Eur Biophys J (2014) 43:613–622
DOI 10.1007/s00249-014-0991-8

ORIGINAL PAPER

EBSA
Biophysics in Europe

Structural states of the flexible catalytic loop of *M. tuberculosis* tyrosyl-tRNA synthetase in different enzyme–substrate complexes

Vasyl V. Mykuliak · Anatoliy I. Dragan · Alexander I. Kornelyuk

Received: 8 April 2014 / Revised: 17 September 2014 / Accepted: 28 September 2014 / Published online: 6 November 2014

Abstract Tyrosyl-tRNA synthetase from *Mycobacterium tuberculosis* (*Mt*TyrRS) is an enzyme that belongs to class I of aminoacyl-tRNA synthetases, which catalyze the attachment of L-tyrosine to its cognate tRNATyr in the preribosomal step of protein synthesis. *Mt*TyrRS is incapable of cross-recognition and aminoacylation of human cytoplasmic tRNATyr, so this enzyme may be a promising target for development of novel selective inhibitors as putative antituberculosis drugs. As a class I aminoacyl-tRNA synthetase, *Mt*TyrRS contains the HIGH-like and KFGKS catalytic motifs that catalyze amino acid activation with ATP. In this study, the conformational mobility of *Mt*TyrRS catalytic KFGKS loop was analyzed by 100-ns all-atoms molecular dynamics simulations of the free enzyme and its complexes with different substrates: tyrosine, ATP, and the tyrosyl–adenylate intermediate. It was shown that in the closed state of the active site, the KFGKS loop, readily adopts different stable conformations depending on the type of bound substrate. Molecular dynamics simulations revealed that the closed state of the loop is stabilized by dynamic formation of two antiparallel β-sheets at flanking ends which hold the KFGKS fragment inside the active center. Prevention of β-sheet formation by introducing point mutations in the loop sequence results in a rapid (<20 ns) transition of the loop from its functional "closed" M-like structure to an inactive "open" O-like structure, i.e. rapid diffusion of the catalytic loop outside the active site. The flexibility and rapid dynamics of the wild-type aaRS catalytic loop structure are crucial for formation of protein–substrate interactions and subsequently for overall enzyme functional activity.

Keywords Tyrosyl-tRNA synthetase · *Mycobacterium tuberculosis* · Active site · Molecular dynamics · β-strands · Grid technology

Introduction

Tyrosyl-tRNA synthetase (TyrRS) is a class I enzyme of the aminoacyl-tRNA synthetases (aaRS) which catalyzes the attachment of tyrosine to the 3′ end of the cognate tRNATyr in the preribosomal step of protein synthesis. The aminoacylation reaction consists of two steps. In the first step, the L-tyrosine is activated by ATP, forming the enzyme-bound tyrosyl–adenylate intermediate. In the second step of the reaction, the activated L-tyrosine is transferred to tRNATyr to form the tyrosyl–tRNATyr complex (Bedouelle 1990; Abergel et al. 2007; Kornelyuk 1998; Bonnefond et al. 2005).

In general, bacterial TyrRSes are regarded as promising targets for developing of new types of drugs (Qiu et al. 2001), including anti-tuberculosis drugs. In particular tyrosyl-tRNA synthetase from *Mycobacterium tuberculosis* (*Mt*TyrRS) could serve as an antimicrobial target of novel enzyme inhibitors (Hoffmann and Torchala 2009; Eitner et al. 2007). *Mt*TyrRS is one of the most vital proteins of bacterial cells, and its inhibition could effectively suppress the growth of pathogenic

Electronic supplementary material The online version of this article (doi:10.1007/s00249-014-0991-8) contains supplementary material, which is available to authorized users.

V. V. Mykuliak · A. I. Dragan · A. I. Kornelyuk
Institute of High Technologies, Taras Shevchenko National University of Kyiv, 64, Volodymyrs'ka St., Kiev 01601, Ukraine

V. V. Mykuliak (✉) · A. I. Dragan · A. I. Kornelyuk
Department of Protein Engineering and Bioinformatics, Institute of Molecular Biology and Genetics, NASU, 150, Akademika Zabolotnoho St., Kiev 03680, Ukraine
e-mail: v.mykuliak@imbg.org.ua

bacteria. Remarkably, the structure of *Mt*TyrRS significantly diverges from human cytoplasmic TyrRS. The identity of amino acid sequences between these two enzymes is <20 % (Odynets and Kornelyuk 2008). Whereas the mitochondrial enzyme is more similar to eubacterial TyrRSs (Bonnefond et al. 2005), the identity of amino acid sequences between mitochondrial TyrRS and *Mt*TyrRS is only approximately 42 % (Odynets and Kornelyuk 2008). Also, *Mt*TyrRS has another type of protein spatial organization, significant differences in the active center structure, and is incapable of cross-recognition and aminoacylation of mammalian tRNATyr (Bedouelle 1990; Odynets and Kornelyuk 2008).

TyrRS is a Class I aaRS which contains a Rossman-fold catalytic domain. The active site contains two highly conserved sequence motifs—HIGH and KMSKS (Odynets and Kornelyuk 2008; First and Fersht 1993; Leatherbarrow et al. 1985). It is believed that the catalytic KMSKS-loop is in a disordered state connecting the catalytic and anticodon binding domains. The *Mt*TyrRS enzyme also has the KMSKS-like motif, specifically the sequence KFGKS (Lys231–Phe232–Gly233–Lys234–Ser235), in which the K231 and K234 residues are involved in binding and stabilization of ATP in the active center. The positively charged groups of these Lys residues effectively interact with negatively charged phosphate groups of ATP, thereby stabilizing the complex (Kamijo et al. 2008). Also K231 and K234 of the KFGKS catalytic loop are important for initial binding of tRNATyr to TyrRS (Xin et al. 2000). The eukaryotic KMSKS sequence participates in catalysis of tyrosyl–adenylate formation by interacting with the pyrophosphate moiety of ATP (Xin et al. 2000; Austin and First 2002). The K231 and K234 residues of the bacterial tyrosyl-tRNA synthetase behave in the same way, i.e. interact with the pyrophosphate moiety of the ATP substrate, resulting in stabilization of the transition state responsible for tyrosine activation (Austin and First 2002). It is believed that inside the active center of the TyrRSes the catalytic KMSKS loop is in the so-called "open" state or adopts specific conformations (First and Fersht 1995), "semi-open" and "closed", depending on the type of bound substrate (Kobayashi et al. 2005a; Li et al. 2008; Palencia et al. 2012).

Bacterial TyrRSes have been studied for many years, and the Protein Data Bank (PDB) contains the crystal structures of the TyrRS from *E. coli* (1VBM, 1VBN, 1WQ3, 1WQ4, 1X8X) (Kobayashi et al. 2005a, b), *T. thermophilus* (1H3E, 1H3F) (Yaremchuk et al. 2002), *S. aureus* (1JII, 1JIJ, 1JIK, 1JIL) (Qiu et al. 2001), *B. stearothermophilus* (1TYD, 2TS1, 3TS1, 4TS1) (Brick et al. 1989), and *M. tuberculosis* (2JAN) (Hartmann et al., unpublished).

Recently, MD simulations have been widely used for investigation of aminoacyl-tRNA synthetases and the flexibility of their active sites; examples include tyrosyl-tRNA synthetase (Li et al. 2008; Yesylevskyy et al. 2011; Savytskyi et al. 2013), methionyl-tRNA synthetase (Budiman et al. 2007), aspartyl-tRNA synthetase (Thompson et al. 2006; Thompson and Simonson 2006), lysyl-tRNA synthetase (Hughes et al. 2006), and tryptophanyl-tRNA synthetase (Kapustina and Carter 2006).

In this work we investigated the dynamic properties of tyrosyl-tRNA synthetase from *M. tuberculosis* (*Mt*TyrRS), with emphasis on the structural dynamics of the catalytic loop (KFGKS) in the substrate-free enzyme and in complexes with substrates—tyrosine, ATP, and the tyrosyl–adenylate intermediate—by using MD simulations over a long (100 ns) time interval by means of modern computational grid techniques (Salnikov et al. 2010; Savytskyi et al. 2011).

Special attention was devoted to the flexibility of the catalytic loop in the closed state, i.e. inside the *Mt*TyrRS active center cavity, and its capacity for dynamic formation of two antiparallel β-sheets in the KMSKS loop flanking regions. The β-hairpins could have important structure–functional activity in holding the catalytic loop in the closed M-like conformation, thus ensuring it functions within the active site cavity.

Materials and methods

Preparation of an initial unligated protein structure (all-atom coordinates) and protein in complexes with cognate substrates for MD simulations

The crystalline *Mt*TyrRS dimer structure was used as initial structure (PDB code 2JAN) (Hartmann et al., unpublished). The *Mt*TyrRS monomer contains three domains: the catalytic domain (1–248), the anticodon-binding domain (249–341), and the C-terminal domain (342–424) (Fig. 1). The *Mt*TyrRS KFGKS sequence (231–235) corresponds to the

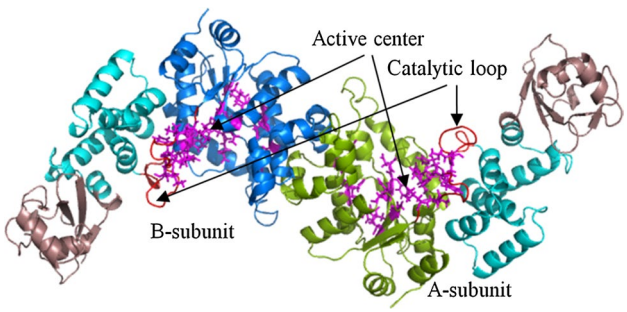

Fig. 1 Three-dimensional structure of the *Mt*TyrRS dimer. The catalytic domains are in *green* (A subunit) and in *blue* (B subunit) and the anticodon binding domains are in *cyan*. The C-terminal domains are in *brown*, the catalytic loops are in *red*, and the active site residues are in *magenta*

Table 1 Initial structures of the *Mf*TyrRS protein for MD simulations

Simulating structures	Initial PDB files	Constructed structures
Uncomplexed *Mf*TyrRS	2JAN	Constructed by completing 2JAN
*Mf*TyrRS + Tyr	1X8X	Constructed by superposition of 2JAN and 1X8X, keeping the protein of 2JAN and the Tyr ligand of 1X8X
*Mf*TyrRS + Tyr + ATP	1H3E	Constructed by superposition of 2JAN, 1X8X, and 1H3E, keeping the protein of 2JAN, the Tyr ligand of 1X8X, and the ATP ligand of 1H3E
*Mf*TyrRS + tyrosyl–adenylate	1VBM	Constructed by superposition of 2JAN and 1VBM, keeping the protein of 2JAN and the Tyr-AMS ligand of 1X8X, replacing atom S to P, to obtain the tyrosyl–adenylate but not its analog

conservative catalytic loop KMSKS motif inherent to Class I aaRS (Kobayashi et al. 2005a; Yaremchuk et al. 2002).

The substrate-free *Mf*TyrRS dimer structure coordinates were obtained by adding missing residues to the crystal structure (1–4 and 424 at the A subunit and 1–4 and 81–93 at the B subunit) by use of the ModLoop server (Fiser et al. 2000; Fiser and Sali 2003). The coordinates of ligands (substrates) in the active center were taken from crystal structures of bacterial TyrRSs (Bonnefond et al. 2005; Kobayashi et al. 2005a; Retailleau et al. 2001). To construct the enzyme structure we superimposed the atomic coordinates of the protein—2JAN (Hartmann et al., unpublished) and ligand (Tyr)—1X8X (*E. coli* TyrRS) (Kobayashi et al. 2005a), keeping the protein structure and Tyr invariable. The same strategy was used to generate the complex of *Mf*TyrRS with ATP (1H3E—*T. thermophilus* TyrRS) (Yaremchuk et al. 2002), and with the tyrosyl–adenylate intermediate (1VBM—*E. coli* TyrRS) (Kobayashi et al. 2005a). In latter case we replaced atom S by P, to obtain the tyrosyl–adenylate but not its analog (Table 1).

To study the effect of the β-sheets on the catalytic loop dynamics and structure–functional state we designed a variant of the *Mf*TyrRS loop with proline substitutions (proline residues effectively break the β-sheets) at positions 225, 226, 230, 231, 234, 235, and 240, i.e. at sites involved in β-hairpin formation (Fig. 6).

Molecular dynamics simulations

All-atom MD simulations were performed by use of the GROMACS 4.5 package (Hess et al. 2008; Van Der Spoel et al. 2005; Berendsen et al. 1995). Each system was simulated with 100-ns trajectories at least three times with the Amber ff99SB-ILDN force field (Hornak et al. 2006) and, separately, three times with the CHARMM27 force field (Bjelkmar et al. 2010). The protein structure of substrate-free mutant *Mf*TyrRS was calculated by use of the Amber ff99SB-ILDN force field only. The ligand topologies for the Amber ff99SB-ILDN force field were prepared by use of acpype (AnteChamber PYthon Parser interfacE) scripts (Wang et al. 2004, 2006), in the antechamber suite of AmberTools12 (Case et al. 2005). The ligand topologies for the CHARMM27 force field were prepared by using the SwissParam web-service (Zoete et al. 2011). The protein was placed in a triclinic water box with the minimum distance between the *Mf*TyrRS and the box wall of 1 nm. Explicit TIP3P water molecules (Jorgensen and Madura 1983) were used. All simulations were performed under periodic boundary conditions. Na^+ and Cl^- counterions were added to completely neutralize the system at 150 mM NaCl salt concentration. Each system was energy-minimized then equilibrated with positioning restraints on heavy atoms of the protein before the simulations were initiated. The leap-frog integration algorithm was used, together with a 2 fs timestep. All bond lengths were constrained using the LINCS algorithm (Hess et al. 1997). Unless otherwise stated, long-range electrostatic interactions were computed by use of the fourth-order particle mesh Ewald (PME) (Essmann et al. 1995) method with a Fourier spacing of 0.16 nm. The real space coulombic interactions and the pair-list calculations were set to 1.0 nm. A twin-range cutoff of 1 nm was used for the Van der Waals interactions. The temperature and pressure were maintained by coupling temperature and pressure baths, by using the V-rescale (Bussi et al. 2007) and Parrinello-Rahman methods (Parrinello and Rahman 1981) with relaxation times of 0.1 and 0.5 ps, respectively. A temperature of 310 K and pressure of 1 atm were used. All MD simulations were calculated using the services of the MolDynGrid virtual laboratory (http://moldyngrid.org) at the ICYB and ISMA clusters of the Ukrainian National Grid infrastructure (Salnikov et al. 2010; Savytskyi et al. 2011).

Graphical and structural analysis

PyMOL (2011) and VMD (Humphrey et al. 1996) software were used for trajectory visualization and graphical structure analysis. The secondary structure of the enzyme was analyzed by use of the GROMACS analysis tool do_dssp based on the Dictionary of Secondary Structure of Protein (DSSP) algorithm (Joosten et al. 2011). The root mean square deviations (RMSD) and root mean square fluctuations (RMSF) were calculated by use of the g_rms and g_rmsf routines,

respectively, of GROMACS software. It should be noted that the difference between RMSD and RMSF is that with the latter the average is taken over time, giving a value for each Cα atom of a protein, whereas with RMSD the average is taken over all the Cα atoms, giving time-specific values. Hydrogen bonds were monitored with the g_hbond routine. Mutations in the protein sequence were introduced by use of the PyMOL (2011) Mutagenesis Wizard application.

Results and discussion

Analysis of global protein structure stability during MD simulations

To verify the global structural stability of the protein during MD simulations we analyzed the Cα atoms RMSD which were calculated as the deviation of the current atomic coordinates from the initial coordinates on application of the force field. Calculated RMSD values demonstrate global stability of the protein structure in the course of MD simulations (0–100 ns) and also show insignificant difference between Cα RMSD data calculated for the two force fields used, Amber ff99SB-ILDN and CHARMM27, i.e. the *Mt*TyrRS dimer is quite stable in both applied fields (Fig. 2a).

The MD simulation trajectories obtained can be divided into two parts: deviations calculated during the first ~10 ns of simulations (relaxation step) and subsequent productive MD. For both parts the calculated global RMSD values are in the range 2.2–6.1 Å reflecting the protein global structure stability.

Interestingly, the calculated RMSF values for Cα atoms are not uniform. The amplitude of fluctuations (RMSF)

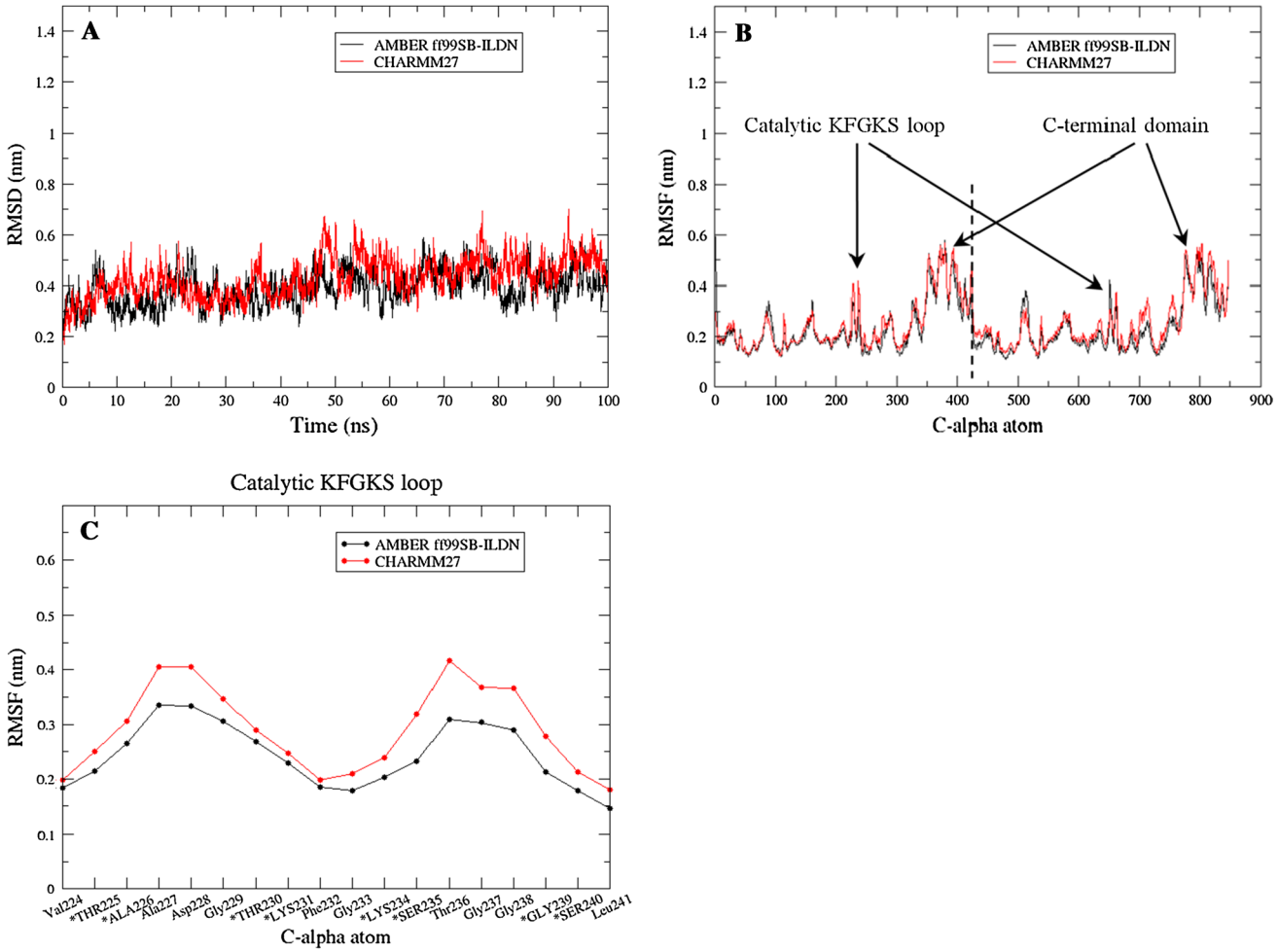

Fig. 2 RMSD (**a**) and RMSF (**b**, **c**) of the Cα atoms from the starting structure of the uncomplexed *Mt*TyrRS during MD simulations with the Amber ff99SB-ILDN (*black*) and CHARMM27 (*red*) force fields. The RMSDs can be divided into two parts: deviations calculated during the first ~10 ns of simulations (relaxation step) and subsequent productive MD. For both parts, the calculated global RMSDs are in the range 2.2–6.1 Å. RMSF show that the most flexible elements of the dimer are the catalytic KFGKS loops (RMSF ~4 Å) and the C-terminal domains (RMSF is 2.3–5.8 Å). Residues labeled with an *asterisk* form the β-hairpin structure

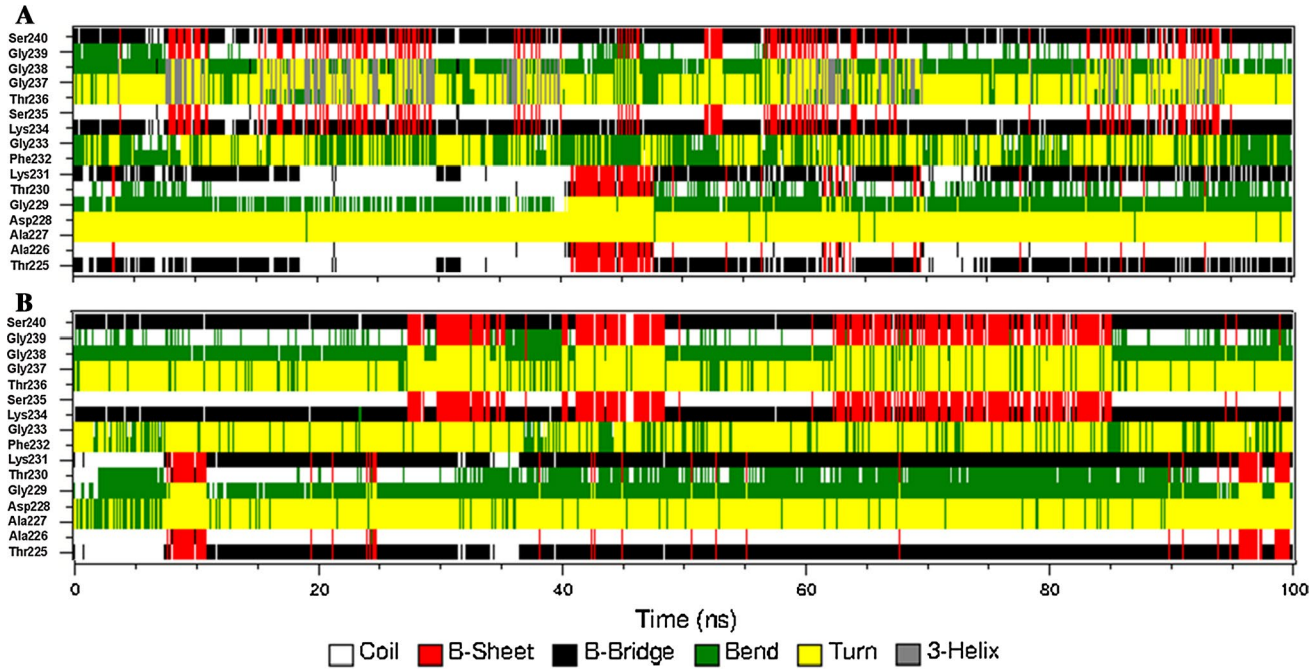

Fig. 3 A schematic representation of the secondary structure changes of the catalytic loop KFGKS (Thr225–Ser240 residues) in the A-subunit of free *Mt*TyrRS in the course of MD simulations with the Amber ff99SB-ILDN (**a**) and CHARMM27 (**b**) force fields. Antiparallel β-strands form from Thr225–Ala226 with Thr230–Lys231 and from Lys234–Ser235 with Gly239–Ser240

changes along the polypeptide chain revealing enhanced fluctuation of the C-terminal domains and catalytic KFGKS loops, i.e. these domains are the most flexible parts of the *Mt*TyrRS dimer (Fig. 2b, c). Substrate-dependent interdomain movements have been shown in tryptophanyl-tRNA synthetase (Weinreb et al. 2014). In contrast, we have shown that the C-terminal domain has high flexibility both in the substrate-free enzyme state and in complexes with different substrates. High flexibility of the C-terminal domain is in agreement with the high temperature factors for the domain in the *Mt*TyrRS crystal structure (PDB code: 2JAN). One can assume that flexibility of this domain is a general property of that class of enzymes, which explains problems in resolving its structure in other TyrRS (Kobayashi et al. 2005a, b; Qiu et al. 2001; Brick et al. 1989).

Formation of dynamic β-hairpin structures in the catalytic loop flanking regions

It is typical for aaRSs of Class I that the catalytic KMSKS-like loop (KFGKS for bacterial *Mt*TyrRS) connecting catalytic and anticodon binding domains is in the unfolded state, so the 3D structure of the uncomplexed *Mt*TyrRS dimer (PDB code: 2JAN) reveals that the KFGKS-loop has the M-like conformation but without detectable β-hairpins in the crystal structure. However, in the course of our MD simulations dynamic formation of two antiparallel β-sheets can clearly observed in the flanking regions of the catalytic KFGKS loop (Fig. 3; video-1S file in supplemental material). Video-1S shows the dynamics (flickering) of β-hairpin formation in the loops over a simulation time of 100 ns. The residence time of folded and unfolded states is in the ps–ns range, i.e. comparable with the lifetime of H-bonds between water molecules in solution (Keutsch and Saykally 2001). Interestingly, fast dynamic interactions of water molecules with DNA (the average residence time of water molecules on DNA is approx. 20 ps) resulting in stabilization of the DNA duplex in the B-form (Pal et al. 2003a, b). Consequently, one can assume that the observed dynamic, rapidly flickering β-hairpins stabilize the catalytic loop in a specific M-like conformation. This would enable rapid and efficient binding to the active center loop of other molecules that also use hydrogen bonds for interaction. The length of these β-hairpins is quite short, consisting of only four residues. The first antiparallel β-hairpin is formed by the Thr225–Ala226 and Thr230–Lys231 residues, the second contains Lys234–Ser235 and Gly239–Ser240. It should be noted that the catalytic K231, K234, and S235 residues of the loop are involved in β-hairpin formation, and the catalytic motif is positioned between the flanking β-structures in the active site. The β-hairpins in the *Mt*TyrRS catalytic loop are not a part of the protein stable secondary structures, but are unique dynamic, fast-flickering structures. The time interval of stability of the observed

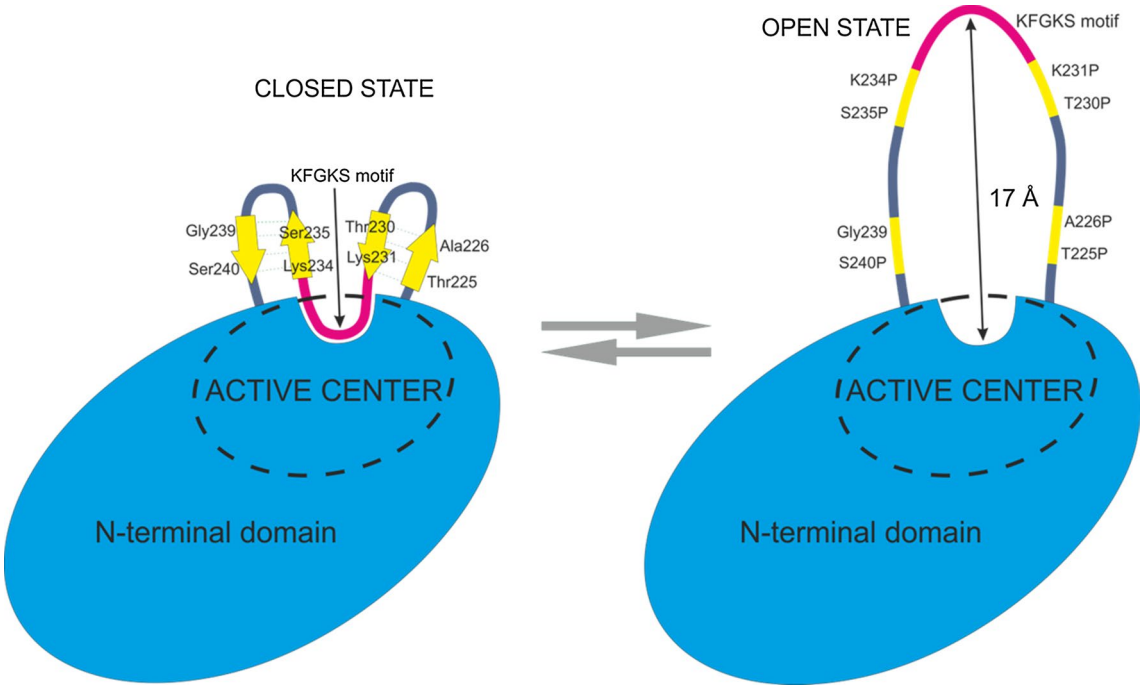

Fig. 4 Schematic representation of the catalytic KFGKS loop in two states: the "closed" M-like structure and the "open" ring-like structure (O-structure). Formation of β-strands in the loop keeps it in the "closed" M-like state directing the KFGKS motif toward the active center. In the absence of the β-strands the KFGKS motif moves out of active site, changing the loop conformation to the "open" O-structure

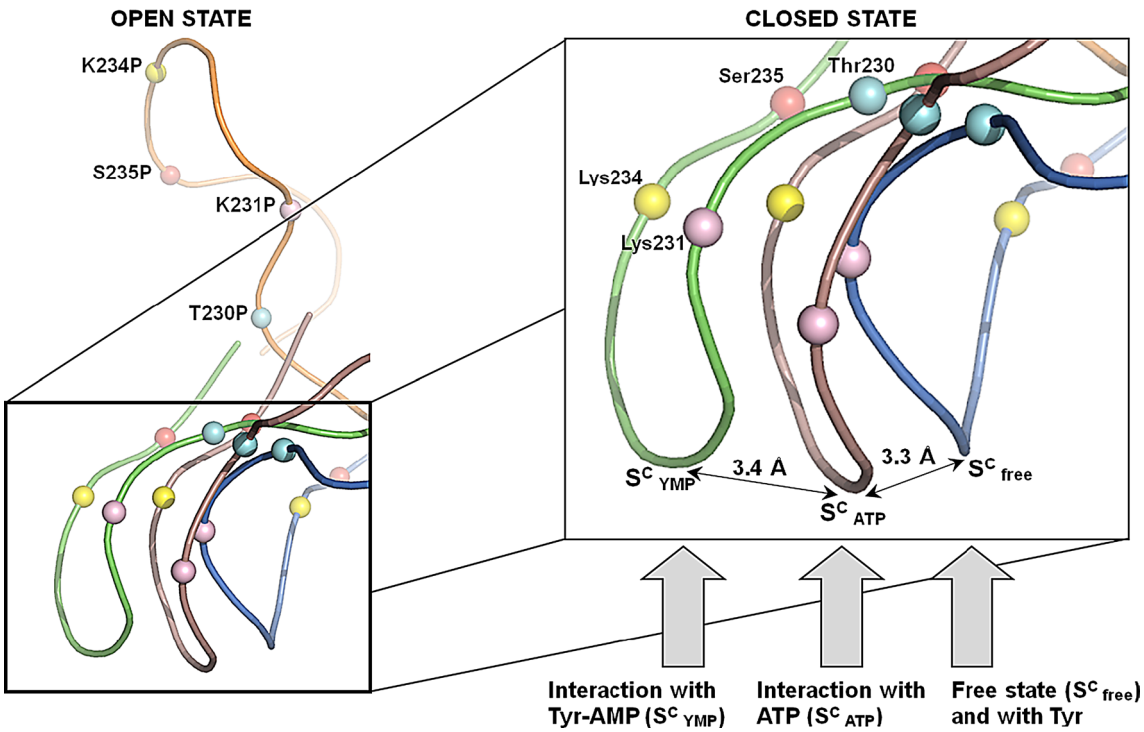

Fig. 5 The "open" O-like and "closed" M-like conformations of the catalytic loop. Formation of β-strands in catalytic loop keeps the loop in a "closed" M-shape conformation. Depending on the substrate in the active center the loop adopts different conformations: in the free state it relaxes (S^C_{free}), when interacting with ATP the loop shifts by 3.4 Å (S^C_{ATP}), and when interacting with tyrosyl–adenylate it shifts by 6.7 Å (S^C_{YMP}). The different conformations of the catalytic loop interacting with substrates are the final structures after 100-ns MD simulations. Without β-strands in the catalytic loop, it moves from the active site into its "open" O-like conformation

β-sheets is short and varies in the range from picoseconds to several nanoseconds (Fig. 3). It is interesting to note that similar short β-hairpins were also found in the catalytic loop of the crystal structure of *E. coli* TyrRS in the complex with tyrosyl–adenylate analog (Fig. 1S, supplemental material) (crystal structure PDB file: 1VBM) (Kobayashi et al. 2005a), a fact which supports the findings of our MD simulations.

For the wild type TyrRS the catalytic loop has an M-like structure, the central part of which points to the active center of the enzyme (Fig. 4). However, such an M-shape of the disordered polypeptide fragment is energetically unfavorable and without specific factors to stabilize the loop in this dynamic M-conformation it undoubtedly unfolds into a ring-like conformation. The β-hairpin structures observed in our MD simulations at sites flanking the loop may restrict the free movement (degree of freedom) of the chain and thereby play the role of such specific factors. The energy of β-hairpin formation can effectively stabilize the M-structure, directing the catalytic part of the loop toward the active site ("closed" state) and make it ready to bind the substrate.

Conformation of the *Mt*TyrRS catalytic loop in substrate-bound states

Analysis of catalytic loop conformation in substrate-bound states revealed that it is in the closed state (inside the protein active site) and the position and conformation of the central loop strongly depend on the type of substrate (Fig. 5). Being in a closed state, the flexibility of the loop enables adoption of a conformation that is specific for strong interaction with different cognate substrates. Analysis of our MD data revealed that the catalytic loop can be located in three closed states which are designated as substrate free (S^C_{free}), ATP-bound (S^C_{ATP}), and tyrosyl–adenylate-bound (S^C_{YMP}) states, respectively. The similarity of these three states is that the catalytic loop located inside the enzyme active center (closed state) and its position and conformation within the cavity are governed by interactions with the substrates. According to other results (Kobayashi et al. 2005a; Li et al. 2008), these states are entitled "open", "semi-open" and "closed", respectively. Recent work by Datt and Sharma (2014) has classified two conformational states of the KMSKS loop, compact and extended, depending on the distance between the KMSKS motif and the Rossmann fold domain. They also found that the catalytic loop has a dynamic and flexible structure. Our findings are in good agreement with the results obtained by Datt and Sharma (2014).

The conformation of the KFGKS loop in the tyrosine-bound state is similar to the substrate-free conformation (S^C_{free}). This can be explained by the binding of L-tyrosine at the opposite side of the active center cavity, and the KFGKS loop is far away from amino acid substrate, not forming direct bonds with the L-tyrosine residue. Direct interactions of K231 and K234 residues with ATP shift the loop position within the active center cavity by 3.3 Å, relative to its position in the S^C_{free} state, stabilizing the S^C_{ATP} conformation. In tyrosyl–adenylate binding the loop specifically interacts with this intermediate adopting a new conformation, S^C_{YMP}, and shifting the loop from the substrate-free conformation by 6.7 Å (Fig. 5). It should be noted that the loop forms strong interactions with the tyrosyl–adenylate intermediate. This complex contains three H-bonds between the catalytic loop and tyrosyl–adenylate: two H-bonds with V224 and one H-bond with the F232 residue (Table 2) (Fig. 2S, supplemental material).

In all considered closed states (substrate-free *Mt*TyrRS and substrate-bound states) the M-like conformation of the loop is also stabilized by the dynamic β-hairpin flanking structures.

Table 2 Hydrogen bond formation between the substrates and active center of *Mt*TyrRS during the MD simulations

Hydrogen bond		Distance (Å)
*Mt*TyrRS	Tyr	
Tyr36-OH	–OH	2.82
Asp40-OD2	–H1N	2.71
Gln175-OE1	–H2N	2.73
Asp178-OD2	–HO	2.98
Gln197-NE2H	–OC2	2.89
Gln197-OE1	–H3N	2.87
*Mt*TyrRS	ATP	
His50-NE2H	–O2'	3.28
Val224-O	–H20N6	2.81
Val224-NH	–N1	3.03
Phe232-O	–H1N6	2.86
Lys234-NZHZ2	–O2A	2.61
*Mt*TyrRS	Tyr-AMP	
Tyr36-OH	–OH	3.11
Gly38-O	–H24OAE	2.89
Asp40-NH	–OAD	3.04
Gln175-OE1	–H1N	2.69
Asp178-OD2	–HO	2.71
Gly194-NH	–O2'	2.67
Asp196-OD1	–HO3'	2.62
Gln197-NE2H	–O5'	3.23
Gln197-OE1	–H2N	2.86
Val224-O	–H1N6	3.01
Val224-NH	–N1	3.28
Phe232-O	–H2N6	3.12

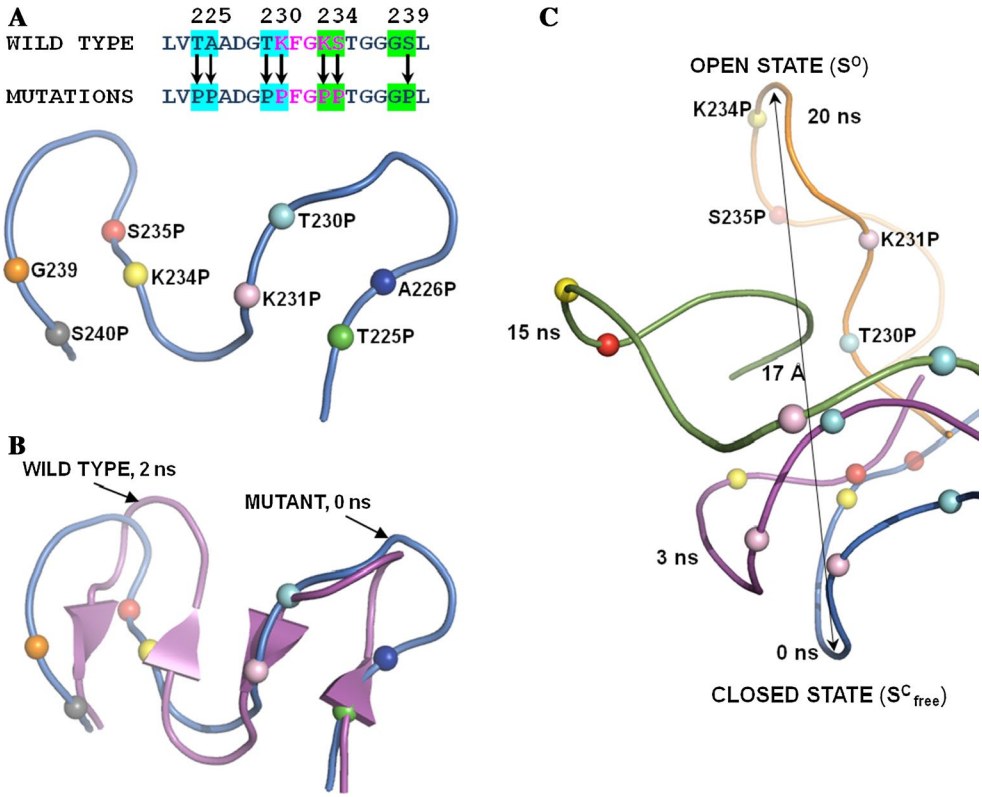

Fig. 6 Mutations of the catalytic loop that prevent the formation of β-structures and lead to opening of the KFGKS loop. **a** Sequence and tertiary structure of the catalytic loop with mutations T225P, A226P, T230P, K231P, K234P, S235P, and S240P. Residues of the KFGKS sequence are shown in *magenta*, residues that form β-structures are highlighted in *cyan*. **b** The wild type catalytic loop dynamically forms β-structures which keep the loop in the "closed" M-like state (*magenta*, 3 ns of MD simulations). The initial conformation (no β-structures) of the mutant loop is shown in *blue*. **c** The introduced mutations prevent formation of β-structures in the catalytic loop, which results in the conformation changing into the "open" O-like state in 20 ns of MD simulations. The difference between positions in the "closed" and "open" states is approximately 17 Å

Structural state and dynamics of the *Mt*TyrRS catalytic loop in the absence of flanking β-hairpins

To verify our hypothesis that flanking β-hairpins in the catalytic loop keep it in the closed M-like state, which directs the catalytic KFGKS motif toward the active center cavity, we performed computational mutagenesis of *Mt*TyrRS substituting residues involved in β-sheet formation with proline: T225P, A226P, T230P, K231P, K234P, S235P, and S240P (Fig. 6a). It is known that proline could abolish the propensity of the peptide to form stable secondary structure (Deber et al. 2010).

Figure 6c shows the results of MD simulations performed with the mutant *Mt*TyrRS. Remarkably, the catalytic loop starts to move out of the active site cavity from the very beginning of MD simulation. The M-shape of the loop (initial structure) rapidly changes (in a 20-ns time interval) to an "open" ring-like structure (O-structure) which is fully exposed to the solvent (Video-2S, supplemental material). The O-like structure is energetically favorable in the absence of flanking β-hairpins (Fig. 6b), and is under control of inner diffusion processes. The transition from the "closed" M-state to the "open" O-state is accompanied by large-scale changes in the RMSD of the KFGKS residues of the loop. As a result, this fragment is spatially displaced from its initial "closed" state by ~17 Å (Fig. 6c).

The results of MD simulations suggest that the main forces stabilizing the closed state of the KMSKS-like loop are interchain H-bonds that stabilize two flanking β-hairpins and consequently the whole M-like conformation of the catalytic loop. In the closed state the KFGKS loop has no significant contacts with the rest of the protein, which explains its capacity to make "pestle in a mortar" movements inside the cavity, depending on the type of bound substrate.

This result elucidates the important role of β-hairpin structures in maintaining the active M-state of the catalytic KFGKS loop inside the enzyme active center.

Analysis of forces stabilizing complexes of TyrRS with different cognate substrates

Hydrogen bonds between the *Mt*TyrRS active site and different substrates revealed by MD simulations are listed in Table 2. It is notable that there is no direct contact between the L-tyrosine and the catalytic loop of the enzyme, whereas the complex of L-tyrosine with TyrRS is stabilized by six hydrogen bonds with Tyr36, Asp40, and Asp178, as shown by Li et al. (2008), in addition to those with Gln175 and Gln197.

In contrast with L-tyrosine binding, ATP interacts electrostatically with the positively charged Lys231 and Lys234 residues of the loop, and also forms H-bonds with His50, Val224 (two H-bonds), Phe232, and Lys234 (three of

them are from the catalytic loop; Table 2) (Fig. 2S, supplemental material). The hydrogen bond network between the tyrosyl–adenylate and *Mt*TyrRS contains 12 hydrogen bonds. Some form the H-bond network that stabilizes the complex with L-tyrosine and another one that is involved in complex with ATP. Tyrosyl–adenylate forms also additional specific H-bonds with *Mt*TyrRS via Gly38, Gly194, and Asp196 residues.

Conclusion

We have presented an analysis of the structural stability and dynamics of the *Mt*TyrRS dimer in solution according to MD simulation data. Specifically, we have investigated the structural and dynamical properties of the *Mt*TyrRS catalytic KFGKS loop in the substrate-free state (S^C_{free}) and in complexes with different substrates: L-tyrosine (S^C_{Tyr}), ATP (S^C_{ATP}), and tyrosyl–adenylate intermediate (S^C_{YMP}). The conformation of the catalytic loop in the "closed" state depends on the type of cognate substrate in the active center. The flexibility of the loop enables adoption of several specific conformations stabilized by interactions with different ligands. On transition from the S^C_{free} state to the S^C_{ATP} state the catalytic loop shifts by 3.3 Å, whereas from the S^C_{free} to the S^C_{YMP} state it shifts by 6.7 Å. It was found in the course of MD simulations that the catalytic loop dynamically forms short antiparallel β-sheets involving the Lys231, Lys234, and Ser235 residues of the universal catalytic KMSKS-like motif of Class I aaRSs. The β-hairpins in the uncomplexed enzyme, being in rapid dynamic equilibria (ps to ns range), keep the loop in the "closed" state, orienting the KFGKS sequence toward the active center and making it ready for substrate binding. Prevention of β-hairpin formation in a mutant *Mt*TyrRS drives the catalytic loop into the "open" inactive conformation, moving the catalytic KFGKS motif away from the active site cavity.

Our MD simulations revealed that the flexibility and rapid dynamics of the wild-type *Mt*TyrRS catalytic loop structure is crucial for formation of protein–substrate interactions and subsequently for overall enzyme functional activity.

Acknowledgments This work was supported by grant no 15/2013 of the State target scientific and technical program "Implementation and application grid technologies 2009–2013".

References

Abergel C, Rudinger-Thirion J, Giegé R, Claverie JM (2007) Virus-encoded aminoacyl-tRNA synthetases: structural and functional characterization of mimivirus TyrRS and MetRS. J Virol 81:12406–12417. doi:10.1128/JVI.01107-07

Austin J, First E (2002) Comparison of the catalytic roles played by the KMSKS motif in the human and *Bacillus stearothermophilus* tyrosyl-tRNA synthetases. J Biol Chem 277:28394–28399. doi:10.1074/jbc.M204404200

Bedouelle H (1990) Recognition of tRNA(Tyr) by tyrosyl-tRNA synthetase. Biochimie 72:589–598

Berendsen HJ, Van Der Spoel D, Van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun 91:43–56

Bjelkmar P, Larsson P, Cuendet M, Hess B, Lindahl E (2010) Implementation of the CHARMM force field in GROMACS: analysis of protein stability effects from correction maps, virtual interaction sites, and water models. J Chem Theory Comput 6:459–466. doi:10.1021/ct900549r

Bonnefond L, Giegé R, Rudinger-Thirion J (2005) Evolution of the tRNATyr/TyrRS aminoacylation systems. Biochimie 87:873–883. doi:10.1016/j.biochi.2005.03.008

Brick P, Bhat TN, Blow DM (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 A resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. J Mol Biol 208:83–98. doi:10.1016/0022-2836(89)90090-9

Budiman ME, Knaggs MH, Fetrow JS, Alexander RW (2007) Using molecular dynamics to map interaction networks in an aminoacyl-tRNA synthetase. Proteins 68:670–689. doi:10.1002/prot.21426

Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. J Chem Phys 126:014101

Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The amber biomolecular simulation programs. J Comput Chem 26:1668–1688

Datt M, Sharma A (2014) Conformational landscapes for KMSKS loop in tyrosyl-tRNA synthetases. J Struct Funct Genomics. doi:10.1007/s10969-014-9178-x

Deber CM, Brodsky B, Rath A (2010) Proline residues in proteins. Encycl Life Sci. doi:10.1002/9780470015902.a0003014.pub2

Eitner K, Gaweda T, Hoffmann M, Jura M, Rychlewski L, Barciszewski J (2007) eHiTS-to-VMD interface application. The search for tyrosine-tRNA ligase inhibitors. J Chem Inf Model 47:695–702. doi:10.1021/ci600392r

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8592

First EA, Fersht AR (1993) Involvement of threonine 234 in catalysis of tyrosyl adenylate formation by tyrosyl-tRNA synthetase. Biochemistry 32:13644–13650

First EA, Fersht AR (1995) Analysis of the role of the KMSKS loop in the catalytic mechanism of the tyrosyl-tRNA synthetase using multimutant cycles. Biochemistry 34:5030–5043

Fiser A, Sali A (2003) ModLoop: automated modeling of loops in protein structures. Bioinformatics 19:2500–2501. doi:10.1093/bioinformatics/btg362

Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. Protein Sci 9:1753–1773. doi:10.1110/ps.9.9.1753

Hess B, Bekker H, Berendsen HJ, Fraaije JG (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 Theory Comput 4:435–447. doi:10.1021/ct700301q

Hoffmann M, Torchala M (2009) Search for inhibitors of aminoacyl-tRNA synthases by virtual click chemistry. J Mol Model 15:665–672. doi:10.1007/s00894-008-0421-x

Hornak V, Abel R, Okur O, Strockbine B, Roitberg A, Simmerling C (2006) Comparison of multiple amber force fields and

development of improved protein backbone parameters. Proteins 65:712–725. doi:10.1002/prot.21123
Hughes SJ, Tanner JA, Miller AD, Gould IR (2006) Molecular dynamics simulations of LysRS: an asymmetric state. Proteins 62:649–662. doi:10.1002/prot.20609
Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Mol Graph 14:33–38
Joosten RP, te Beek TA, Krieger E, Hekkelman ML, Hooft RW, Schneider R, Sander C, Vriend G (2011) A series of PDB related databases for everyday needs. Nucleic Acids Res 39:411–419. doi :10.1093/nar/gkq1105
Jorgensen WL, Madura JD (1983) Solvation and conformation of methanol in water. J Am Chem Soc 105:1407–1413
Kamijo S, Fujii A, Onodera K, Wakabayashi K, Kobayashi T, Sakamoto K (2008) Investigateion of requirements for the KMSKS loop in aminoacyl-tRNA synthetase by random PCR method. J Proteomics Bioinform S2:227. doi:10.4172/jpb.s1000135
Kapustina M, Carter CW (2006) Computational studies of tryptophanyl-tRNA synthetase: activation of ATP by induced-fit. J Mol Biol 362:1159–1180. doi:10.1016/j.jmb.2006.06.078
Keutsch FN, Saykally RJ (2001) Water clusters: untangling the mysteries of the liquid, one molecule at a time. Proc Natl Acad Sci USA 98:10533–10540. doi:10.1073/pnas.191266498
Kobayashi T, Takimura T, Sekine R, Kelly VP, Kamata K, Sakamoto K, Nishimura S, Yokoyama S (2005a) Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase. J Mol Biol 346:105–117. doi:10.1016/j.jmb.2004.11.034
Kobayashi T, Sakamoto K, Takimura T, Sekine R, Kelly VP, Kamata K, Nishimura S, Yokoyama S (2005b) Structural basis of nonnatural amino acid recognition by an engineered aminoacyl-tRNA synthetase for genetic code expansion. Proc Natl Acad Sci USA 102:1366–1371. doi:10.1073/pnas.0407039102
Kornelyuk AI (1998) Structural and functional investigation of mammalian tyrosyl-tRNA synthetase. Biopolym Cell 14:349–359
Leatherbarrow RJ, Fersht AR, Winter G (1985) Transition-state stabilization in the mechanism of tyrosyl-tRNA synthetase revealed by protein engineering. Proc Natl Acad Sci USA 82:7840–7844
Li T, Froeyen M, Herdewijn P (2008) Comparative structural dynamics of tyrosyl-tRNA synthetase complexed with different substrates explored by molecular dynamics. Eur Biophys J 38:25–35. doi:10.1007/s00249-008-0350-8
Odynets KO, Kornelyuk OI (2008) A model of three-dimensional structure of *Mycobacterium tuberculosis* tyrosyl-tRNA synthetase. Ukr Biokhim Zh 80:62–75
Pal SK, Zhao L, Xia T, Zewail AH (2003a) Site- and sequence-selective ultrafast hydration of DNA. Proc Natl Acad Sci USA 100:13746–13751
Pal SK, Zhao L, Zewail AH (2003b) Water at DNA surfaces: ultrafast dynamics in minor groove recognition. Proc Natl Acad Sci USA 100:8113–8118
Palencia A, Crépin T, Vu M, Lincecum T, Martinis S, Cusack S (2012) Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. Nat Struct Mol Biol 19:677–685. doi:10.1038/nsmb.2317
Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 52:7182
PyMOL (2011) The PyMOL molecular graphics system, version 1.4.1. Schrödinger, LLC
Qiu X, Janson CA, Smith WW, Green SM, McDevitt P, Johanson K, Carter P, Hibbs M, Lewis C, Chalker A, Fosberry A, Lalonde J, Berge J, Brown P, Houge-Frydrych CS, Jarvest RL (2001) Crystal structure of Staphylococcus aureus tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors. Protein Sci 10:2008–2016. doi:10.1110/ps.18001
Retailleau P, Yin Y, Hu M, Roach J, Bricogne G, Vonrhein C et al (2001) High-resolution experimental phases for tryptophanyl-tRNA synthetase (TrpRS) complexed with tryptophanyl-5′AMP. Acta Crystallogr D Biol Crystallogr 57:1595–1608. doi:10.1107/S090744490101215X
Salnikov AO, Sliusar IA, Sudakov OO, Savytskyi OV, Kornelyuk AI (2010) Virtual laboratory moldyngrid as a part of scientific infrastructure for biomolecular simulations. Int J Comput 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. In: 2011 IEEE 6th international conference on intelligent data acquisition and advanced computing systems (IDAACS), vol 1, pp 208–211. doi:10.1109/IDAACS.2011.6072742
Savytskyi OV, Yesylevskyy SO, Kornelyuk AI (2013) Asymmetric structure and domain binding interfaces of human tyrosyl-tRNA synthetase studied by molecular dynamics simulations. J Mol Recognit 26:113–120. doi:10.1002/jmr.2259
Thompson D, Simonson T (2006) Molecular dynamics simulations show that bound $Mg2^{+}$ contributes to amino acid and aminoacyl adenylate binding specificity in aspartyl-tRNA synthetase through long range electrostatic interactions. J Biol Chem 281:23792–23803. doi:10.1074/jbc.M602870200
Thompson D, Plateau P, Simonson T (2006) Free-energy simulations and experiments reveal long-range electrostatic interactions and substrate-assisted specificity in an aminoacyl-tRNA synthetase. ChemBioChem 7:337–344. doi:10.1002/cbic.200500364
Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) GROMACS: fast, flexible and free. J Comput Chem 26:1701–1718
Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general AMBER force field. J Comput Chem 25:1157–1174. doi:10.1002/jcc.20035
Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25:247–260. doi:10.1016/j.jmgm.2005.12.005
Weinreb V, Li L, Chandrasekaran SN, Koehl P, Delarue M, Carter CW Jr (2014) Enhanced amino acid selection in fully evolved tryptophanyl-tRNA synthetase, relative to its urzyme, requires domain motion sensed by the D1 switch, a remote dynamic packing motif. J Biol Chem 289:4367–4376. doi:10.1074/jbc.M113.538660
Xin Y, Li W, First E (2000) The 'KMSKS' motif in tyrosyl-tRNA synthetase participates in the initial binding of tRNATyr. Biochemistry 39:340–347
Yaremchuk A, Kriklivyi I, Tukalo M, Cusack S (2002) Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. EMBO J 21:3829–3840. doi:10.1093/emboj/cdf373
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. doi:10.1016/j.bpc.2011.01.005
Zoete V, Cuendet MA, Grosdidier A, Michielin O (2011) SwissParam: a fast force field generation tool for small organic molecules. J Comput Chem 32:2359–2368. doi:10.1002/jcc.21816