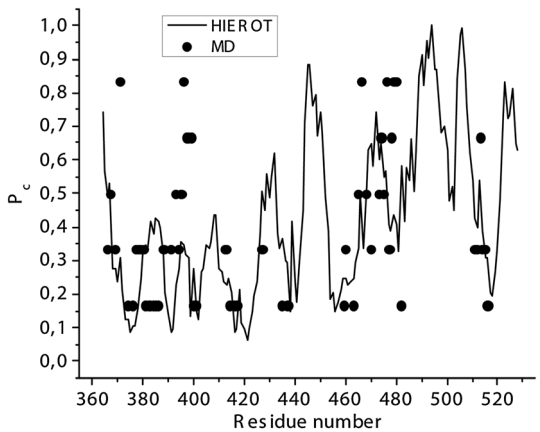

Figure 7. Probabilities of contacts formation P_c obtained by MD simulations in C-module and by the HIEROT technique (Yesylevskyy *et al*., 2011).

DISCUSSION

The crystal structure of the whole *Hs*TyrRS is still unknown, which makes MD simulation the method of choice for the study of equilibrium structure and dynamics of this enzyme. Obtaining functional compact structure of *Hs*TyrRS is a challenging task because of the large conformational freedom of C-modules connected by very flexible linkers to the catalytic dimer of mini-TyrRS. Thus, it is expected that different MD trajectories of *Hs*TyrRS will sample different regions of the conformational space and will lead to different compact structures of the enzyme. Visual inspection of the final structures (Figure 2) shows that C-modules bind in various orientations and positions, which are different in all simulations. Both C-terminal and N-terminal modules do not undergo significant conformational changes during simulations except the motions of flexible loops (data not shown), thus the dominant motion in the system was the diffusion of the C-modules and their binding to the N-modules dimer.

The RMSD values and their fluctuations are different in different MD trajectories. Trajectories 4 and 6 show very small fluctuations of RMSD, whereas the absolute values of RMSD for these trajectories are the largest and reaches 1.7 and 1.5 nm, respectively. In contrast, other four trajectories exhibit very large fluctuations of RMSD around mean values ranging from 0.8 to 1.3 nm. This means that equilibrated structures in simulations 4 and 6 are rather rigid, whereas the structures from all other simulations remain flexible.

The usage of the tRMSF method allows detecting changes of mobility of individual residues during simulations. Small reversible

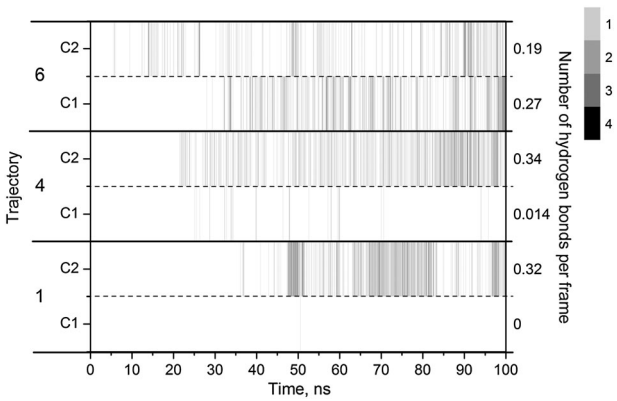

Figure 8. Evolution of the hydrogen bonds between both ELR motifs (residues 91–93 in each monomers) and the surfaces of individual C-modules and corresponding interdomain linkers (residues 343 to 528 in each monomer). The number of trajectory and the numbers of C-modules are shown at left. Mean number of the hydrogen bonds per frame is shown at right. Color code corresponds to the number of hydrogen bonds observed in each trajectory frame.

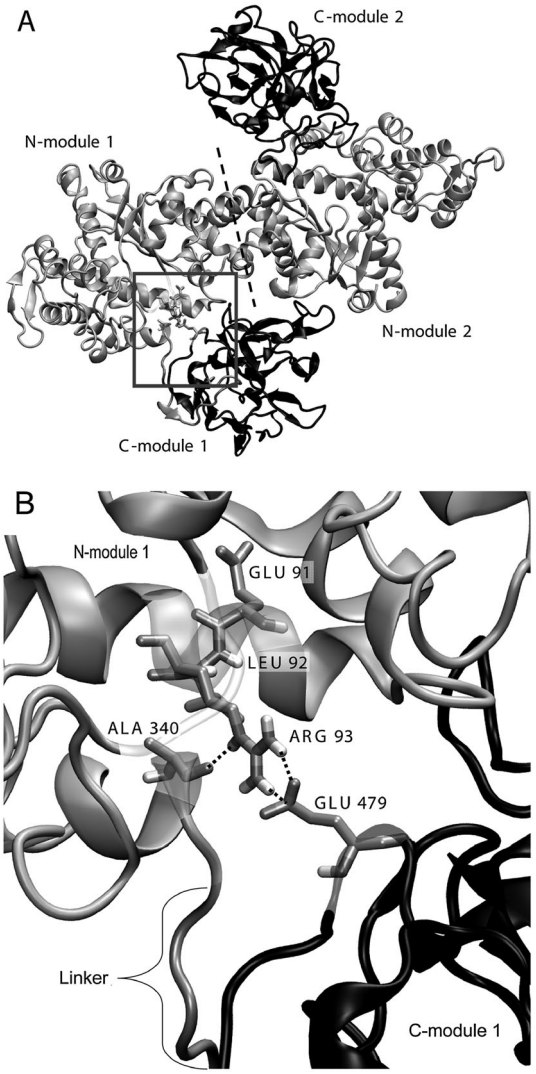

Figure 9. Typical configuration of the hydrogen bonds between the ELR motif of N-module and the C-module (trajectory 6, simulation time 96 ns). (A) Whole structure of *Hs*TyrRS dimer. Rectangle shows the region enlarged in B. Dashed line shows an approximate interface between two N-modules. (B) Close-up view of the ELR motif. N-modules are gray, C-modules are black. Residues of the ELR motif and the residues involved in the hydrogen bonding are shown as sticks. Hydrogen bonds are shown as dashed lines in B.

fluctuations are clearly visible for all residues but no systematic changes in RMSF are observed. In contrast, trajectories 1, 4, and 6 show pronounced decrease of mobility of the first C-module at times ~22, ~24, and ~27 ns, respectively (indicated by the arrows in Figure 3). This decrease is abrupt and occurs during single 200 ps window of tRMSF calculation. Comparison of Figures 2 and 3 shows that C_α RMSD grow for 15–20 ns after this sudden decrease of mobility, thus it does not marks the overall stabilization of the protein structure. However, it is plausible that very small amplitude of RMSD fluctuations in trajectories 4 and 6 correlates with the decrease of mobility of the first C-module visible in Figure 2. In trajectories 1 and 6, only one of the C-modules binds strongly, whereas in trajectory 4, both C-modules exhibit almost identical binding energies of ~400–500 kJ/mol. These data suggest that the domain interactions in MD simulations 2, 3, and 5 are rather unspecific, whereas trajectories 1, 4, and 6 lead to strong specific binding between the C-modules and the dimer of N-modules.

It is remarkable that sudden decrease of mobility of the first C-module in trajectories 1, 4, and 6, visible in Figure 3, correlates well with pronounced increase of interaction energies in these trajectories (indicated by the arrows in Figure 4). Both events occur at times ~22, ~24 and ~27 ns for trajectories 1, 4, and 6, respectively. These data reveal that strong binding of the first C-module in these simulations dramatically decreases its mobility, whereas all binding events observed in other trajectories does not lead to such effect. It can be suggested that such asymmetric change of mobility reflects functional asymmetry of the native *Hs*TyrRS monomers, but additional simulations with the ligands and bound tRNA molecules should be performed to clarify this question.

The interdomain interfaces obtained in our simulations overlap significantly, but exact pairing of contacting residues in N-modules and C-modules remains different (data not shown). The interface types differ mainly by the set of contacting residues of C-module, whereas the contacting residues of N-module are within the well-defined "binding spot" on the same side of the dimer of N-modules. Obtained domain interfaces are directly comparable with the interfaces computed using HIEROT coarse-grained technique in our previous paper (Yesylevskyy *et al*., 2011). HIEROT computations are very "cheap" in comparison to MD; therefore, the data of 400 independent HIEROT trajectories from Yesylevskyy *et al*. (2011) are averaged in Figures 5 and 6, which leads to detailed continuous and detailed curve. In contrast, there are only six MD simulations, thus the P_c values are discrete in the case of MD.

There is a remarkable correlation between the P_c values obtained in both techniques for the N-module (Figure 6). All major peaks revealed in HIEROT simulations are also present in MD simulations, and none of MD points falls into the minima of the HIEROT curve. There are nine well-defined regions in each N-module, which are continuous in sequence and possess high P_c values (marked by shaded bars in Figure 6). Lower panel of Figure 6 shows these regions on the surface of the dimer of N-modules. The regions located on the opposite side of the globule in the first monomer are located in the shown side of the second monomer, thus the continuous binding surface is formed. This surface includes all three "binding hot spots" revealed by HIEROT simulations in Yesylevskyy *et al*. (2011). The first hot spot (Tyr79–Leu89) corresponds to region 3, the second hot spot (Pro200–Tyr204) corresponds to region 7, and the third hot spot (Lys335, Ser338, Ala339) corresponds to region 9. In contrast to N-module, there is no correlation between P_c values obtained by HIEROT and MD for the C-module (Figure 7).

Our data suggest that the *Hs*TyrRS molecules in solution coexist in a number of compact asymmetric conformations, which differ significantly by their general rigidity, mobility of C-modules, and the strength of their binding to the dimer of N-modules. The orientation of bound C-modules is rather unspecific while there is a pronounced set of binding hot spots on the surface of N-modules.

Recently, Schimmel and co-workers found that in the human mini-TyrRS, the cytokine ELR motif is masked by the terminal α-helix 14, which is tethered by a hydrogen bond between the hydroxyl of the Y341 side chain and the main chain carbonyl of G46 (Yang *et al*., 2007). The molecular shapes of the wild type TyrRS and Y341A TyrRS mutant in solution were characterized by the small angle X-ray scattering profiles recently (Lee *et al*., 2012). It was proposed that the ELR area is masked by the C-module in

the wild type TyrRS, but is exposed in Y341A mutant. It was suggested that the full length TyrRS is inactive as a cytokine because the interdomain linker and the C-terminal domain occlude the ELR motif (Lee *et al*., 2012).

In view of these findings, it was interesting to analyze the hydrogen bonding of ELR motif with the linker and C-module residues (D343 to S528) in our MD simulations. The analysis confirmed the presence of strong hydrogen bonding in three MD trajectories, which persists for 32–46% of the simulation time. Figure 9 shows typical position of the ELR-motif and its putative interactions with other residues in the compact structure (obtained in trajectory 6 after 96 ns of simulation). It is clearly seen that the ELR motif is confined between the C-module and the linker, and it is immobilized by three hydrogen bonds. Thus, our MD simulations support the idea of Schimmel and co-workers that the full length TyrRS lacks its cytokine activity because of the interactions with the interdomain linker and the C-terminal domain, which protects the ELR cytokine motif (Lee *et al*., 2012).

Although MD simulations in this work are performed in the absence of the ligands, they provide a solid basis for comparison of the compact states and interdomain interfaces of *Hs*TyrRS, which will be obtained in the future for *Hs*TyrRS complexes with various ligands.

CONCLUSION

MD simulations of the full length *Hs*TyrRS have shown that the binding of C-modules to mini-TyrRS catalytic dimer is highly asymmetric in terms of the interface structure and dynamics of the formed complex. Strong binding of one of C-modules leads to abrupt decrease of its internal mobility, whereas the second C-module remains relatively flexible. Three well-defined binding hot spots on the surface of mini-TyrRS (Tyr79-Leu89, Pro200-Tyr204; Lys335, Ser338; Ala339) were revealed for N-module. These regions of preferable interdomain contacts are in perfect agreement with the coarse-grained HIEROT simulations reported earlier (Yesylevskyy *et al*., 2011). Strong hydrogen bonding between the residue R93 of the ELR motif and the residues A340 and E479 in C-module was observed in our MD simulations. This supports the idea that the full length TyrRS lacks its cytokine activity because of the interactions between N-terminal and the C-terminal modules, which protect the ELR cytokine motif.

Acknowledgements

Ievgen Sliusar, Andrii Salnikov, Nikolay Pydiura, Vladimir Kazakov, Dmitriy Sukhomlynov, Dr. Hugo Verli, Prof. Valery N. Kharkyanen, and Prof. Alexander P. Demchenko are acknowledged for their technical support and useful discussions. Computational resources were provided by Ukrainian National Grid infrastructure (IMBG, ISMA, KPI, MAO, KNU, ICYB SCIT-3 clusters). This work was supported by the project 15/2010–2012 of the State Scientific and Technical Program of Implementation and Application of Grid Technologies in Ukraine. Development of the Pteros library was partially supported by the STCU grant 5525 and the project 15/2010-2012. O. S. was supported by the Federation of European Biochemical Societies for Youth Travel Funds (Y/10/35, Y/11/38).

REFERENCES

Alexander RW, Schimmel P. 2001. Domain-domain communication in aminoacyl-tRNA synthetases. *Prog. Nucleic Acid Res. Mol. Biol.* **69**: 317–349.

Bennion BJ, Daggett V. 2004. Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: a chemical chaperone at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* **101**: 6433–6438.

Berger AC. 2000. Endothelial monocyte-activating polypeptide II, a tumour-derived cytokine that plays an important role in inflammation, apoptosis, and angiogenesis. *J. Immunother.* **23**: 519–527.

Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**: 12–21.

Eswar N, Marti-Renom MA, Webb B, Madhusudhan MS, Eramian D, Shen M, Pieper U, Sali A. 2006. Comparative protein structure modeling with MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30: 200.

Feenstra KA, Hess B, Berendsen HJC. 1999. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**: 786–798.

Gnatenko DV, Kurochkin IV, Ribkinska TA, Kornelyuk AI, Matsuka GK. 1991. Purification and characterization of functionally active proteolytically modified form of tyrosyl-tRNA synthetase from bovine liver, Ukr. *Biokhim. Zhurn. (Kiev)* **63**: 61–67.

Goette M, Stumpe MC, Ficner R, Grubmüller H. 2009. Molecular determinants of snurportin 1 ligand affinity and structural response upon binding. *Biophys. J.* **97**: 581–589.

Guo M, Schimmel P, Yang XL. 2010. Functional expansion of human tRNA synthetases achieved by structural inventions. *FEBS Lett.* **584**: 434–442.

Guo M, Yang XL, Schimmel P. 2010. New functions of aminoacyl-tRNA synthetases beyond translation. *Nat. Rev. Mol. Cell Biol.* **11**: 668–674.

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. 1997. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**: 1463–1472.

Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theor. Comp* **4**: 435–447.

Humphrey W, Dalke A, Schulten K. 1996. VMD – Visual Molecular Dynamics. *J. Molec. Graphics* **14**: 33–38.

Kleeman TA, Wei D, Simpson KL, First EA. 1997. Human tyrosyl-tRNA synthetase shares amino acid sequence homology with a putative cytokine. *J. Biol. Chem.* **272**: 14420–14425.

Kornelyuk AI, Kurochkin IV, Matsuka GK. 1988. Tyrosyl-tRNA synthetase from the bovine liver. Isolation and physico-chemical properties. *Molekulyarnaya Biologiya (Moscow)* **22**: 176–186.

Kornelyuk AI, Tas MPR, Dybrovsky A, Murray C. 1999. Cytokine activity of the non-catalytic EMAP-2-like domain of mammalian tyrosyl-tRNA synthetase. *Biopolym. Cell (Kiev)* **15**: 168–172.

Lee PS, Zhang HM, Marshall AG, Yang XL, Schimmel P, Schimmel P. 2012. Uncovering of a short internal peptide activates a tRNA synthetase procytokine. *J. Biol. Chem.* **287**: 20504–20508.

Levanets OV, Naidenov VG, Odynets KA, Woodmaska MI, Matsuka GK, Kornelyuk AI. 1997. Homology of C-terminal non-catalytic domain of mammalian tyrosyl-tRNA synthetase with cytokine EMAP II and non-catalytic domains of methionyl- and phenylalanyl-tRNA synthetases. *Biopolym. Cell (Kiev)* **13**: 474–478.

Marti-Renom MA, Stuart A, Fiser A, Sánchez R, Melo F, Sali A. 2000. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**: 291–325.

Mirande M. 1991. Aminoacyl-tRNA synthetase family from prokaryotes and eukaryotes: structural domains and their implications. *Prog. Nucleic Acid Res. Mol. Biol.* **40**: 95–142.

Miyamoto S, Kollman PA. 1992. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **13**: 952–962.

Reznikov AG, Chaykovskaya LV, Polyakova LI, Kornelyuk AI. 2007. Antitumor effect of endothelial monocyte-activating polypeptide-II on human prostate adenocarcinoma in mouse xenograft model. *Exper. Oncol* **29**: 267–271.

Salnikov AO, Sliusar IA, Sudakov OO, Savytskyi OV, Kornelyuk AI. 2009. MolDynGrid virtual laboratory as a part of Ukrainian Academic Grid infrastructure. Proceedings of the 5th IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. IDAACS 2009: 237–240.

Salnikov A, Sliusar I, Sudakov O, Savytskyi O, Kornelyuk A. 2010. Virtual laboratory MolDynGrid as a part of scientific infrastructure for biomolecular simulations. *Intl. J. Comp.* **9**: 294–300.

Savytskyi OV, Sliusar IA, Yesylevskyy SO, Stirenko SG, Kornelyuk AI. 2011. Integrated tools for molecular dynamics simulation data analysis in the MolDynGrid virtual laboratory. Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2011. 1: 209–211.

Shen MY, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**: 2507–2524.

Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. 2005. GROMACS: fast, flexible and free. *J. Comp. Chem.* **26**: 1701–1718.

Verli H, Guimarães JA. 2005. Insights into the induced fit mechanism in antithrombin-heparin interaction using molecular dynamics simulations. *J. Mol. Graph. Model.* **24**: 203–212.

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

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

Weimer KM, Shane BL, Brunetto M, Bhattacharyya S, Hati S. 2009. Evolutionary basis for the coupled-domain motions in *Thermus thermophilus* leucyl-tRNA synthetase. *J. Biol. Chem.* **284**: 10088–10099.

Woese CR, Olsen GJ, Ibba M, Söll D. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* **64**: 202–236.

Yang X-L, Skene RJ, McRee DE, Schimmel P. 2002. Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl. Acad. Sci. U. S.A.* **99**: 15369–15374.

Yang X-L, Liu J, Skene RJ, McRee DE, Schimmel P. 2003. Crystal structure of an EMAP-II-like cytokine released from a human tRNA synthetase. *Helv. Chim. Acta* **86**: 1246–1257.

Yang XL, Kapoor M, Otero FJ, Slike BM, Tsuruta H, Frausto R, Bates A, Ewalt KL, Cheresh DA, Schimmel P. 2007. Gain-of-function mutational activation of human tRNA synthetase procytokine. *Chem. Biol.* **14**: 1323–1333.

Yesylevskyy SO. 2012. Pteros: fast and easy to use open-source C++ library for molecular analysis. *J. Comput. Chem.* **33**: 1632–1636.

Yesylevskyy SO, Savytskyi OV, Odynets KA, Kornelyuk AI. 2011. Interdomain compactization in human tyrosyl-tRNA synthetase studied by the hierarchical rotations technique. *Biophys. Chem.* **154**: 90–98.

Zhang CM, Hou YM. 2005. Domain-domain communication for tRNA aminoacylation: the importance of covalent connectivity. *Biochemistry* **44**: 7240–7249.