A study of communication pathways in methionyl-tRNA synthetase by molecular dynamics simulations and structure network analysis

Amit Ghosh and Saraswathi Vishveshwara*

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

Edited by John Kuriyan, University of California, Berkeley, CA, and approved August 21, 2007 (received for review May 15, 2007)

The enzymes of the family of tRNA synthetases perform their functions with high precision by synchronously recognizing the anticodon region and the aminoacylation region, which are separated by ≈70 Å in space. This precision in function is brought about by establishing good communication paths between the two regions. We have modeled the structure of the complex consisting of *Escherichia coli* methionyl-tRNA synthetase (MetRS), tRNA, and the activated methionine. Molecular dynamics simulations have been performed on the modeled structure to obtain the equilibrated structure of the complex and the cross-correlations between the residues in MetRS have been evaluated. Furthermore, the network analysis on these simulated structures has been carried out to elucidate the paths of communication between the activation site and the anticodon recognition site. This study has provided the detailed paths of communication, which are consistent with experimental results. Similar studies also have been carried out on the complexes (MetRS + activated methonine) and (MetRS + tRNA) along with ligand-free native enzyme. A comparison of the paths derived from the four simulations clearly has shown that the communication path is strongly correlated and unique to the enzyme complex, which is bound to both the tRNA and the activated methionine. The details of the method of our investigation and the biological implications of the results are presented in this article. The method developed here also could be used to investigate any protein system where the function takes place through long-distance communication.

dynamic cross-correlations | methionyl-AMP | protein structure network | shortest pathways of communication | stacking

A crucial step in the translation of the genetic code is the aminoacylation of tRNA, which involves the molecular recognition between the aminoacyl-tRNA synthetases (aaRS) and their cognate tRNA. Each synthetase consists of the catalytic domain and the anticodon domain that are separated by ≈70 Å. Each tRNA connects these two regions with its anticodon and the acceptor stems. The mechanism of differentiations between cognate and noncognate tRNAs depends on contacts of anticodon domain of synthetase and anticodon stem of tRNA. The efficiency of the selection mechanism controls the overall accuracy of protein synthesis (1, 2). Recognition of the protein (aaRS) and the tRNA is explained by using the induced-fit mechanism, which suggests conformational changes in protein, tRNA, or both, leading to the final bound complex (3). However, the details of communication between the anticodon region and the aminoacylation region are less understood.

In all living cells, protein synthesis starts with methionine. Methionyl-tRNA synthetase (MetRS) binds to free methionine in its catalytic site, and the charging of amino acid takes place in the presence of ATP, which then is finally transferred to 2′-OH moiety of the terminal adenosine of tRNA. MetRS acylates two tRNAs ($tRNA_f^{Met}$ and $tRNA_m^{Met}$), which have different nucleotide sequences. However, their common anticodon (CAU) is responsible for discriminating other noncognate tRNAs against cognate ones. The C-terminal α-helical domain (anticodon domain) of MetRS is responsible for the recognition of three anticodon bases of $tRNA^{Met}$ (4). Extensive experiments have been carried out to establish the importance of several amino acid residues in MetRS and the nucleotides of tRNA in the recognition processes. For instance, the loop containing highly conserved residue Trp-461 plays a crucial role in the recognition of anticodon bases, as confirmed by affinity-labeling experiments (5), isolation of second site mutants (6), and docking studies of glutaminyl-tRNA on MetRS (7). Substitution of highly conserved residues Trp-461, Asn-452, or Arg-395 shows their involvement in the binding of MetRS to anticodon site of tRNA (8–10). Two negatively charged residues, Asp-449 and Asp-456, also contribute to the overall specificity of the anticodon recognition, although they do not make direct contact with tRNA (11). Because of their acidic nature, these residues are able to reject all other noncognate tRNAs from binding to the anticodon sites of MetRS.

Various investigations have been carried out to understand the mode of communication between the anticodon and the activation site in aaRS. Interdomain signaling has long been recognized to be important in establishing contacts (12, 13). It also has been pointed out that the interface region is sensitive to structural perturbations, which affect the noncovalent packing interaction between the two functional domains (14, 15) in MetRS. The importance of covalent connectivity in domain–domain communication also has been investigated in CysRS (16). The idea of communication between the two sites through a network of residues, interacting in space has been put forward (12). Thus, it is a challenging task to identify such networks in aaRS. The identification of cooperative network of interactions between residues in the protein (U1A)–RNA complex has been carried out by NMR experiments (17, 18). More details of communications in this system have been obtained through molecular dynamics (MD) simulations (19, 20). This investigation has confirmed that the collective atomic fluctuations obtained from MD simulations agree very well with the experimentally observed cooperativity.

In this article, we have identified the communication paths between the anticodon binding region and the aminoacylation site in MetRS. Two approaches are combined to achieve this goal. First, the dynamically cross-correlated residues in MetRS have been identified from MD simulations by using a method similar to the one used in U1A–RNA complex (19, 20). Although this method is powerful in identifying the highly correlated residues involved in the

Author contributions: A.G. and S.V. designed research; A.G. performed research; A.G. and S.V. analyzed data; and A.G. and S.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: aaRS, aminoacyl-tRNA synthetase; MetRS, methionyl-tRNA synthetase; MD, molecular dynamics; MetAMP, methionyl-AMP; RMSD, root mean square deviation; DCCM, dynamic cross-correlation map; PSN, protein structure network; PSG, protein structure graphs.

*To whom correspondence may be addressed. E-mail: sv@mbu.iisc.ernet.in.

This article contains supporting information online at www.pnas.org/cgi/content/full/0704459104/DC1.

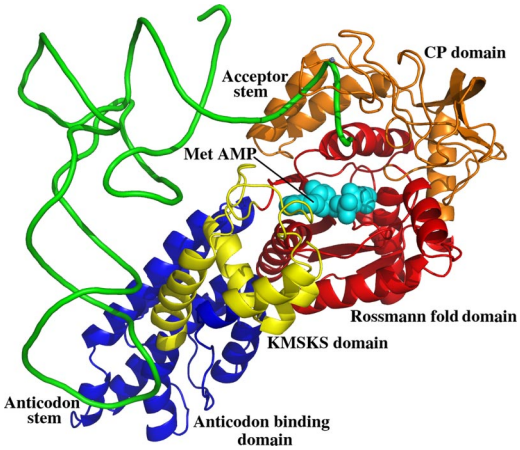

Fig. 1. The overall docked structure of *E. coli* MetRS–$tRNA_f^{Met}$ complex along with Met AMP.

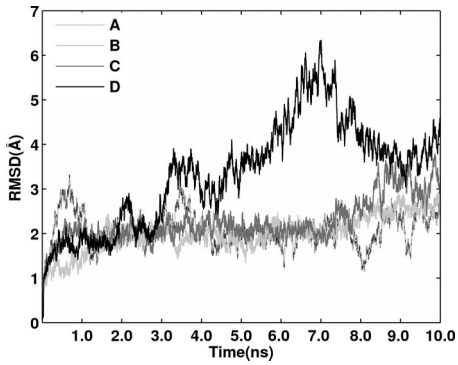

Fig. 2. MD trajectories of RMSD of the four systems (A, B, C, and D) of MeRS with reference to the minimized crystal structure. A, B, C and D in this figure and in subsequent figures correspond to the simulations on four systems: MetRS (A), MetRS-MetAMP (B), MetRS–$tRNA_f^{Met}$ (C), and MetRS–$tRNA_f^{Met}$–MetAMP (D).

communication path, the correlated residues are not necessarily connected in space; hence, the paths that are spatially connected cannot be detected by this method alone. In the present work, we bridge this gap by analyzing the protein structure graphs (PSG). This approach, developed in our laboratory (21–25), identifies the residues connected spatially by noncovalent interactions. The noncovalently interacting direct links, connecting the anticodon region and the aminoacylation region, have been identified by the analysis of the protein-structure graph of MetRS. In the final step, the shortest paths of connected networks, consistent with the dynamically cross-correlated set of residues, have been selected as the paths of communication between the anticodon and the aminoacylation regions. Such paths have been identified in the ligand-free form of *Escherichia coli* MetRS (system A) and also in three complexed structures of *E. coli* MetRS: with methionyl-AMP (MetAMP) (system B), with tRNA (system C), and with both MetAMP and tRNA (system D). A comparison of the paths in these four systems has shown that the set of communication paths are shortest and strongly correlated in MetRS bound to both MetAMP and tRNA (system D). Several of the residues identified in these paths have been shown by experiments to be important in recognition. The methodology followed in the identification of paths and the important results are presented in this article.

Results and Discussion

Modeling of the $tRNA_f^{Met}$-bound structure of *E. coli* MetRS. The crystal structure of the complex of *E. coli* MetRS with $tRNA^{Met}$ is not available, although the structures of $tRNA^{Met}$ (26) and MetRS (27, 28) are available. Hence, we modeled the structure of MetRS–$tRNA_f^{Met}$ complex by using the *Aquifex aeolicus* (29) as a template. A major difference between the structures of the bound form of *Aquifex* tRNA and the unbound form of *E. coli*–tRNA was found in the anticodon loop consisting of 11 bases. Therefore we replaced the conformation of *E. coli* tRNA in this region by that of the *Aquifex* tRNA in the complex, and some details of this modeled structure of *E. coli* tRNA are given in supporting information (SI) Fig. 7. The structure of *E. coli* MetRS–$tRNA^{Met}$ complex thus modeled is shown in Fig. 1.

The communication paths between the anticodon recognition site and the activated amino acid binding region (active site region) in *E. coli* MetRS has been identified by analyzing and comparing the MD trajectories of four systems (systems A, MetRS; B, MetRS bound to MetAMP; C, MetRS bound to tRNA; and D, MetRS bound to MetAMP and tRNA). The simulation results along with the analyses of the dynamic cross-correlations and the protein structure networks (PSN) are presented in the following sections.

Root Mean Square Deviation (RMSD) Profiles. MD simulations were carried out on the four systems A, B, C, and D in aqueous medium. The trajectories of RMSDs with respect to the minimized starting structure are shown in Fig. 2. The C_α RMSD values are within 4 Å throughout the simulation for the systems A, B, and C. However, it reaches a maximum value of 6.4 Å at ≈7 ns in system D. This increase in RMSD clearly indicates a drastic conformation change in MetRS structure in the presence of tRNA and MetAMP. The removal of the CP domain reduced the RMSD to ≈1.5 ± 0.5 Å in all of the simulations, including system D (SI Fig. 8). Thus, the sharp increase in the RMSD in system D (Fig. 2) is attributable to the opening of the active site pocket, which resulted because of a large displacement of the CP domain (SI Fig. 9). Such an opening allows the tRNA acceptor stem to enter inside the active site for aminoacylation.

Conformation of tRNA and the Recognition of Anticodon by MetRS. The conformation of tRNA in the acceptor stem region has changed in both the simulations C and D, in comparison with the starting conformation (SI Fig. 10). In the presence of MetAMP (system D), the acceptor stem unwinds and extends to reach out to the ligand–MetAMP. However, in the absence of MetAMP (system C), the acceptor stem retains its original backbone conformation, which is stabilized further by the interaction of the bases C1 and A72. There also are subtle differences in the conformation at the anticodon binding region. These differences are brought out in terms of the stacking energies and the interactions with MetRS. The anticodon loop is characterized by the stacking of bases C34–A35–A38, and the base U36 flips to the outside of the loop in the starting structure, a feature that is retained in both the simulations (Fig. 3 *a* and *b*). This feature also is reflected in the stacking energies (SI Table 3). However, the stacking of bases A37–A38 is significantly different in the tRNA structures from the two simulations. An increase in the stacking of A37–A38 in system C also is reflected in the mutual orientation of the two bases as shown in Fig. 3*a*. Thus, the long-distance influence of MetAMP is felt at the level of the orientation of the base (A37) in the anticodon region. Furthermore, differences also can been seen in the dynamically stable hydrogen bonds (defined as hydrogen bonds that are present in >50% of the simulation time) between MetRS and the anticodon region in systems C and D (SI Table 4 and Fig. 3 *a* and *b*). Specifically, the hydrogen bonding patterns of A35 (ribose oxygen O2′), U36 (base nitrogen N3), and A37 (phosphate oxygen) of tRNA with the residues of MetRS are different in the presence and absence of MetAMP. The amino acid residues Asn-391, Asn-452, Asp-456, Arg-395, Arg-533, Tyr-531, and Trp-461 of MetRS show differences in hydrogen bonding patterns with tRNA

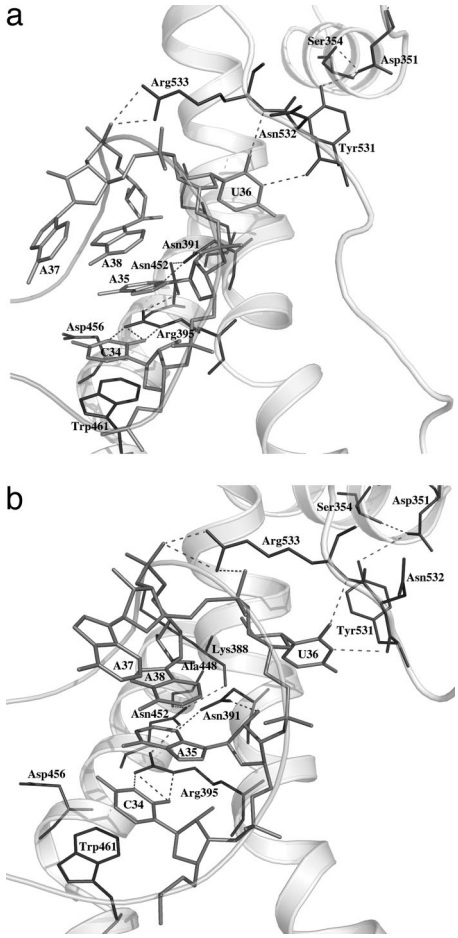

Fig. 3. Recognition of the anticodon C34–A35–U36 of $tRNA_f^{Met}$ by MetRS for systems C (*a*) and D (*b*) at 8.0 ns of simulation. The interacting protein region is shown as a transparent gray ribbon. The bases in the anticodon region and the interacting amino acids are shown as lines. The dynamically stable hydrogen bonds are shown by dotted lines.

either directly or indirectly mediated through other amino acids. The difference in these interactions results in the overall shift of tRNA from MetRS in system C, although the specific interactions of the anticodon bases with the protein are retained, as can be seen in Fig. 3 *a* and *b*. It is interesting to note that several amino acid residues such as Trp-461, Asn-452, Asp-456, and Arg-395 also have been shown to influence binding/activity of MetRS (5–11) and some of them (Trp-461 and Asn-452) are highly conserved.

Interactions at the Active Site. Dynamically stable hydrogen bonds made by MetAMP with MetRS residues are given in Table 1 for systems B and D, and they are schematically represented in SI Fig. 11 for system B. Overall stability and the number of hydrogen bonds in system B, however, are more when compared with those of system D. The amide nitrogen of methionine in MetAMP consistently forms hydrogen bonds with the carbonyl oxygen of Leu-13 (which is considered a representative of the active site, for identifying the network paths between the active site and the anticodon region in further analysis) and with the carboxylic oxygen of Asp-52 in both the systems. Further, the O3′ atom of MetAMP also exhibits a stable hydrogen bond with the carboxylic oxygen atom of Glu-27 in both the systems. The phosphate oxygen of MetAMP interacts with His-24 and with Tyr-15 in system B. However, the interaction with His-24 has diminished and the hydrogen bond with Tyr-15 is lost (present only for 21%) in system D. Based on the crystal

Table 1. Dynamically stable hydrogen bonds between MetAMP and MetRS in the active site

Donor–acceptor	B	D
MetAMP(O2′)–Glu-27(OE2)	99.65	63.38
MetAMP(O3′–Glu-27(OE2)	100.00	99.45
MetAMP(N)–Leu-13(O)	99.72	95.95
MetAMP(N)–Asp-52(OD1)	98.26	88.19
His-24(NE2)–MetAMP(O1P)	99.52	58.76
Tyr-15(N)–MetAMP(O1P)	51.67	
Gly-294(N)–MetAMP(O3′)	75.52	

The numbers in the table represent the percentage occurrence of the hydrogen bonds.

structure, it was suggested (28) that the residue Tyr-15 plays an important role in closing the active site in the presence of MetAMP. Our result shows the opening of active site in the presence of tRNA as Tyr-15 loses its interaction with MetAMP. Thus, the presence of the tRNA acceptor stem around the active site region makes some of the hydrogen bonds between MetAMP and MetRS less stable.

Dynamic Cross-Correlations. Cross-correlation coefficients were computed from the MD trajectories as described in *Materials and Methods*. The dynamic cross-correlation map (DCCM) of system D is presented in Fig. 4 (similar maps for the other three systems are provided in SI Fig. 12). Positive and negative correlations are shown in the upper and lower triangles of the DCCM, respectively. Note that the correlated fluctuations are obtained not only within the protein residues but also between residues of protein–tRNA and tRNA–tRNA. The correlation analysis pertaining to system D is presented below.

Correlations from DCCM. The correlations/anticorrelations within the protein residues are shown in the large square box between the markings (4–550) in DCCM (Fig. 4). The expected cross-correlation of residues within the helices are seen along the diagonal. A significant amount of correlation in the regions away from the diagonal also are seen, indicating a large amount of fluctuational correlation between noncontiguous residues in MetRS. It should be noted that some of the residues across domains are strongly correlated (e.g., residue 363 of KMSKS domain and

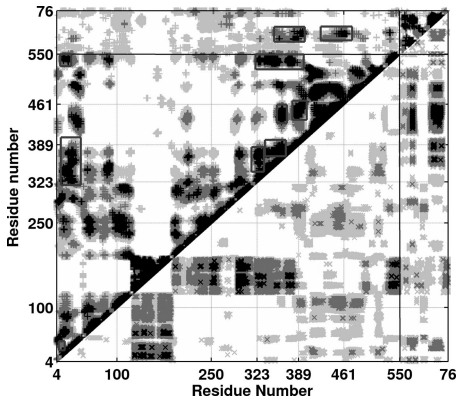

Fig. 4. DCCM representing the collective atom fluctuation for system D. The strong (C_{ij} = ±0.7–1.0), moderate (C_{ij} = ±0.5–0.7), and weak (C_{ij} = ±0.3–0.5) ones are represented by black, dark gray, and light gray, respectively. The lower and upper triangles correspond to negative and positive correlations, respectively. The numbers 4–550 correspond to MetRS residues. The residue numbers corresponding to different domains are marked: 4–99, 251–322, Rossmann fold domain; 100–250, CP domain; 323–388, KMSKS domain; 389–550, anticodon binding domain; 550–76, tRNA. The correlated residues enclosed in rectangular boxes are involved in the shortest pathways of communication, the details of which are given in Table 2.

Table 2. Network of correlated residues identified from DCCM for system D

Strong	Moderate
Leu-13 His-323 Leu-363	**Leu-13** His-28
Asn-452 **Trp-461**	Lys-388 **Trp-461**
Leu-13 Val-326 Leu-355	**Leu-13** Ile-534
Val-381 Asn-391 **Arg-395**	**Arg-395**
Leu-13 Val-326 Leu-355	**Leu-13** His-28
Val-381 **Asn-391**	**Asn-391**

Correlation coefficient values for strong (C_{ij} = 0.7–1.0) and moderate (C_{ij} = 0.5–0.7) levels. The paths are evaluated between the residues shown in boldface.

Asn-452 of the anticodon binding domain). The correlated residues involved in the shortest pathways of communication (which is discussed in a later section) are enclosed in rectangular boxes. There also is a significant amount of anticorrelated fluctuation within the protein. The residues (101–250) of the CP domain are moderately anticorrelated with the residues (1–100) of the catalytic domain and the stem-contact fold domain (323–389), which means that the CP domain moves away from the catalytic and the stem-contact fold domains by an anticorrelated fashion.

The correlated and the anticorrelated fluctuations of tRNA with MetRS are shown in the rectangles between the markings (550–76) and (4–550), respectively, at the top and on the right of Fig. 4. It is interesting to see that the loop and the stem of the anticodon region of tRNA (bases 33 to 38) are strongly correlated with the residues 452–465 and 388–395 of the anticodon binding domain of MetRS. This feature may assist in binding to the cognate tRNA containing the bases C34–A35–U36. The rest of tRNA (bases 14 to 21 and 48 to 63) is strongly anticorrelated with the anticodon domain region of MetRS, indicating that the interaction between tRNA and MetRS is dominated mainly through the cognate stem region of the tRNA.

The fluctuational correlations of tRNA within itself are shown in the small box in the top right of Fig. 4. Correlations and anticorrelations are negligible between bases of the acceptor stem (bases 39 to 44 and 27 to 31) and the anticodon loop (bases 32 to 38) of tRNA. Thus intramolecular long-range signaling from the anticodon loop of tRNA to the active site of MetRS is very poor along the length of the L-shaped tRNA. This finding is consistent with the observation made earlier (14).

Correlated fluctuations and shortest pathways of cooperativity. The collective atomic fluctuations for MetRS–tRNA complex form cooperative networks among residues of MetRS. DCCMs are analyzed to obtain networks (paths) of such correlated residues. The shortest network of correlated residues between two selected residues in MetRS is extracted from DCCM. As mentioned earlier, the carbonyl oxygen of Leu-13 makes a strong hydrogen bond with the nitrogen of methionine part of MetAMP, which is stable throughout the simulation, and a set of residues (Asn-391, Arg-395, and Trp-461) interacts with the anticodon part of tRNA. Hence, Leu-13 from the activation site and the residues such as Trp-461, Arg-395, and Asn-391 from the anticodon region have been selected as the terminal residues between which the shortest paths have been explored These paths have been identified at different levels of correlations (strong: C_{ij} = 1.0–0.7; moderate: C_{ij} = 0.7–0.5; and weak: C_{ij} = 0.3–0.5). The correlated residues involved (at all levels) in the shortest paths between Leu-13 and Trp-461, Arg-395, and Asn-391 are marked as rectangles in the DCCM of system D (Fig. 4), and all of the identified paths are given in Table 2 (some of the correlated paths are depicted pictorially in SI Fig. 13).

Influence of tRNA and MetAMP on correlations within MetRS. The DCCMs also have been generated for systems A, B, and C (SI Fig. 12 *a–c*). A striking contrast between the DCCM of system D and those of systems A, B, and C is the reduction in the magnitude of cross-correlated/anticorrelated fluctuations. The correlations are reasonably high in the tRNA bound structure C, although they are not as high as in system D. Drastic reduction in the correlations is observed in systems A and B, both of which lack tRNA as the ligand, which clearly indicates that the binding of tRNA is required for efficient communication between the active site and the anticodon binding region in MetRS, and this communication is enhanced in the presence of MetAMP. Shortest pathways of communication between the two regions have been evaluated in all of the three systems (A, B, and C) (SI Table 5). It is interesting to note that the strongly correlated paths between the active site and the anticodon region are present only in system D and not in the other three systems.

Communication Pathways from PSN. The patterns of cross-correlation between different residues in MetRS have provided valuable information from which the paths of communications between the activation site and the anticodon binding region have been constructed. However, it does not provide the details of complete connections. Here we make use of the paths derived from cross-correlation (Table 2) and find the missing links from PSN analysis (tRNA is not included in this calculation, because there was no correlation between the anticodon and the acceptor stem regions of tRNA). First, we evaluate all noncovalent residue–residue connections as described in *Materials and Methods*. Then, we select important residues interacting with MetAMP and the anticodon region as the end residues of the communication path, as mentioned in the previous section. Finally, we identify the shortest paths that connect the two end members by a series of noncovalent interactions and also incorporate the residues with strong or medium cross-correlations. Such an analysis has been carried out for all four systems, and the detailed results are given in SI Table 6. Average shortest path length between the two end residues is obtained by averaging the lengths of the shortest path obtained from all of the snapshots. The shortest of the shortest path obtained from the snapshots that incorporated the correlated residues also is given in SI Table 6 along with path details. Interestingly, this is much smaller for system D than that of the other three systems, indicating an efficient mode of communication in system D. In addition, the correlated shortest paths in system D are significantly shorter than the average shortest path lengths, in contrast to systems A and B, which again emphasizes the fact that the correlated paths are shortest in system D.

Because strong correlations are seen only in the case of system D, extensive analysis is presented for this system. Several paths (SI Table 6) identified from the analysis have been summarized in Fig. 5. Four paths (I, II, III, and IV) have been identified, which are depicted pictorially in Fig. 6. As mentioned in *Materials and Methods*, these paths were identified on the basis of the set of residues connecting the domain interface regions. Paths (Figs. 5 and 6) are more diverse toward the activation site, whereas they are more focused toward the anticodon binding region because of the fact that the residue connections from the anticodon region to the interdomain region are mainly through helices as shown in Fig. 6. This is achieved through connections Leu-363–Phe-437–Trp-432, Tyr-539–Leu-355–Phe-377, Trp-346–Phe-350–Gln-538, and Arg-36–Leu-495, respectively, in paths I, II, III, and IV (Fig. 5). Interestingly, several of these residues are in the linker peptide (348–362), which is at the domain (active site)–domain (anticodon recognition) interface and have been shown to play an important role in the communication (12). Furthermore, it may be noted that paths I and II (Figs. 5 and 6) incorporate the strongly correlated residues, whereas paths III and IV contain the moderately correlated residues, and the path lengths are smaller (Leu-13–Asn-452 is

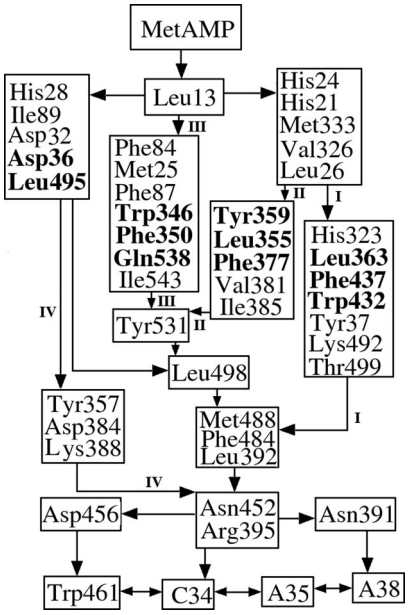

Fig. 5. Four communication pathways between the active site and the anticodon binding region as detected by the PSN analysis of noncovalent interaction in system D. The residues in bold are in the interface of the anticodon binding domain and the catalytic plus KMSKS domains.

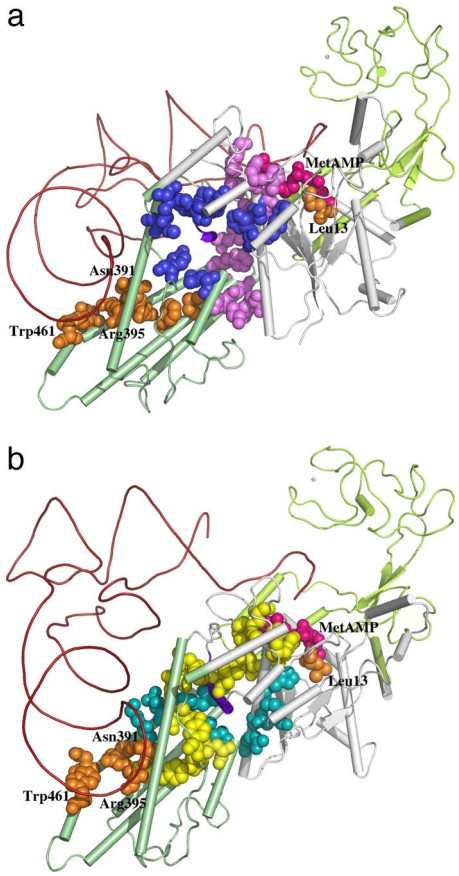

Fig. 6. Pictorial representation of the shortest paths (between Leu-13 and the anticodon binding residues Trp-461, Asn-391, and Arg-395) shown in Fig. 5 [(*a*) paths I and III; (*b*) paths II and IV)]. In both *a* and *b*, the residues participating in the communication paths and MetAMP are shown as spheres, MetAMP is shown in magenta, and Leu-13 and Trp-461 are shown in orange. The residues common between the paths I and III in *a* and between II and IV in *b* also are shown in orange. The residues that are exclusive to paths I and III are shown in violet and blue, respectively, in *a*, and the residues that are exclusive to paths II and IV are shown in yellow and cyan, respectively, in *b*. Light green, anticodon binding domain; gray, active site and KMSKS domains; purple blue, linker peptide; lemon green, CP domain; brown, tRNA.

connected by 10–14 residues) in the cases derived from moderate correlation than the strongly correlated paths (Leu-13–Asn-452 is connected by 17 residues). However, it is interesting to note that the path length is smallest in path IV (SI Table 6), and it is the path that is most frequently observed in the simulation (percentage occurrence of I, II, III, and IV is 2.3%, 6.7%, 2.7%, and 43.3%, respectively).

A comparison of the results of the path analysis of systems A, B, and C (SI Table 6) with that of system D indicates interesting differences. As discussed earlier, the striking feature is that there are no strong correlations in systems A, B, and C, contrary to that seen in system D. Furthermore, the shortest path lengths in the moderately correlated cases in systems A, B, and C are much longer (19 to 33) than those of system D (13 to 15). The paths near the anticodon region are more or less intact in most of the cases, which perhaps is caused by the fact that the anticodon domain is made up of helices and the communication proceeds through a narrow path. Diversity is more toward the MetAMP region. The paths in system C are similar to those of system D (paths I and II). However, similar paths are not seen between system D and those of systems A and B. Interestingly, the shortest path (path IV in Fig. 5, depicted in Fig. 6*b*) involving the subnetwork Tyr-357–Asp-384–Lys-388, which is present in the moderately correlated case D, does not seem to exist in the other systems, suggesting that this path is created only when both the tRNA and MetAMP are bound to the enzyme.

To summarize, it appears that the segments of paths are inherently present in the enzyme, and their interlinking to establish efficient communication of 70 Å between the anticodon region and the activation site requires the binding of both the tRNA and MetAMP. This finding is likely to be achieved through subtle movements of the two domains, which establish specific residue–residue contacts. In addition, the system has multiple paths, which makes the communication system robust. Our results exhibit significant correlation with experimental observations. For instance, the highly conserved residues Trp-461 and Asn-452 and the residues Asp-456 and Arg-395, which are known to interact directly or indirectly with the anticodon stem of tRNA (5–11), are part of the communication networks. More importantly, experiments have shown the vital participation of interface of domains (12–15) in the communication process. Our results not only confirm these observations but also provide the details of linking residues involved in this processes. Our findings may help in reexamining the experimental results from different perspectives. Furthermore, additional insights may be gained if the efficiency of aminoacylation is probed by designing experimental strategies to disrupt all of the multiple communication pathways.

Summary

To elucidate the modes of communication between the anticodon region and the aminoacylation region in MetRS, extensive MD simulations were carried out on different liganded states of MetRS. The simulation trajectories were analyzed in terms of correlated fluctuations and PSN.

The following important results have been derived from these investigations:

Strong communication paths within the residues of MetRS are established when the protein binds to both the tRNA and MetAMP, as evidenced by dynamic cross-correlations. The correlation/anticorrelation coefficients between MetRS and tRNA suggest that the protein residues have correlated movements only with the anticodon region of tRNA, whereas the rest of the tRNA moves

away from the protein. Furthermore, the communication does not proceed through tRNA because correlation within tRNA is poor.

Shortest paths of connections (connected through noncovalent interaction) between the anticodon region and the aminoacylation region of MetRS have been determined by combining the dynamic cross-correlation information with PSN analysis of the simulation snapshots. From such an analysis, it has been possible to identify about four communication paths between the active site region and the anticodon region of MetRS. Although fragments of paths are present in MetRS, efficient connections between the subnetworks are established only when the enzyme binds to both the tRNA and the activated MetAMP. The identified paths contain several amino acids found by experiments to be important for the activity. However, the identification of multiple paths of communication in detail suggests that further experiments can be designed to gain additional insights to the process of aminoacylation in MetRS.

Materials and Methods

Modeling, Simulations, and Cross-Correlations. MD simulations were performed at 300 K by using the AMBER9 package (30) with parm99 (31) parameters on four systems of *E. coli* MetRS. These four systems correspond to the unbound MetRS (27) (A), MetRS complexed with MetAMP (28) (B), MetRS–$tRNA_f^{Met}$ complex (C), and MetRS–$tRNA_f^{Met}$–MetAMP complex (D). The tRNA bound structures were modeled. The details of modeling, the simulation protocols, the evaluation of dynamic cross-correlations, and the generation of DCCMs from the simulation trajectories are presented in *SI Materials and Methods*.

PSG. The noncovalent residue-wise interactions in protein structures can be captured effectively by the PSG/PSN. The details of the construction of such a graph on the basis of interaction strength (I_{min}), and the implications of such graphs have been given earlier (22). The utilities of PSG in following MD trajectories have been discussed recently (32). Some of these details, including the selection of optimal strength for connections (I_{min}), are presented in *SI Materials and Methods*. The shortest paths between any two residues in the protein are evaluated as given below, assuming nonbonded connections at $I_{min} = 3.2\%$.

Shortest Path Between Vertices in PSG. The communication path between the anticodon region and the active site region is obtained by selecting important residues from both the regions and then identifying the shortest noncovalently connected path between them. The Floyd–Warshal algorithm (33) was used to determine the shortest path between selected pairs of vertices in a graph. The distance between connected residues was considered to be one, and the shortest path was identified as the path in which the two concerned residues were noncovalently connected by the smallest number of intermediate residues. Important residues from the active site (Leu-13) and the anticodon binding region (Trp-461, Asn-391, and Arg-395) were selected as the end members, and the shortest paths, along with the intermediately connecting residues between them, were evaluated from the snapshots of the MD trajectories. Our aim, however, was to select the paths that were the shortest as well as dynamically correlated, which was achieved by combining the information from DCCM in selecting the dynamically relevant shortest paths (DCCM provides correlations between residues that are not spatially connected directly to each other).

The following procedure was adopted for the identification of the shortest path, including the dynamically correlated residues. First, the dynamically correlated residues were identified from DCCM at two levels [(a) highly correlated ($C_{ij} = 0.7–1.0$) and (b) moderately correlated ($C_{ij} = 0.5–0.7$)]. This information was used in identifying the network of correlated residues connecting the residues selected from two regions of interest. Second, the shortest noncovalently connected path between the selected two residues was identified from the PSG, from all of the snapshots of the MD trajectories. Third, the snapshots were selected in which the shortest path included the correlated residues.

An analysis of the correlated shortest paths from the MD trajectories showed that a wide range of residues were part of the shortest path near the active site and the anticodon binding site regions. Thus, the conformational fluctuations along the trajectories were reflected in the variations of the length and the composition of the shortest path. Interestingly, most of the snapshots contained residues around the interface region (i.e., interface between catalytic plus KMSKS and anticodon binding domains) in their communication paths. Here, we were able to identify four different sets of residues in this region. Based on this result, we inferred that there are four different communication paths, and we designated them I, II, III, and IV. The average shortest path lengths and the shortest of the shortest paths incorporating the dynamically correlated residues were obtained for each of the four communication networks. These analyses were carried out on all of the systems (A, B, C, and D), and the results for system D were reported in detail.

We acknowledge support from the Computational Genomics Initiative at the Indian Institute of Science, funded by the Department of Biotechnology (DBT), India, and the computational facilities at the Supercomputer Education and Research Center (SERC), Indian Institute of Science, Bangalore.

1. Lee JW, Beebe K, Nangle LA, Jang J, Longo-Guess CM, Cook SA, Davisson MT, Sundberg JP, Schimmel P, Ackerman SL (2006) *Nature* 443:50–55.
2. Sekine S, Nureki O, Shimada A, Vassylyev DG, Yokoyama S (2001) *Nat Struct Biol* 8:203–206.
3. Sauter C, Lorber B, Cavarelli J, Moras D, Giege R (2000) *J Mol Biol* 299:1313–1324.
4. Kim HY, Pelka H, Brunie S, Schulman LH (1993) *Biochemistry* 329:10506–10511.
5. Leon O, Schulman LH (1987) *Biochemistry* 26:5416–5422.
6. Meinnel T, Mechulam Y, LeCorre D, Panvert M, Blanquet S, Fayat G (1991) *Proc Natl Acad Sci USA* 88:291–295.
7. Perona JJ, Rould MA, Steitz TA, Risler JL, Zelwer C, Brunie S (1991) *Proc Natl Acad Sci USA* 88:2903–2907.
8. Meinnel T, Mechulam Y, Blanquet S, Fayat G (1991) *J Mol Biol* 220:205–208.
9. Ghosh G, Pelka H, Schulman LH (1990) *Biochemistry* 29:2220–2225.
10. Ghosh G, Kim HY, Demaret JP, Brunie S, Schulman LH (1991) *Biochemistry* 30:11767–11774.
11. Schmitt E, Meinnel T, Panvert M, Mechulam Y, Blanquet S (1993) *J Mol Biol* 233:615–628.
12. Alexander RW, Schimmel P (1999) *Biochemistry* 38:16359–16365.
13. Weygand-Durasevic I, Rogers MJ, Soll D (1994) *J Mol Biol* 240:111–118.
14. Alexander RW, Schimmel P (2001) *Prog Nucleic Acid Res Mol Biol* 69:317–349.
15. Burbaum JJ, Schimmel P (1991) *Biochemistry* 30:319–324.
16. Zhang CM, Hou YM (2005) *Biochemistry* 44:7240–7249.
17. Kranz JK, Hall KB (1999) *J Mol Biol* 285:215–231.
18. Showalter SA, Hall KB (2005) *Biophys J* 89:2046–2058.
19. Kormos BL, Baranger AM, Beveridge DL (2006) *J Am Chem Soc* 128:8992–8993.
20. Kormos BL, Baranger AM, Beveridge DL (2006) *J Struct Biol* 157:500–513.
21. Sathyapriya R, Vishveshwara S (2007) *Proteins* 68:541–550.
22. Kannan N, Vishveshwara S (1999) *J Mol Biol* 292:441–464.
23. Vishveshwara S, Brinda KV, Kannan N (2002) *J Th Comp Chem* 1:187–211.
24. Brinda KV, Vishveshwara S (2005) *Biophys J* 89:4159–4170.
25. Brinda KV, Vishveshwara S (2005) *BMC Bioinformatics* 6:696.
26. Schmitt E, Panvert M, Blanquet S, Mechulam1 Y (1998) *EMBO J* 17:6819–6826.
27. Mechulam Y, Schmitt E, Maveyraud L, Zelwer C, Nureki O, Yokoyama S, Konno M, Blanquet S (1999) *J Mol Biol* 294:1287–1297.
28. Crepin T, Schmitt E, Mechulam Y, Sampson PB, Vaughan MD, Honek JF, Blanquet S (2003) *J Mol Biol* 332:59–72.
29. Nakanishi K, Ogiso Y, Nakama T, Fukai S, Nureki O (2005) *Nat Struct Mol Biol* 12:931–932.
30. Case DA, Darden TA, Cheatham TE, III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, *et al*. (2006) AMBER 9 (University of California, San Francisco).
31. Cheatham TE, III, Cieplak P, Kollman PA (2002) *J Biomol Struct Dyn* 16:845–861.
32. Ghosh A, Brinda KV, Vishveahwara S (2007) *Biophys J* 92:2523–2535.
33. Thomas HC, Charles EL, Ronald LR (1990) *Introduction to Algorithms* (MIT Press, Cambridge, MA), 1st Ed.