BIOCHEMISTRY
including biophysical chemistry & molecular biology

ARTICLE

pubs.acs.org/biochemistry

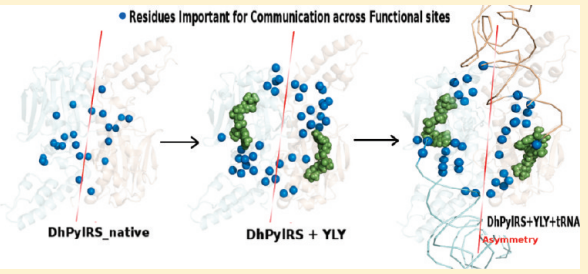
Probing the Allosteric Mechanism in Pyrrolysyl-tRNA Synthetase Using Energy-Weighted Network Formalism

Moitrayee Bhattacharyya and Saraswathi Vishveshwara*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Supporting Information

ABSTRACT: Pyrrolysyl-tRNA synthetase (PylRS) is an atypical enzyme responsible for charging $tRNA^{Pyl}$ with pyrrolysine, despite lacking precise tRNA anticodon recognition. This dimeric protein exhibits allosteric regulation of function, like any other tRNA synthetases. In this study we examine the paths of allosteric communication at the atomic level, through energy-weighted networks of *Desulfitobacterium hafniense* PylRS (DhPylRS) and its complexes with $tRNA^{Pyl}$ and activated pyrrolysine. We performed molecular dynamics simulations of the structures of these complexes to obtain an ensemble conformation−population perspective. Weighted graph parameters relevant to identifying key players and ties in the context of social networks such as edge/node betweenness, closeness index, and the concept of funneling are explored in identifying key residues and interactions leading to shortest paths of communication in the structure networks of DhPylRS. Further, the changes in the status of important residues and connections and the costs of communication due to ligand induced perturbations are evaluated. The optimal, suboptimal, and preexisting paths are also investigated. Many of these parameters have exhibited an enhanced asymmetry between the two subunits of the dimeric protein, especially in the pretransfer complex, leading us to conclude that encoding of function goes beyond the sequence/structure of proteins. The local and global perturbations mediated by appropriate ligands and their influence on the equilibrium ensemble of conformations also have a significant role to play in the functioning of proteins. Taking a comprehensive view of these observations, we propose that the origin of many functional aspects (allostery and half-sites reactivity in the case of DhPylRS) lies in subtle rearrangements of interactions and dynamics at a global level.



Aminoacyl-tRNA synthetases (aaRS) are primarily responsible for the fidelity in the translation of the genetic code, ensuring charging of amino acids to their cognate tRNAs by efficient long-range communication and editing mechanisms.[1] The response to the identification of the correct tRNA is elicited in the release of the amino acid to the 3′ end of cognate tRNA. Such efficient communication between distant functional sites is the hallmark of allosteric proteins. A network of interacting residues contributes to the cross-talk between tRNA binding and active sites, making aaRS an excellent model for studying the allosteric mechanism. The canonical stop codon UAG is translated to pyrrolysine (22nd ribosomally inserted amino acid) in members of archaeal family Methanosarcinaceae and *Desulfitobacterium hafniense* through pyrrolysyl-tRNA synthetase (PylRS).[2] PylRS has been shown to exhibit "diffused substrate recognition" by recognizing a wide variety of lysine derivatives as substrates, mediated by a loose recognition of their α-amino moieties.[3] As a consequence, this orthogonal aaRS−tRNA pair (PylRS−$tRNA^{Pyl}$) is well suited for the engineered expansion of the genetic code.[3] Furthermore, *D. hafniense* PylRS (DhPylRS) is an exception to canonical aaRS also at the level of tRNA base recognition. Patches of surface residues of DhPylRS make contact with the acceptor stem and D arm of the tRNA,[4] and it has also been demonstrated that the anticodon arm of $tRNA^{Pyl}$ comprising the "CUA" triplet anticodon is not a determinant of aminoacylation efficiency.[4,5] The mutations of some basic residues of the interface patch of recognition, leading to reduction or abolishment of catalytic activity, further validated this observation.[4] A recently determined crystal structure of DhPylRS (dimer) bound to $tRNA^{Pyl}$ has given further structural insights into such unique interactions with its cognate tRNA, accounting for the orthogonality of this aaRS−tRNA pair.[6] The discrimination of other canonical tRNAs by DhPylRS arising from shape complementary recognition of the tightly packed tertiary core of $tRNA^{Pyl}$ (largely through tRNA binding domain 1 unique to PylRS) is also evident from the crystal structure of DhPylRS complexed with $tRNA^{Pyl}$.[6]

Classical models of allostery (MWC[7] and KNF[8]) have made a great impact on enzymology. However, the allosteric phenomenon has been revisited and redefined (population shift model) primarily to gain understanding at the molecular level.[9] The key question is how to delineate the train(s) of interacting amino acid residues that transmits the signal between distal sites. In systems like aaRS, it is also important to understand the role of

Received: March 1, 2011
Revised: June 8, 2011
Published: June 08, 2011

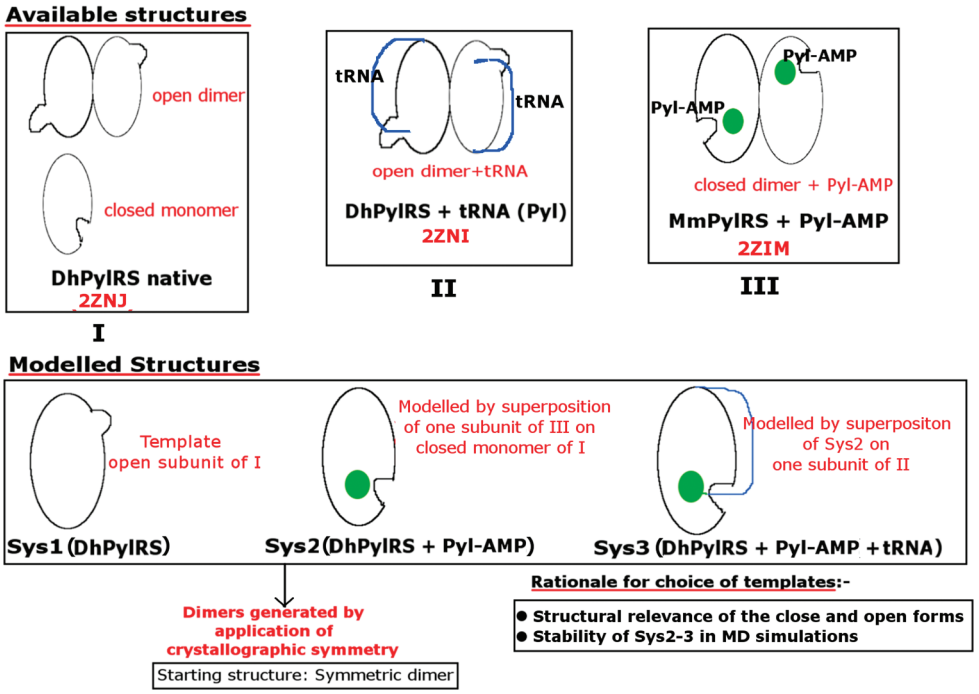


Figure 1. Detailed schematic representation of the modeling exercise used to obtain the starting structures for Sys1−3. The cartoons in the top panel represent the different available structures for PylRS (in various ligand-bound states). The middle panel shows the modeled structures for Sys1−3. The binding of Pyl-AMP to PylRS is known to induce a closed state. Therefore, Sys2 and Sys3 are modeled on the basis of the closed form from PDB entry 2znj, whereas the native structure is modeled using the open form. The bottom panel gives a general rationale for the modeling procedure.

nucleotides in such long-range signaling. In addition, dynamics is crucial for function, and hence, the dynamical properties such as equilibrium conformations and correlated movements of residues need to be explored. Some of these issues have been recently addressed in the literature. For example, the involvement of multiple preexisting pathways of communication has been postulated.[10] It is proposed that key allosteric residues are not restricted to definite communication pathways but are the ones responsible for overall conformational and dynamical transitions upon perturbation by ligand binding and/or environmental changes.[10] It is also proposed that a network of coupled amino acid residues is responsible for the effective transmission of perturbation from one site to another, leading to a conformational (enthalpic) or dynamical (entropic) change in the protein, at both the backbone and the side chain levels.[9,11] Several experimental[12,13] and computational techniques[14,15] have been applied to identify the set of residues responsible for such long-range information transfer. We have probed the allosteric pathways of communication by graph theoretical approaches along with MD simulations.[16,17] The contribution of conserved nodes to the characteristic path length in a network has been used to identify residues important for global communication.[18] A correlation-weighted protein−tRNA complex network has been used to recognize the pathways of communication in glutaminyl-tRNA synthetase.[19] The contribution of residues to allosteric cross-talk has also been estimated in terms of equilibrium fluctuations from elastic network models (ENM).[20]

The encoding of function involves precise interaction among relevant components (aaRS, tRNA, and activated amino acid in this case). The subtle structural changes are communicated through basic physicochemical principles such as the energy of interaction among the interacting partners. In this study, we explore allosteric communication by studying interaction energy-weighted protein−tRNA structure networks and the dynamical properties of different ligand-bound states of DhPylRS. The incorporation of chemistry in terms of interaction energies, use of specialized weighted graph parameters for analysis, and exploring biologically relevant paths of communication in different liganded states of DhPylRS are the main features introduced in this study. Further, a realistic equilibrium ensemble perspective is provided because these analyses were performed on a set of structures obtained from MD simulations. We have chosen the dimeric DhPylRS in different states of ligation (Sys1, native DhPylRS; Sys2, DhPylRS−Pyl-AMP; Sys3, DhPylRS−Pyl-AMP−tRNA) to gain insight into the functioning of this noncanonical aminoacyl-tRNA synthetase as well as to understand the mechanism of communication. A comparison of dynamics and network parameters among the three systems provides detailed information about the ligand induced conformational changes and subtle variations in structures that lead to an efficient functional outcome in DhPylRS. We have probed different weighted network parameters to elucidate the key residues for signal propagation across distant sites. Our analysis reveals the importance of side chain reorganization with minimal backbone conformational changes in the three systems being studied. The transfer of an amino acid to the 3′ end of cognate tRNA seems to be mediated by alternation of global rigidity and flexibility, and a general mechanistic insight for aminoacylation is presented. Further, long-range residue coupling is examined, and its functional contribution to global signaling is explored. We also correlate our results with energy-based residue correlation and available experimental observations for DhPylRS.[5]

The concept of half-sites reactivity is discussed extensively in the literature in the context of the functioning of multimeric proteins,[21,22] the underlying principles for which are yet to

emerge. Half-sites reactivity is an extreme case of negative cooperativity in which activity at one site completely precludes reaction at another site,[22] giving rise to asymmetry in function. In this study, we have probed the relevance of the dimeric structure of DhPylRS for its function. We relate half-sites reactivity to the asymmetry of interactions and residue correlations between the two subunits in the ensemble of equilibrium structures. Importantly, it appears that the rearrangement of interactions upon perturbation is the key to diverse biological functions. For instance, in this case of DhPylRS, both allostery and half-sites reactivity are guided by this kind of rearrangement. Here we bring out the underlying physical principles for such phenomena by the investigation of the energy-weighted protein−tRNA network and capturing the dynamics from the MD ensemble.

METHODS

Choice and Modeling of Systems (Sys1−3) for MD Simulations. The three systems of interest are Sys1 (DhPylRS), Sys2 (DhPylRS−Pyl-AMP), and Sys3 (DhPylRS−Pyl-AMP−$tRNA^{Pyl}$). The crystal structure of native DhPylRS [Protein Data Bank (PDB) entry 2znj] is used as the starting structure for Sys1. The asymmetric unit of 2znj contains a dimer in the open form and a monomer in the closed form with respect to the Pyl recognition loop.[6] Sys1 is the open state of 2znj (subunit B), and Sys2 and Sys3 are modeled from the closed state of 2znj (subunit C) using the Pyl-AMP from PDB entry 2zim[23] and $tRNA^{Pyl}$ from PDB entry 2zni,[6] respectively, as schematically explained in Figure 1(details of modeling given in Supplemental Methods of the Supporting Information).

MD Simulations and Residue−Residue Interactions. Molecular dynamics (MD) simulations are performed at 300 K using AMBER9[24] with parm99 parameters[25] on the three systems of DhPylRS, Sys1−3 as explained above (details of simulation protocols given in the Supplemental Methods of the Supporting Information). Simulations are conducted for 20 ns each in aqueous medium using the TIP3P water model, and the trajectories from 10 to 20 ns are used for analyses.

Residue−residue interaction energies are calculated from MD trajectories using the MM-PBSA decomposition module in Amber9[24] with an implicit solvent model, to account for solvent effects. This module splits the free energies into individual residue-wise contributions by performing energy decompositions for gas phase energies. The desolvation free energies are calculated with GB (generalized Born solvation models), and nonpolar contributions to desolvation are evaluated using the ICOSA method in the Sander module of AMBER9 (details of specific parameters are discussed in the Supplemental Methods of the Supporting Information). The noncovalent interaction energy (E_{ij}) between residues i and j is defined as $E_{ij} = E_{ij}^{ele} + E_{ij}^{vdw} + E_{ij}^{solvent}$ ($|i-j|>1$), where E_{ij}^{ele}, E_{ij}^{vdw}, and $E_{ij}^{solvent}$ are the electrostatic, van der Waals, and solvent contributions to the interaction energy between i and j, respectively.

Construction of PcEN. Recently, protein energy networks (PEN) have been used in our lab to probe general structural properties of proteins.[26] Here we extend the concept of PEN to protein complexes with other macromolecule(s) (protein, RNA, or DNA) in terms of protein complex energy networks (PcEN) (the interacting partner being tRNA in this study). PcENs are constructed for each system (Sys1−3) by considering amino acid residues and/or nucleotides as nodes. Edges are constructed between all pairs of nodes (excluding the sequence neighbors) with interaction energies (E_{ij}) of >1 kcal/mol. PcENs are constructed for every 50 ps snapshot for each system. PcENs are weighted networks with the normalized weight of an edge between nodes i and j (W_{ij}) defined as

$$W_{ij} = -0.1|E_{ij}| + 20.2 \quad (1)$$

where E_{ij} is the interaction energy between i and j. The interaction energies between two residues are a measure of the ease of information transfer between them. We have used Dijkstra's algorithm for the shortest path computation (detailed in a later subsection) that is based on an efficient optimization of cost. Therefore, we normalize the edge weights as cost by a simple linear transformation (eq 1; see the Supplemental Methods of the Supporting Information for quantitative details of normalization), so that the higher interaction energy connection (higher E_{ij}) is read as the lowest cost (lower W_{ij}) and vice versa to modify our PcEN for calculations using Dijkstra's algorithm. Further, $PcEN_{avg}$ for a system is constructed by averaging the E_{ij} values for each PcEN from individual snapshots.

$$\overline{E}_{ij} = \frac{1}{f}\sum_{i=1}^{f} E_{ij}^{t} \quad (|i-j|>1) \quad (2)$$

where E_{ij}^{t} is the interaction energy between i and j in the tth snapshot and f is the total number of snapshots involved.

Dynamic Correlations: Interaction Energy Correlation, Residue Correlation Matrix, and Binding Pocket Correlation Factor. The valuable information about dynamical correlations of the residue−residue interaction is embedded in simulation snapshots. Extracting such information through constructions of a pairwise interaction energy correlation matrix and an energy-based residue correlation matrix and their significance have been discussed previously in the literature.[27] The correlation of nonbonded interaction energies of all residue pairs provides insights into the conformational couplings in the system.[27,28] The correlation between a pair of residue−residue interactions (ij and kl) is defined as

$$C_{ij|kl} = \frac{\sum_{i=1}^{f}(E_{ij}^{t}-\overline{E}_{ij})(E_{kl}^{t}-\overline{E}_{kl})}{\sum_{i=1}^{f}\sqrt{(E_{ij}^{t}-\overline{E}_{ij})^{2}(E_{kl}^{t}-\overline{E}_{kl})^{2}}} \quad (3)$$

An interaction is included only if the absolute average value exceeds 1 kcal/mol, and covalently bonded neighboring residues are excluded from this calculation. The rows and columns of such an interaction energy correlation matrix depict pairs of interactions. This matrix is projected back onto the residue space to generate an energy-based residue correlation matrix. The correlation between residues i and j (RC_{ij}) is defined as

$$RC_{ij} = \sum_{m=1}^{N}\sum_{n=1}^{N}|C_{m|n} \times \delta_{m|n}^{ij}| \quad (4)$$

where N is equal to the total number of residues in the system and $\delta_{m|n}^{ij}$ is set to 1 if i and j are involved in interactions m and n; otherwise, $\delta_{m|n}^{ij}$ is equal to 0. $C_{m|n}$ is the correlation between interaction pairs m and n. For example, if the correlation between interaction pair [3, 15|17, 33] is 0.8 and that between correlation pair [3, 62|17, 22] is 0.6, then the residue correlation between residues 3 and 17 is considered to be 1.40.[28] The interaction energy correlation and residue correlation matrices are constructed for all

three systems (Sys1−3) (details in Supplemental Methods of the Supporting Information).

Residues with interaction energies greater than 1 kcal/mol with the ligand Pyl-AMP comprise the ligand binding pockets of each subunit, and the overall coupling of other residues to the ligand binding pockets can be estimated by the binding pocket correlation factor (BPCF),[27] which is defined as

$$CF_i = \sum_{j=1}^{N} RC_{ij}^{bp} \quad (5)$$

where CF_i is the BPCF of residue i and RC_{ij}^{bp} is the correlation of residue i with residue j in the binding pocket.

The rearrangements in interaction and correlation for the different ligand-bound states can be quantified by the cumulative differences ($\Delta E_i/\Delta C_i$) in the interaction energy (E_i) and residue correlation (C_i) of each residue. On the basis of the average values of interaction energy and the residue correlation matrix, the deviation in these parameters between two systems (e.g., Sys1 and Sys2) for residue i ($\Delta E_i/\Delta C_i$) is defined as

$$\Delta E_i = \sum_{j=1}^{N} |\overline{E}_{ij}^{Sys1} - \overline{E}_{ij}^{Sys2}| \quad (6)$$

$$\Delta C_i = \sum_{j=1}^{N} |\overline{C}_{ij}^{Sys1} - \overline{C}_{ij}^{Sys2}| \quad (7)$$

where E_{ij} and C_{ij} are the average interaction energy and the correlation between residues i and j, respectively. In this study, the cumulative differences in interaction energy and correlation for Sys1 and Sys3 are computed with Sys2 as the reference.

Shortest Paths, Suboptimal Paths, and Network Parameters. The shortest paths (SPs) or optimal paths (OPs) of communication (lowest-cost path) between two residues in the PcENs are determined using Dijkstra's algorithm. The cost of a SP between pairs of residues of interest, i and j ($cost_{ij}$), lacking direct noncovalent interaction (hereafter termed termini) is the sum of the edge weights (W_{kl}) between the consecutive nodes (kl) constituting the path and is defined as $cost_{ij} = \sum_{kl} W_{kl}$. The lower the cost, the higher the efficiency of communication along a path. The suboptimal paths (SOPs) are the alternate routes of communication with costs greater than those of the OPs. SOPs are determined in this study by systematically removing all interactions of an OP node(s), thus forcing the traversal of a less than optimal path.

Several network parameters are used in this study to unravel the "social structure" in PcENs. A brief description of each parameter and its general physical significance in terms of the network follow (detailed quantitative description in the Supplemental Methods of the Supporting Information).

Local Node Betweenness (LNB). In the literature, node betweenness (NB) of node i is defined as the number of SPs between all pairs of nodes passing through i.[29] We have defined local NB (LNB) as the number of SPs between residues in defined pairs of functional modules [one module being the core-binding surface (CBS) in DhPylRS (i.e., tRNA recognition residues) and the other module consisting of residues involved in the ligand binding pocket] that pass through node i for every MD snapshot considered. The nodes with high node betweenness (NB) represent residues that are most crucial to information flow in the network;[29] the ones with high LNB influence the transfer of information between the chosen modules.

Edge Betweenness (EB) and Closeness Index (CI). EB is the total number of SPs that pass through an edge. Similar to NB, EB represents edges that control information transmission in the network.[19] CI, on the other hand, is defined as the average weighted distance of a node from all others in the network and is a measure of the centrality of the node in terms of its access to information.[29]

Funneling. The concept of "funneling" has been previously used in analyses of social networks.[29] According to this concept, the communication from a particular node to the rest of the network is mediated by a few mediator/collaborator nodes that have strong ties (i.e., interactions) with the node of interest. Mediator nodes are ranked on the basis of the number of SPs from a particular node to all others nodes in the network passing through them. In this study, the frequency of occurrence of a mediator node in the list of top 20 interacting partners (F_{med}) for each node in the network is calculated from $PcEN_{avg}$ (schematically explained in Figure S18 of the Supporting Information). The nodes with an F_{med} of >10% are identified as important for global communication in a system.

Cliques and Communities: Region of Higher Connectivity. In general network terminology, the parameters cliques and communities represent highly connected regions of the network. In the context of PcEN, these parameters are used to identify the rigid regions in the protein structures and to recognize the ligand-induced conformational changes.[17,30] A detailed description of cliques and communities is given elsewhere[31] (also in the Supplemental Methods of the Supporting Information). A k clique is defined as a set of k nodes (points represented by amino acids) in which each node is connected to all the other nodes. A community is defined as a union of smaller k cliques that share a node or nodes.[31] The cliques and communities for Sys1−3 are evaluated from the $PcEN_{avg}$ with an interaction energy cutoff of 2 kcal/mol using Cfinder.[32]

RESULTS

DhPylRS belongs to class IIc of aminoacyl-tRNA synthetases with a class II catalytic domain fold. The signal to release aa-AMP upon binding of cognate tRNA is mediated through interactions with nucleotides known as identity elements (IEs). Experimental studies of DhPylRS identified the discriminator base (G73), the first base pair in the acceptor stem (G1-G72), D-stem base pairs (G10-C25 and A11-U24), and G9 as the identity elements.[5] The $tRNA^{Pyl}$ structure is unique with a tightly packed tertiary core that is recognized by a core-binding surface (CBS) in DhPylRS (composed of tRNA binding domain 1, C-terminal tail, and helix α6 from the opposing protomer) (Figure S1 of the Supporting Information). The residues in CBS and the IEs of the DhPylRS complex offer novel complementary interaction sites with $tRNA^{Pyl}$ to ensure orthogonality.[6] Results presented in the following sections are based on parameters like energetically coupled dynamic correlations between distal sites from MD simulations and energy-weighted network parameters, a detailed description of which is given in Methods.

Allosteric Tertiary Coupling: Cross-Talk across Distant Sites. The application of the interaction correlation method (developed by the Karplus group) to identify energetically coupled regions in DhPylRS provides information about long-range signal transfer.[27] The interaction correlation is translated to residue-based correlation (see Methods) for ease of mapping back the results on the DhPylRS structure. Strikingly, the energy-based

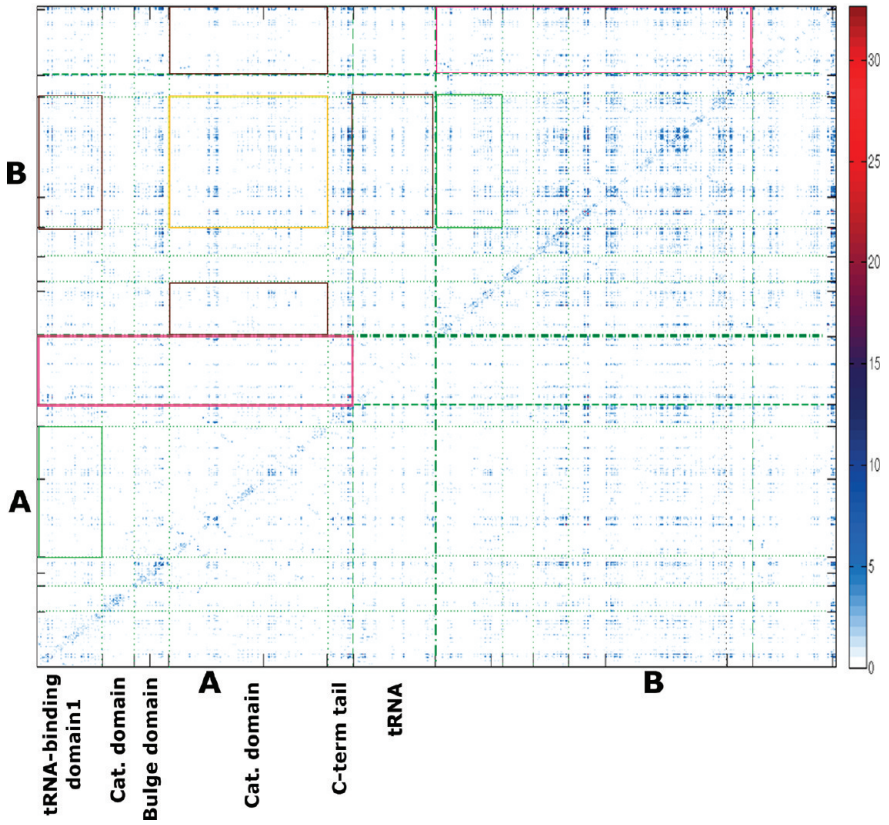


Figure 2. Energy-based residue correlation map for Sys3. Quadrants AA and BB represent intrasubunit correlations, whereas quadrants AB and BA represent intersubunit correlations. Regions AA and BB show marked asymmetry. Energetic coupling between the active site and the tRNA binding domains and/or nucleotides within and across subunits and between the two active sites is highlighted by colored rectangles.

residue correlation maps for Sys1−3 clearly exhibit asymmetry in spite of the symmetric nature of the starting structures (Table S1 of the Supporting Information), with stronger dynamic exchange of information in one subunit as compared to another (Figure 2 and Figure S2 of the Supporting Information), and this asymmetry is maximized in the pretransfer complex (Sys3) (Figure 2 and Figure S15 of the Supporting Information). Efficient cross-talk between the active site and the tRNA-binding residues and/or tRNA nucleotides within and across subunits and between the two active sites is evident from the correlation map for Sys3 (Figure 2). Energetic coupling of the two catalytic sites in Sys1−3 of the DhPylRS dimer (Figure 2 and Figure S2 of the Supporting Information), to various extents, indicates an allosteric communication between them, thus making ligand (Pyl-AMP) binding a cooperative process. This implies that the state of oligomerization of the protein precedes cooperativity in function, in an asymmetric manner.

The cumulative correlation factors of each residue (amino acids and nucleotides) with the ligand binding pockets [binding pocket correlation factor (BPCF)] can provide valuable information. BPCF values of Sys1−3 are compared (Figure 3 and Figure S3 of the Supporting Information) to relate the effect of perturbation (ligand binding in our case) to energetic rearrangements. When ligand binds, the local strain is permeated throughout the network and the residues with a high BPCF with respect to a particular subunit are the ones that act as efficient mediators to disperse this strain. Figure 3 clearly shows that in Sys1−3, BPCF values for most of the residues in the protein dimer are preferentially greater with respect to the binding pocket of one subunit over the other. This increases the possibility of a more efficient remote signal transmission toward a particular catalytic site resulting in increased asymmetry of communication. However, in Sys3, intrasubunit remote coupling values are stronger than the intersubunit ones toward the ligand binding pocket of subunit B (Figure 3 and Table S4 of the Supporting Information). A reduction in overall remote coupling is observed in Sys2, which is restored in Sys3 (Figures 2 and 3). This alternation in long-range coupling can act as a potential signal of binding of a particular ligand. Nucleotides in both tRNAs also show stronger coupling to the binding pocket of one subunit (Figure S4). Interestingly, BPCF values for the CUA tRNA anticodon with respect to both subunits are zero. This observation is consistent with the mutation experiments[5] that have shown that the CUA motif is not important for aminoacylation and the recognition of tRNA by a patch of surface residues identifying moderate identity elements in DhPylRS.[5,6] It was also proposed previously that the communication between distant sites is controlled by a set of residues with high BPCF values.[27] Thus, the pathways of signal transmission seem to be favored along one subunit in DhPylRS with a higher affinity of residues for the binding pocket, especially in the pretransfer complex, Sys3 (Figure 3). The asymmetry noted here in the context of communication is analogous to the concept of half-sites reactivity in multimeric proteins (even in structurally symmetric ones[22]) reported on the basis of kinetic and thermodynamic experiments.

The potential of a residue to influence the network is evaluated by considering all interactions/correlations of that residue [cumulative interaction energy (E_i)/correlation (C_i)], the capacity of which may change upon ligand binding. The differences in

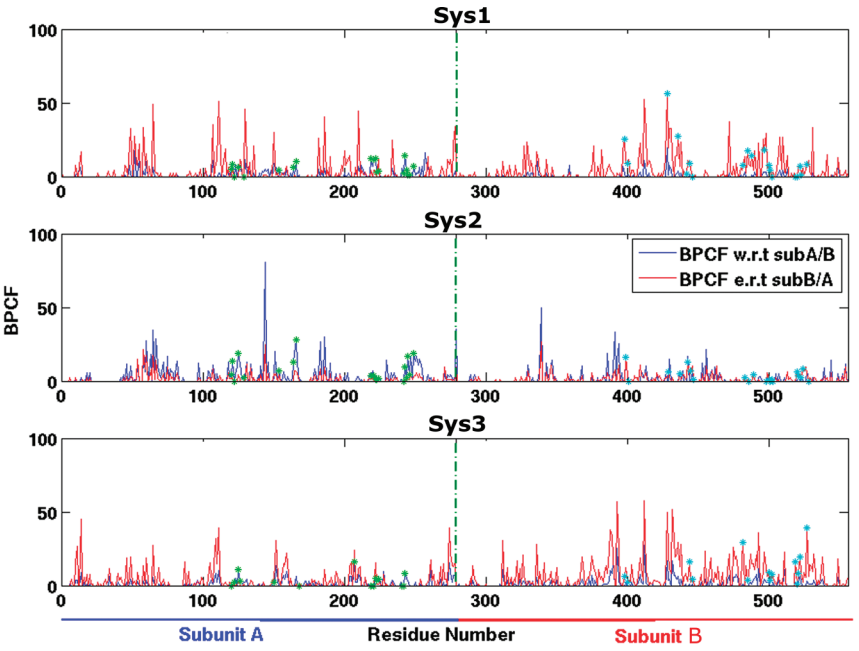


Figure 3. Correlation factor of amino acid residues with the ligand binding pocket of both subunits (BPCF with respect to subunit A, blue; BPCF with respect to subunit B, red) for Sys1−3. The residues involved in the ligand binding pockets of subunits A and B are highlighted with green and cyan asterisks, respectively. The average BPCF, standard deviation, and maximum value of BPCF of the residues in subunits A and B toward both the ligand binding pockets are listed in Table S4 of the Supporting Information.

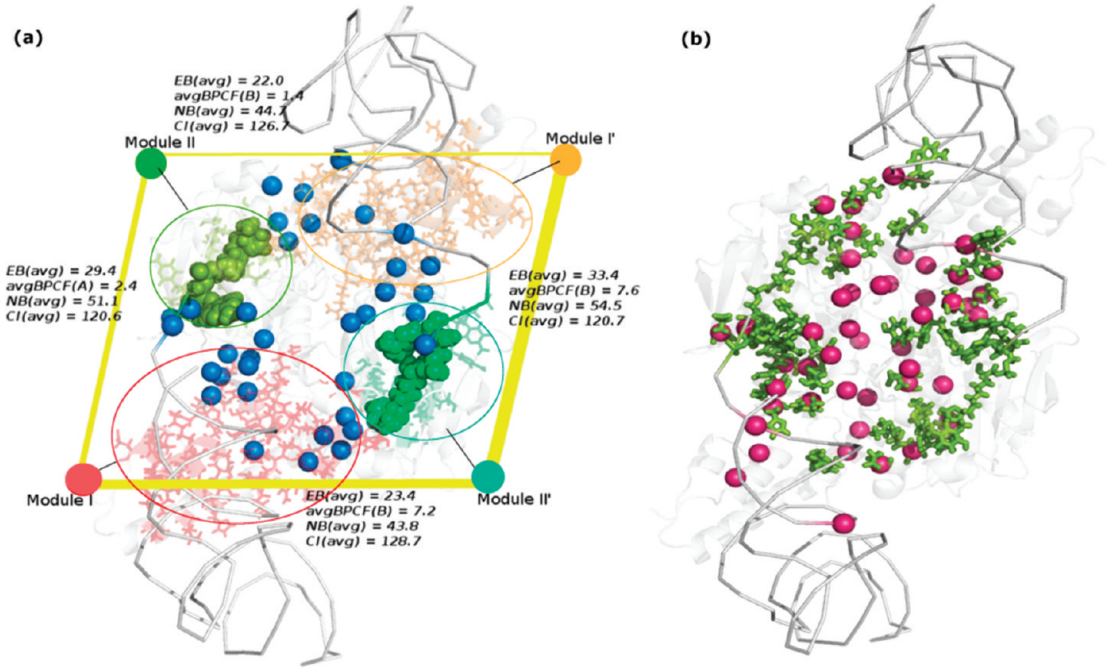


Figure 4. (a) Key residues in communication across functional modules in Sys3 (blue spheres). The four functional modules, I, I′, II, and II′, are shown as light red, orange, green, and lime green sticks, respectively. The thickness of the yellow lines is proportional to the average values of EB, LNB, CI, and BPCF (see Methods) for every communication pathway. (b) Top funneling nodes in Sys3 depicted as deep pink spheres. Important nodes for communication across modules (as shown in panel a) are shown as green sticks. The protein and tRNA backbone are shown as white cartoons and ribbons, respectively. The ligand is highlighted as van der Waals spheres. A detailed residue-wise summary is given in Figures S6 and S8 of the Supporting Information.

E_i/C_i between different ligand-bound states (see Methods) quantitatively estimate the effect of ligand induced perturbation (details in Figure S5 of the Supporting Information). The binding of ligand(s) to the native DhPylRS causes changes in

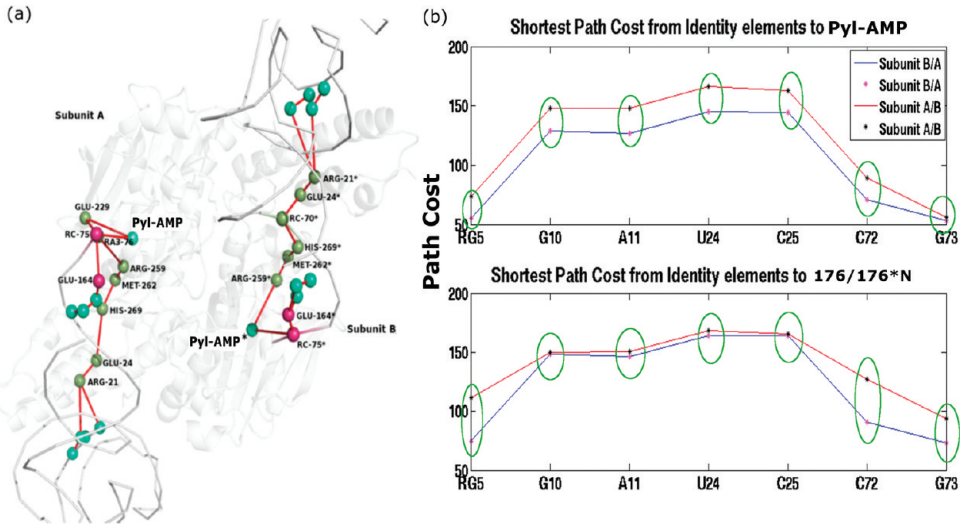


Figure 5. (a) Shortest paths (SP) of communication from seven identity elements to the ligand Pyl-AMP (cyan spheres) in subunits A and B of Sys3. Most of the paths are degenerate with two unique paths (shown by green and deep pink spheres and labeled). (b) Plot of the cost of communication for these seven SPs for both subunits.

both E_i and C_i of the binding pocket residues of both subunits, however, in an asymmetric manner. The tRNA binding region also shows effects of ligand (Pyl-AMP) binding, even before tRNA is bound to the system (in Sys2), showing that the tRNA binding sites are allosterically coupled to the catalytic sites.

Social Structure in Protein−tRNA Complex Energy Networks (PcEN): Key Residues in Communication. Terminologies like betweenness and funneling are known to convey the importance of different nodes (actor) and edges (ties) for communication in social networks.[29] Here we have adopted these concepts to identify key residues and connections for communication in Sys1−3 of DhPylRS using PcENs, which effectively portrays the interaction energy-weighted global connectivity and its rearrangements from a molecular perspective. We use parameters like node betweenness, edge betweenness, closeness index, and the concept of funnelling to inspect the structural framework of the PcENs and identify the residues important for communication across distant sites in the protein.

Communication between Distant Functional Sites. A pair of functional modules composed of tRNA recognition residues [CBS (module I/I′)] and ligand binding pockets (module II/II′) from each subunit are chosen as termini for the study of allosteric communication, both within and across the subunits. The weighted shortest paths (SP) between all the nodes in any pair of these modules are computed, and the frequency of these SPs passing through each node [local node betweenness (LNB)] is determined for the MD ensemble. LNB is a measure of a node's control over the transfer of information between the module pairs. The nodes with an LNB of >20% and also appearing in greater than 20% of MD snapshots are identified as being important for regulating communication between the module pairs (both intra- and intersubunit) for Sys1−3 (Figure 4a and Figure S6 of the Supporting Information). The total numbers of such important nodes in Sys1−3 are 24, 44, and 42, respectively (Figure S6 of the Supporting Information), showing the participation of a larger number of residues with higher transmission capacity in Sys2 and Sys3 as compared to Sys1 (Figures S6 and S7 of the Supporting Information). Similarly, the frequency of SPs passing through an edge [edge betweenness (EB)] is a measure of its capacity to transmit information, and the closeness index (CI) of a node indicates its accessibility within the network. High average values of edge betweenness and BPCF and low average values of the closeness index are known to represent better signal transmission. From the evaluated values in this case, we find a striking asymmetry in the efficacy of transmission between the two subunits of Sys3 (Figure 4a).

Global Communication between Distant Sites. The concept of funneling in networks posits that it is only a few mediator nodes (see Methods) that are central to global communication. We have used this concept to identify the important mediator nodes, if any, in our PcENs. We indeed find that a large fraction of SPs from a node to all other nodes pass through a small number of immediate mediator nodes (we have considered the top 20 mediator nodes connecting each node to all other nodes). We calculate the frequency of such top mediator (F_{med}) nodes considering the entire set of nodes in the graph and identify the important residues (F_{med} > 10%) for global communication for Sys1−3 (Figure S8 of the Supporting Information). Most of the residues identified as being important for the transfer of the signal between the chosen modules (by LNB) coincide with the ones important for global communication through funneling (Figure 4b and Figure S8 of the Supporting Information). A comparison of the important nodes (global and local) (Figures S6 and S8 of the Supporting Information) among Sys1−3 reveals variations upon ligand binding, implying that the potential to be an important node is dictated to some degree by the environment.

Cost of Communication and Relative Fitness of One Subunit. PcEN offers an excellent means of evaluating the cost of communication through interaction energies of the edges constituting the SPs. This allows us to investigate the structural reorganizations in terms of the changes in cost upon binding of different ligands and hence provides a method for correlating the structural and network changes with function. The SPs identified are the minimal cost paths for efficient information transfer in each subunit. The lower the cost, the easier the transmission along the SP. A recently determined crystal structure of the DhPylRS−$tRNA^{Pyl}$ complex[6] reveals that unique interactions between the dimeric DhPylRS and $tRNA^{Pyl}$ are conferred by the

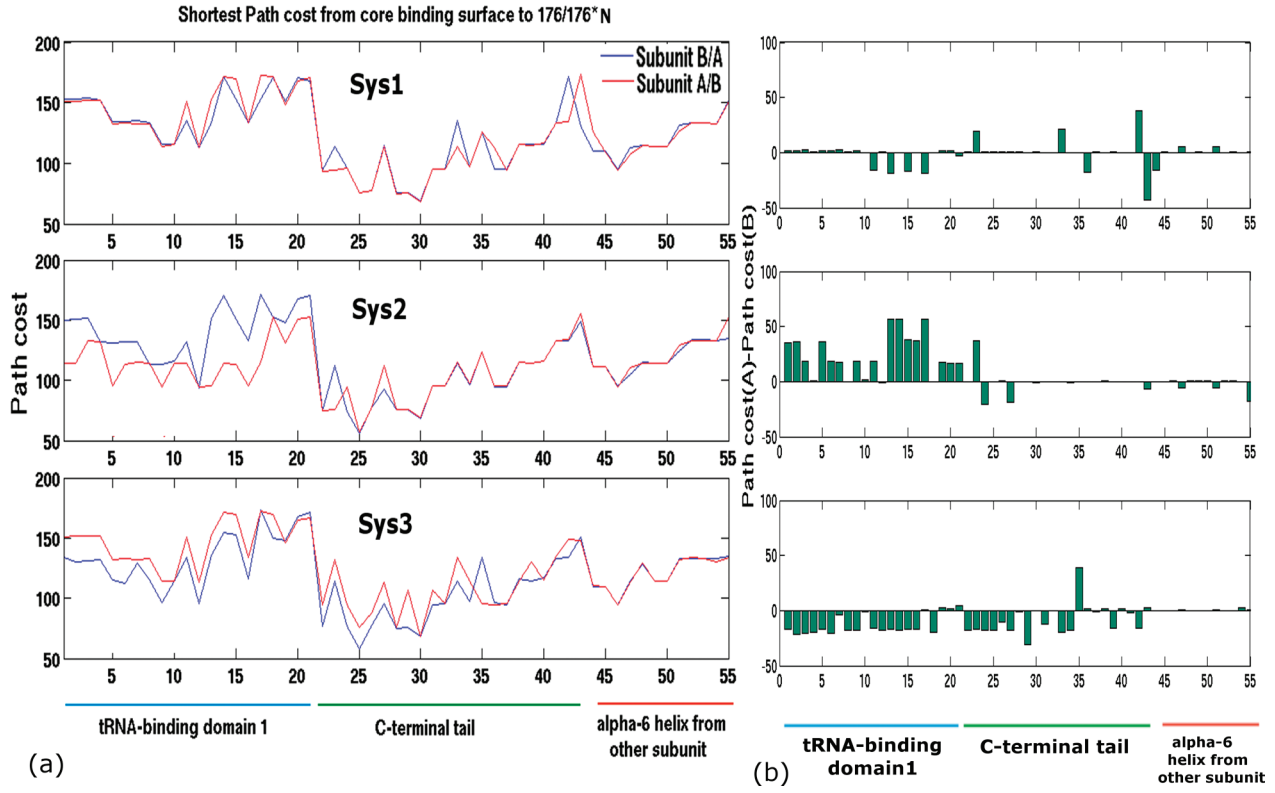


Figure 6. (a) Plot of SP costs from residues in CBS to N176N/N176* (dynamically H-bonded to Pyl-AMP) for both subunits of Sys1−3. (b) Plot of the path cost difference [path cost(A) − path cost(B)] between subunits A and B. It is clearly evident that there is some asymmetry in path cost between the two subunits of Sys1−3, especially in Sys2 and Sys3.

core-binding surface (CBS) of the protein and identity elements (IE) of tRNA. We therefore measure the cost of SPs between the amino acid residues and nucleotides in the CBS and IE, respectively, and selected residues in the ligand binding pocket in both the subunits to capture the unique communications in DhPylRS. The cognate signal from these sites instructs the release of Pyl-AMP from the ligand binding pocket. A residue, N176, forming a dynamically stable H-bond (>90% of the snapshots) with the ligand (Pyl-AMP) and the ligand residue itself (for Sys2 and Sys3) are chosen as the ligand binding pocket terminus in each case. It is seen that the cost of communication between each pair of chosen termini is lower in one subunit compared to the other, with very few exceptions (Figures 5 and 6 and Figure S9 of the Supporting Information). The native protein (Sys1) shows minimal asymmetry, in contrast to Sys2 and Sys3 in which prominent asymmetry is observed in terms of the cost of communication (Figure 6a,b).

The optimal path (OP) (i.e., SP) of communication can be flanked by a large number of suboptimal paths (SOP) with reduced efficiency (higher cost). SOPs provide alternate routes of signal transfer between distant sites. Identification of SOPS is important because the system may adapt these paths under nonoptimal conditions. Computationally, we identify the first-generation SOPs by systematically removing all interactions of an OP node with the rest of the network, thus forcing the detection of the alternate path (interestingly, this concept emulates biological single-point mutations). The increase in the cost of communication along the SOPs with respect to the OP (Δcost) is computed for both subunits of Sys1−3 with residues R21, Q47, E50, and N276 (forming dynamically stable H-bond with identity elements) and N176 as the two end termini. Low Δcost values indicate the existence of SOPs with nearly equivalent costs (i.e., nearly equivalent efficacy). The average Δcost values for each subunit in Sys1 and Sys2 are nearly equal. However, Sys3 is strikingly exceptional with the average Δcost values in the two subunits being highly asymmetric (Figure 7a and Figure S10 of the Supporting Information). The subunit with a low Δcost value has nearly optimal alternate routes (SOPs) and thus seems to be bestowed with increased resilience toward single-point mutations as compared to the other subunit. Incidentally, this comparatively robust subunit in the pretransfer complex is the one with the lower OP cost and stronger long-range coupling. However, upon removal of a pair of OP nodes (second-generation SOPs, mimicking double mutations) for this subunit, Δcost values are as high as those upon removal of single nodes from the other subunit (schematically shown in Figure 7b). This is also in agreement with the concept of half-sites reactivity, with one subunit being more potent than the other in terms of communication.

Preexistence of SOPs of Communication. An ensemble of conformations, generated from MD simulations is classified in terms of their OPs between biologically relevant distal sites as mentioned above. In all three systems (the numbers of pairs of distal sites considered from N176 are 8 in Sys1 and Sys2 and 18 in Sys3), we observe that a majority of snapshots take an OP that is more cost-effective than the ones taken up by smaller numbers of snapshots (Table S2 of the Supporting Information). Here we make an interesting observation that a majority of the OPs evaluated between different pairs of nodes from the average network ($PcEN_{avg}$) are from the major populations in the MD ensemble and the OPs in the less frequently accessed conformations have

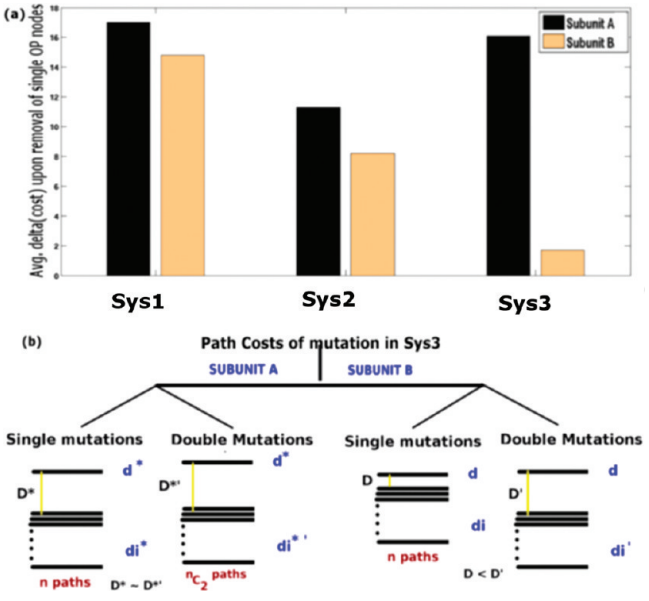


Figure 7. (a) Average difference in the cost of communication (Δcost) between optimal and first-generation suboptimal paths (termini chosen being R21, Q47, E50, and N276 (tRNA binding region) and N176 (ligand pocket) for Sys1−3 (also see Figure S10 of the Supporting Information). (b) Schematic representation of the trends in Δcost values for subunits A and B in Sys3. In subunit A, both first- and second-generation (i.e., single and double mutations of residues on OP) SOPs have costs much greater than that of the OP. In subunit B, the Δcost between the first-generation SOPs and OPs is negligible whereas the second-generation SOPs have high Δcost values.

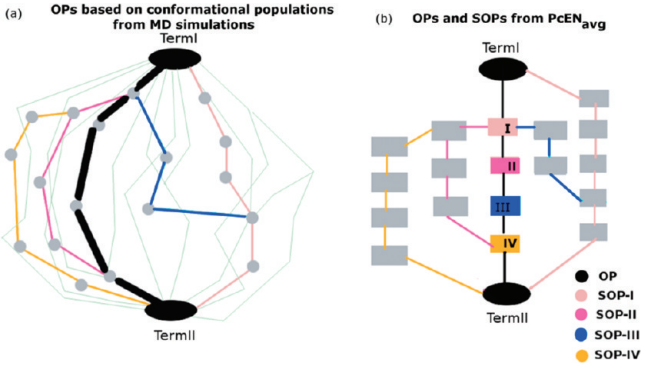


Figure 8. Schematic representation of the unique sets of OPs (between term I and term II). (a) OPs from different conformational populations of the MD ensemble with thick black and fine light green lines representing major and minor populations, respectively. (b) OPs and SOPs between term I and term II from PcEN$_{avg}$. The black line represents the OP. Colored lines depict first-generation SOPs (lines color-coded on the basis of the color of the node removed from OP). Most of these SOPs occur as OPs in a small percentage of the MD ensemble as indicated by thin lines of the same color in panel a.

indeed been captured as SOPs from the knockout exercise on PcEN$_{avg}$ (Table S2 of the Supporting Information). This relation is schematically represented in panels a and b of Figure 8, where the OPs from MD snapshots and OP and SOPs from PcEN$_{avg}$ are shown, respectively. Although the OPs and SOPs are different in three systems, we observe several common pairwise interactions along the OPs and SOPs (Table S2 of the Supporting Information). The observed correlation between the OPs of conformational populations and the SOPs reveals a preexistence of most of these SOPs as OPs in a small percentage of the conformational population. This indicates that a mutation (or ligand induced perturbation) in OP in PcEN$_{avg}$ forces the transmission through a preexisting path by merely shifting the ensemble toward a less accessed conformational population. In a way, this result indicates that the population shift model of allostery[9] and the preexistence of OPs of communication[10] are interlinked through global connectivity and their reorganization upon perturbation.

Alternating Rigidity and Flexibility: Mechanistic Insights into aaRS-Assisted Charging of tRNA. While we discussed extensively the intricate possible mechanism of communication and paths in the preceding sections, here we shift our attention to ligand induced global changes in rigid and flexible regions of the protein complex and its implications for function. The ligand induced rewiring in PcENs is beautifully captured in terms of network parameters like cliques and communities.[33] The cliques and communities for Sys1−3 are evaluated (Table S3 of the Supporting Information). Sys1, with small disjointed communities, is proposed to be flexible facilitating the ligand approach. Sys2 is highly rigid, with large interface communities being formed around the ligand Pyl-AMP in one of the subunits. The significance of this large community may be in protecting the Pyl-AMP from deactivation. However, when the cognate tRNA binds (Sys3), a marked decrease in the size of the community around Pyl-AMP occurs in one of the subunits (Figure S11 of the Supporting Information). This decrease in rigidity in the immediate vicinity of the active site may help in the charging process by allowing an aperture for the cognate tRNA's 3′ end. On the other hand, we find an increased number of communities at the protein−tRNA interface locking the tRNA position tightly by DhPylRS to facilitate thorough recognition for the maintenance of orthogonality and for efficient aminoacylation.

Additionally, the strength of interactions also manifests at the protein−tRNA recognition level. It may be pointed out that some of the identity elements, mainly interacting with residues in the CBS, are part of protein−tRNA interface cliques (Table S3 of the Supporting Information). However, the interaction between the CUA triplet anticodon and DhPylRS is negligible and hence does not even qualify for making a connection with the protein. Thus, the experimentally observed protein−tRNA recognition is also manifested in terms of the strength of interaction and the rigidity imparted through cliques and communities.

DISCUSSION

The results presented above clearly demonstrate that the energy-weighted representation of the structure network and the incorporation of dynamic information from MD simulation is a powerful approach for capturing the subtle conformational rearrangements upon ligand binding. Such an investigation can enhance our understanding of basic biochemical principles such as allosteric communication and half-sites reactivity. We have analyzed three systems of DhPylRS [native enzyme (Sys1) and that bound to Pyl-AMP (Sys2) and Pyl-AMP and tRNA^{Pyl} (Sys3)]. The asymmetry between the two subunits (most prominent in Sys3) is evident from all our results, with indications of differential efficacies for long-range communication and dynamic coupling along one subunit as compared to another. Various weighted graph parameters, both at the local and at the global communication level, vividly portray the social structure of the

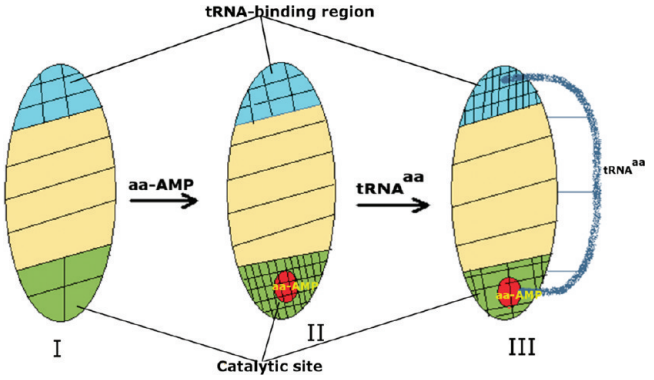


Figure 9. General mechanistic scheme for the aaRS-assisted aminoacylation reaction. State I is the native protein, state II the native protein bound to aa-AMP, and state III the pretransfer complex (native protein bound to aa-AMP and tRNA). The grids indicate rigidity in a region of the structure, grid density being proportional to rigidity. A clear dramatic increase in rigidity around the ligand is evident upon going from state I to state II followed by a slight decrease in state III. The tRNA binding site also undergoes a slight increase in rigidity from state I to state II, which is drastically increased for state III.

energy-weighted networks for Sys1−3 of DhPylRS. Thus, in social network jargon, the protein has assigned a given set of nodes dedicated to information transfer like the "sociometric superstars" in Milgram's experiment.[34] In this study, we have identified a few such important nodes, the mutations of which are expected to affect communication to different extents.

The weighted formalism has provided us with a tool for quantifying the efficacy of long-range communication in terms of costs of communication along the optimal and suboptimal paths. To our surprise, we indeed observe asymmetry in the efficacy of remote signal transfer for the two subunits of a structurally symmetric enzyme, DhPylRS (Table S1 of the Supporting Information). The relative robustness of one subunit in the dimeric pretransfer complex may be the result of an evolutionary optimization of structure for function, in terms of "less but best". Intriguingly, the communication cost is lower (enhanced efficiency) and robustness is higher for the subunit that exhibits higher residue-wise dynamic correlation in a given system, suggesting that energetic long-range coupling complements communication. Furthermore, our analysis also substantiates the concept of preexisting paths of communication in a dynamic ensemble and their reorganizations upon ligand binding and/or environmental variations. The concept of half-sites reactivity or asymmetry in function has been discussed extensively in the literature.[35−37] Specifically in the context of aaRS, Fersht had stated, "these enzymes exhibit negative cooperativity of substrate binding and half-of-the-sites reactivity".[38] Also, it has been tested in several aaRS by rigorous experiments[35,39,40] and also elucidated in TrpRS by a network approach.[17] Here we predict this phenomenon in DhPylRS at a detailed molecular level (including dynamics and energetics), and this computationally derived result would warrant experimental verification.

Additionally, the analysis of the ligand induced global changes in rigid and flexible regions of the protein complex provides general mechanistic insight into the aminoacylation reaction. The aminoacylation process is proposed to be optimized by a deft modulation of rigidity and flexibility at different sites by appropriate ligands. Our previous studies of MetRS[41] and TrpRS[17] also produced similar results in terms of the network parameters, cliques, and communities, which define the higher order connectivity. Thus, a general mechanistic scheme is proposed for aaRS (Figure 9). According to this scheme, the native aaRS exhibits flexibility around the active site to facilitate the approach of ligand(s). Once the activated amino acid is bound, it is protected by the formation of large rigid communities around the ligand. Also, the tRNA binding site shows a slight increase in the number of communities that is further augmented upon tRNA binding. When the cognate tRNA makes its approach, the signal to release the aa-AMP is interpreted by reducing the rigidity around the active site, thus allowing the entry of the 3′ end of cognate tRNA and efficient aminoacylation.

In summary, MD simulations and the analysis of the trajectories through energy-weighted structure networks reveal prominent structural and functional asymmetry between the two subunits in dimeric DhPylRS with the pretransfer complex (bound to tRNA and Pyl-AMP) being maximally asymmetric. One of the subunits of DhPylRS is favored for function, with increased communication efficiency and greater resilience to single-point mutations of optimal path residues. This is in agreement with the concept of half-sites reactivity discussed in the literature. Strikingly, both allosteric communication and half-sites reactivity seem to be guided by alterations in the global network features through a clever rearrangement of pairwise interactions, imparting enhanced functional efficiency to one subunit. Key residues for communication are identified using several weighted network parameters and related concepts. Additionally, we find interdependence between long-range energetic coupling and communication. Our study reveals the preexistence of alternate communication paths in the MD ensemble. A general mechanism of aaRS-assisted charging of cognate tRNA, involving alternation of the rigidity and flexibility at definite functional sites, is also proposed. Thus, a subtle rearrangement of side chains upon binding of different ligands may take place even in the absence of drastic conformational changes at the backbone level. This implies that subtle conformational variations may indeed hold the key to the precise functioning of proteins, and it would be interesting to explore structure−function correlations of other proteins at such a refined level. In particular, the methodology outlined in this study is general and can be used to probe allosteric communications and ligand induced structural rearrangements in proteins as well as protein−protein/RNA/DNA complexes.

■ ASSOCIATED CONTENT

Supporting Information. Supporting figures (Figures S1−S18), supporting tables (Tables S1−S5), and Supplemental Methods. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +91-80-22932611. Fax: +91-80-23600535. E-mail: sv@mbu.iisc.ernet.in.

Notes

We acknowledge support from DST for Mathematical Biology (DSTO773), Government of India. M.B. thanks CSIR for a fellowship.

■ ACKNOWLEDGMENT

We are grateful to the computational facilities at the Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, India.

■ ABBREVIATIONS

DhPylRS, *D. hafniense* pyrrolysyl-tRNA synthetase; MmPylRS, *Methanosarcina mazei* pyrrolysyl-tRNA synthetase; Pyl-AMP, pyrrolysyl-AMP; PcEN, protein complex energy network; CBS, core-binding surface; IE, identity elements; OP and SOP, optimal and suboptimal paths, respectively; SP, shortest path.

■ REFERENCES

(1) Ling, J., Reynolds, N., and Ibba, M. (2009) Aminoacyl-tRNA Synthesis and Translational Quality Control. *Annu. Rev. Microbiol. 63*, 61–78.

(2) Krzycki, J. A. (2005) The direct genetic encoding of pyrrolysine. *Curr. Opin. Microbiol. 8*, 706–712.

(3) Kobayashi, T., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2009) Recognition of Non-α-amino Substrates by Pyrrolysyl-tRNA Synthetase. *J. Mol. Biol. 385*, 1352–1360.

(4) Yanagisawa, T., Ishii, R., Fukunaga, R., Kobayashi, T., Sakamoto, K., and Yokoyama, S. (2008) Crystallographic Studies on Multiple Conformational States of Active-site Loops in Pyrrolysyl-tRNA Synthetase. *J. Mol. Biol. 378*, 634–652.

(5) Herring, S., Ambrogelly, A., Polycarpo, C. R., and Söll, D. (2007) Recognition of pyrrolysine tRNA by the *Desulfitobacterium hafniense* pyrrolysyl-tRNA synthetase. *Nucleic Acids Res. 35*, 1270–1278.

(6) Nozawa, K., O'Donoghue, P., Gundllapalli, S., Araiso, Y., Ishitani, R., Umehara, T., Söll, D., and Nureki, O. (2009) Pyrrolysyl-tRNA synthetase-tRNAPyl structure reveals the molecular basis of orthogonality. *Nature 457*, 1163–1167.

(7) Monod, J., Wyman, J., and Changeux, J. P. (1965) On the nature of allosteric transitions: A plausible model. *J. Mol. Biol. 12*, 88–118.

(8) Koshland, D. E., Némethy, G., and Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry 5*, 365–385.

(9) Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins: Struct., Funct., Bioinf. 57*, 433–443.

(10) del Sol, A., Tsai, C. J., Ma, B., and Nussinov, R. (2009) The Origin of Allosteric Functional Modulation: Multiple Pre-existing Pathways. *Structure 17*, 1042–1050.

(11) Tsai, C. J., del Sol, A., and Nussinov, R. (2008) Allostery: Absence of a Change in Shape Does Not Imply that Allostery Is Not at Play. *J. Mol. Biol. 378*, 1–11.

(12) Gandhi, P. S., Chen, Z., Mathews, F. S., and Di Cera, E. (2008) Structural identification of the pathway of long-range communication in an allosteric enzyme. *Proc. Natl. Acad. Sci. U.S.A. 105*, 1832–1837.

(13) Peterson, E. S., and Friedman, J. M. (1998) A Possible Allosteric Communication Pathway Identified through a Resonance Raman Study of Four β37 Mutants of Human Hemoglobin A. *Biochemistry 37*, 4346–4357.

(14) Chen, J., Dima, R. I., and Thirumalai, D. (2007) Allosteric Communication in Dihydrofolate Reductase: Signaling Network and Pathways for Closed to Occluded Transition and Back. *J. Mol. Biol. 374*, 250–266.

(15) Lockless, S. W., and Ranganathan, R. (1999) Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science 286*, 295–299.

(16) Ghosh, A., and Vishveshwara, S. (2007) A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis. *Proc. Natl. Acad. Sci. U.S.A. 104*, 15711–15716.

(17) Bhattacharyya, M., Ghosh, A., Hansia, P., and Vishveshwara, S. (2010) Allostery and conformational free energy changes in human tryptophanyl-tRNA synthetase from essential dynamics and structure networks. *Proteins: Struct., Funct., Bioinf. 78*, 506–517.

(18) del Sol, A., Fujihashi, H., Amoros, D., and Nussinov, R. (2006) Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol. 2*, 0019.

(19) Sethi, A., Eargle, J., Black, A. A., and Luthey-Schulten, Z. (2009) Dynamical networks in tRNA:protein complexes. *Proc. Natl. Acad. Sci. U.S.A. 106*, 6620–6625.

(20) Chennubhotla, C., and Bahar, I. (2007) Signal Propagation in Proteins and Relation to Equilibrium Fluctuations. *PLoS Comput. Biol. 3*, e172.

(21) Hill, T. L., and Levitzki, A. (1980) Subunit Neighbor Interactions in Enzyme Kinetics: Half-Of-The-Sites Reactivity in a Dimer. *Proc. Natl. Acad. Sci. U.S.A. 77*, 5741–5745.

(22) Shi, J., Dertouzos, J., Gafni, A., Steel, D., and Palfey, B. A. (2006) Single-molecule kinetics reveals signatures of half-sites reactivity in dihydroorotate dehydrogenase A catalysis. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5775–5780.

(23) Kavran, J. M., Gundllapalli, S., O'Donoghue, P., Englert, M., Söll, D., and Steitz, T. A. (2007) Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc. Natl. Acad. Sci. U.S.A. 104*, 11268–11273.

(24) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER9*, University of California, San Fransisco.

(25) Cheatham, T. E., III, Cieplak, P., and Kollman, P. A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn. 16*, 845–862.

(26) Vijayabaskar, M. S., and Vishveshwara, S. (2010) Interaction Energy Based Protein Structure Networks. *Biophys. J. 99*, 3704–3715.

(27) Kong, Y., and Karplus, M. (2009) Signaling pathways of PDZ2 domain: A molecular dynamics interaction correlation analysis. *Proteins: Struct., Funct., Bioinf. 74*, 145–154.

(28) Kong, Y., and Karplus, M. (2007) The Signaling Pathway of Rhodopsin. *Structure 15*, 611–623.

(29) Newman, M. E. J. (2001) Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E 64*, 016132.

(30) Sukhwal, A., Bhattacharyya, M., and Vishveshwara, S. (2011) Network approach for capturing ligand induced subtle global changes in protein structures. *Acta Crystallogr. D67*, 429–439.

(31) Palla, G., Derényi, I., Farkas, I., and Vicsek, T. (2005) Uncovering the overlapping community structure of complex networks in nature and society. *Nature 435*, 814–818.

(32) Adamcsek, B., Palla, G., Farkas, I., Derényi, I., and Vicsek, T. (2006) CFinder: Locating cliques and overlapping modules in biological networks. *Bioinformatics 22*, 1021–1023.

(33) Bhattacharyya, M., and Vishveshwara, S. (2010) Elucidation of the conformational free energy landscape in *H. pylori* LuxS and its implications to catalysis. *BMC Struct. Biol. 10*, 27.

(34) Milgram, S. (1967) The Small World Problem. *Psychol. Today 2*, 60–67.

(35) Huang, C. Y., Rhee, S. G., and Chock, P. B. (1982) Subunit Cooperation and Enzymatic Catalysis. *Annu. Rev. Biochem. 51*, 935–971.

(36) Bloom, C. R., Kaarsholm, N. C., Ha, J., and Dunn, M. F. (1997) Half-Site Reactivity, Negative Cooperativity, and Positive Cooperativity: Quantitative Considerations of a Plausible Model. *Biochemistry 36*, 12759–12765.

(37) Brzovic, P. S., Choi, W. E., Borchardt, D., Kaarsholm, N. C., and Dunn, M. F. (1994) Structural Asymmetry and Half-Site Reactivity in the T to R Allosteric Transition of the Insulin Hexamer. *Biochemistry 33*, 13057–13069.

(38) Fersht, A. R. (1975) Demonstration of two active sites on a monomeric aminoacyl-tRNA synthetase. Possible roles of negative cooperativity and half-of-the-sites reactivity in oligomeric enzymes. *Biochemistry 14*, 5–12.

(39) Trezeguet, V., Merle, M., Gandar, J. C., and Labouesse, B. (1986) Kinetic evidence for half-of-the-sites reactivity in tRNATrp aminoacylation by tryptophanyl-tRNA synthetase from beef pancreas. *Biochemistry 25*, 7125–7136.

(40) Ambrogelly, A., Kamtekar, S., Stathopoulos, C., Kennedy, D., and Söll, D. (2005) Asymmetric behavior of archaeal prolyl-tRNA synthetase. *FEBS Lett. 579*, 6017–6022.

(41) Ghosh, A., and Vishveshwara, S. (2008) Variations in Clique and Community Patterns in Protein Structures during Allosteric Communication: Investigation of Dynamically Equilibrated Structures of Methionyl tRNA Synthetase Complexes. *Biochemistry 47*, 11398–11407.