Exploring Cooperative Domain Dynamics in *Thermus thermophilus* Leucyl-tRNA Synthetase Using Low-frequency Normal Mode Calculations and Statistical Coupling Analysis

Kristina Weimer, Brianne Shane, Michael Brunetto, Sudeep Bhattacharyay, and Sanchita Hati
Department of Chemistry, University of Wisconsin–Eau Claire, WI-54702

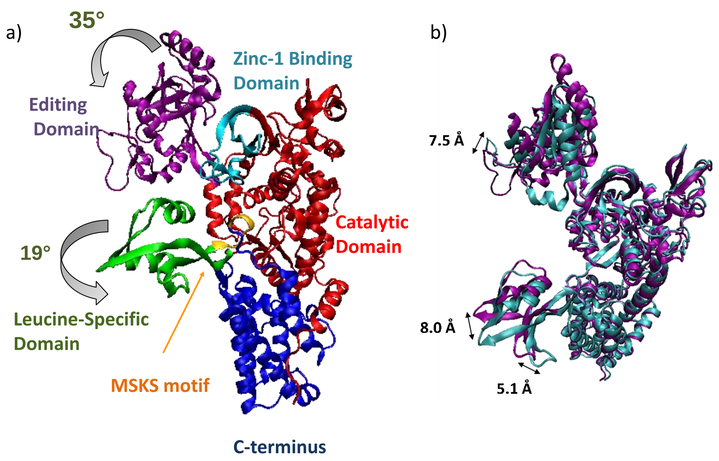
Acknowledgements:
Research Corporation Cottrell College Science Award
UWEC-Office of Research and Sponsored Programs

Abstract

Leucyl-tRNA synthetases are class I synthetases that catalyze the covalent attachment of leucine to the $tRNA^{Leu}$. The three-dimensional crystal structure of *Thermus thermophilus* leucyl-tRNA synthetase (Tt LeuRS) demonstrates a complex modular architecture where three flexible domains [the conserved connective polypeptide 1 (CP1) domain (residues 224-417), the leucine-specific (LS) domain (residues 577-634), and the zinc-1 (ZN-1) binding domain (residues 154-189)] are inserted into the central catalytic domain (*1*). The crystal structure of the Tt LeuRS-$tRNA^{Leu}$ complex (in post-transfer-editing conformation) demonstrated that the CP1 domain undergoes a rotation of 35° from the position observed in the tRNA unbound form. The LS domain, which is critical for aminoacylation, undergoes a rigid-body rotation of 19°(*2*). Various structural elements (including catalytically important $H^{49}MGH$ and $V^{638}MSKS$ loops) in the central catalytic core also undergo considerable conformational changes due to leucyl-adenylate binding (*1*). These substrate induced conformational rearrangements of various structural elements of Tt LeuRS suggest that cooperative domain dynamics play an important role in the enzyme function. In the present work, we have investigated the collective motion of various structural elements in Tt LeuRS using normal mode calculations. In addition, statistical coupling analysis has been performed to examine if the evolutionarily coupled networks of residues have significant contributions to these concerted domain motions. Taken together, these studies demonstrate that domain motions in Tt LeuRS are indeed cooperative in nature and lead to the identification of the network of residues that propagate long-range interdomain communications in this enzyme.

Objectives

To understand the domain dynamics of Tt LeuRS at the molecular level and to identify residue networks that mediate domain-domain communications in this enzyme.



a) A cartoon diagram of the structure of *T. Thermophilus* LeuRS. b) $tRNA^{Leu}$ unbound (1h3n, magenta) and bound (2byt, cyan) structures were superimposed.

Normal Mode Analysis (NMA)

The collective motion of the structural elements of a large biomolecule can be represented by normal modes. It is believed that the lowest frequency (large-amplitude) normal modes of a multi-domain enzyme describe the functionally relevant motions. Normal mode calculation is based on the harmonic approximation of the potential energy function around a minimum energy conformation. In this work NMA was carried out using the elastic network model (*3*). In the elastic network model, protein residues are represented by only their C_α atoms. The C_α atoms on a protein backbone are considered to be connected by uniform springs and the harmonic potential is given by:

$$E_p = \sum_{d^o_{pq} < R_c} C(d_{pq} - d^o_{pq})^2$$

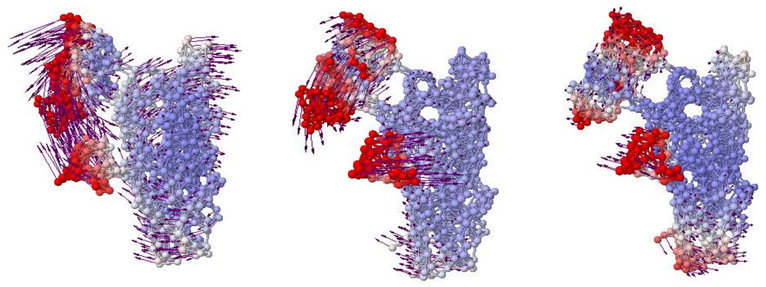
where d_{pq} is the distance between atoms *p* and *q*, $d_{pq}{}^0$ is the distance between these two atoms in the given crystallographic structure, *C* is the strength of the potential, and R_c is an arbitrary cut-off parameter which defines the maximum interaction range between C_α atoms.

In this work Anisotropic Network Model (ANM) is used to describe functionally important collective motions of Tt LeuRS . In ANM the fluctuations are anisotropic and the overall potential of the system is a sum of the harmonic potentials given by:

$$V_{\mathrm{ANM}} = \frac{\gamma}{2}\left[\sum_{q,q\neq p} \Gamma_{pq} (R_{pq} - R^o_{pq})^2\right]$$

where γ represents the uniform spring constant, $R_{pq}{}^0$ and R_{pq} are the original and instantaneous distance vectors between residues *p* and *q*, Γ_{pq} is the *pq*-th element of the connectivity matrix of inter-residue contacts. Based on an interaction cut-off distance of R_c, Γ_{pq} is equal to 1 if $R_{pq}{}^0 < R_c$ and zero otherwise (*4*). The online server http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi was used to analyze the functional motion of Tt LeuRS. The optimal cut-off interactions between C_α atoms was kept at 15Å.

Collective Domain Motions in Tt LeuRS is Best Described by Mode 1

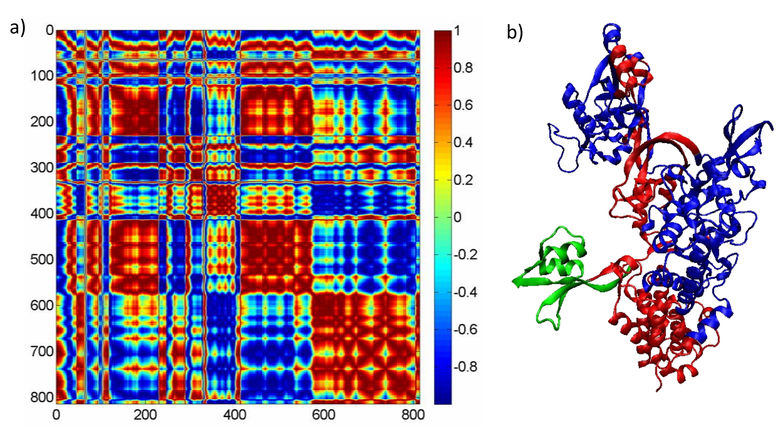


Three lowest-frequency normal modes obtained from ANM calculations using open-form structure of Tt LeuRS

Overlap, Correlation , and collectivity value of the three lowest-frequency modes involved in the conformational change

Properties	Mode 1	Mode 2	Mode 3
Overlap (direction of motion)	0.72	0.60	0.71
Correlation (magnitude of motion)	0.51	0.56	0.31
Collectivity	0.37	0.20	0.42

NMA Study Revealed Anticorrelation Between the Fluctuations of Structural Elements in LS and CP1 Domains



a) Cross-correlations map for residue fluctuations in mode 1 that is most involved in the conformational change. Correlated motion is shown by positive values up to 1 (green to red) and anticorrelated motion is shown by negative values down to -1 (cyan to blue); b) Protein segments engaged in correlated and anticorrelated motion with respect to LS domain (green) are colored in red and blue, respectively.

Statistical Coupling Analysis (SCA)

SCA is based upon the assumption that "coupling of two sites in a protein, whether for structural or functional reasons, should cause those two positions to co-evolve" (*5*). The overall evolutionarily conservation parameter at a position *i* in the sequence of the chosen protein family is calculated and expressed as

$$\Delta G_i^{\mathrm{stat}} = kT * \sqrt{\sum_x [\ln(P_i^x / P_{MSA}^x)]^2}$$

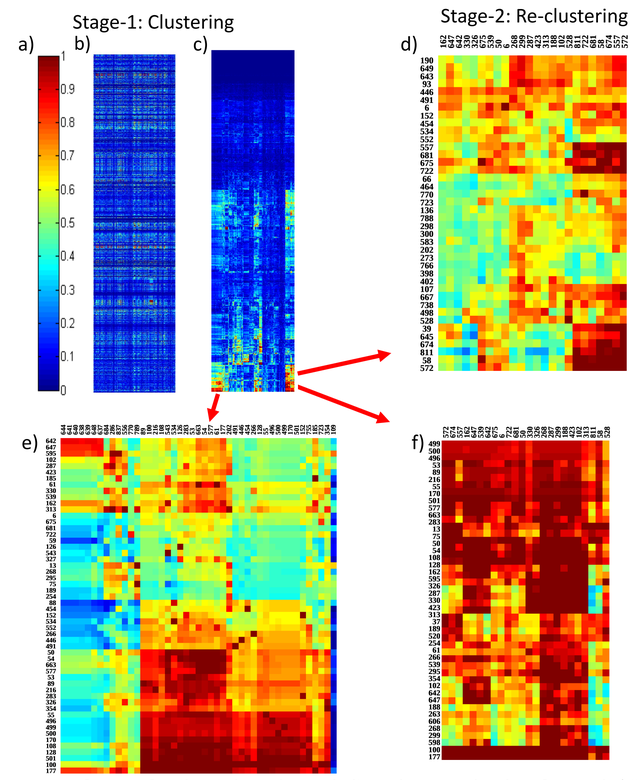
where *kT** is an arbitrary energy unit, P_i^x is the probability of any amino acid *x* at site *i*, and $P_{MSA}{}^x$ is the probability of *x* in the MSA. The coupling of site *i* with site *j* is calculated and expressed as

$$\Delta\Delta G_{i,j}^{stat} = kT * \sqrt{\sum_x [\ln(P_i^x \big| \delta_j / P_{MSA}^x \big| \delta_j) - \ln(P_i^x / P_{MSA}^x)]^2}$$

where $P_i^x \mid \delta_j$ is the probability of *x* at site *i* dependent on perturbation at site *j*.

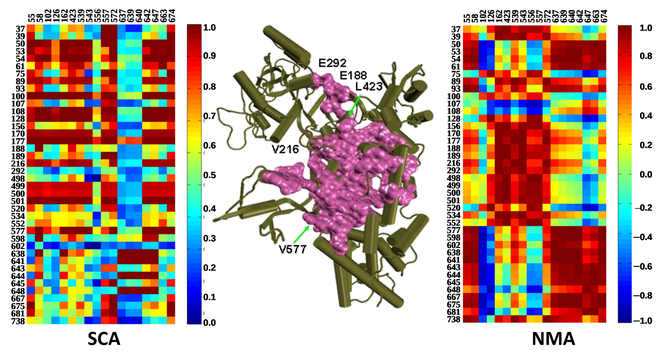
We performed SCA on an alignment of 484 protein sequences of LeuRS family. The SCA was performed by systematically perturbing each position where a specific amino acid was present in at least 50% of the sequences in the alignment. The initial clustering resulted in a matrix with 876 (residue number)× 216 (perturbation site) matrix elements representing the coupling between residues. The SCA on the LeuRS family demonstrates a group of residues which have coevolved in the Tt LeuRS.

Statistically Coupled Residues in Tt LeuRS



SCA of LeuRS family. a) the color scale linearly maps the data from 0 kT^* (blue) to 1 kT^* (red); b) the unclustered matrix; c) statistical coupling matrix where rows represent positions (N to C terminus, top to bottom) and columns represent perturbations (N to C terminus, left to right); d-f) Two dimensional clustering showing three separate co-evolving networks.

Thermally Coupled and Coevolved Residues



Mapping of the thermally and evolutionarily coupled residue network on the 3D structure of Tt LeuRS. Out of 876 residues of Tt LeuRS, only 58 residues were identified which exhibit strong co-evolutionary pattern of variations as well as coupled dynamics.

Conclusions

- *Cooperative Domain Dynamics*: The NMA study demonstrates that mode 1 adequately describes the conformational change in the Tt LeuRS. Analysis of the motion indicates that the LS and CP1 domains are engaged in anticorrelated motion.
- *Coevolved Residue Network*: SCA has identified a core set of residues which are evolutionarily coupled and reside at the domain interface. They form a sparse but contiguous network of interactions between the domains.
- *Identifying a Functionally Relevant Network*: Combined results of the NMA and SCA have produced a subset of residues which are not only correlated by evolution but also are coupled by thermal motions. These residues are within the van der Waals contact and appear to be critical for maintaining key structural scaffolds and domain dynamics in Tt LeuRS.
- *Mutational Data*
  - Existing *E. coli* (Ec) LeuRS mutational data demonstrates that mutation of some of these evolutionarily and thermally coupled residues have a strong impact on enzyme function. For example, mutation of a single residue at the interface of LS and catalytic domains (position 577) alters amino acid discrimination and tRNA aminoacylation (*6*).
  - Mutation of positions 292 and 188 have significant effect on enzyme catalysis.
  - Various constructs of Ec LeuRS, obtained by deleting the "hinge" regions, have a profound effect on editing reaction. Some of these hydrophobic "hinge" residues (position 423 and 216) are thermally and evolutionarily coupled with the main body of the enzyme (*7*).
  - These mutational results support the validity of this combined NMA-SCA approach to identify the important residues which are involved in maintaining the cooperative domain dynamics.
- *Future work*: Mutational studies to further explore the role of networking residues (*identified in this work*) that mediate long-range communications between domains.

References
1. Cusack, S., Yaremchuk, A., and Tukalo, M. (2000) *EMBO J.* **19**, 2351-2361.
2. Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S., and Cusack, S. (2005) *Nat. Struct. Mol. Biol.* **12**, 923-930.
3. Tirion, M. M. (1996) *Phys. Rev. Lett.* **77**, 1905-1908.
4. Eyal, E., Yang, L. W., and Bahar, I. (2006) *Bioinformatics* **22**, 2619-2627.
5. Lockless, S. W., and Ranganathan, R. (1999) *Science 286*, 295-299.
6. Lue, S. W. (2007) *Biochemistry* **46**, *2266-4472.*
7. *Betha, A. K., Williams, A. M., Martinis, S. A.* (2007) *Biochemistry* **46**,*6258-6267.*
8.Williams, A.M., and Martinis, S.A. (2006) *PNAS* **103**, *3586-3591.*
9.Du, X., and Wang, E. (2002) *Biochemistry* **41**, 10623-10628.