HOW ARE CHARGED RESIDUES DISTRIBUTED AMONG FUNCTIONALLY DISTINCT STRUCTURAL DOMAINS OF AMINOACYL-tRNA SYNTHETASES?

Safro M.*1, Tworowski D.,1 Feldman A.2
1 Department of Structural Biology, Weizmann Institute of Science, 76100, Rehovot, Israel;
2 EML Research gGmbH, Schloss-Wolfsbrunnenweg 31c, 69118 Heidelberg, Germany
* Corresponding author: e-mail: mark.safro@weizmann.a.ill

Key words: aminoacyl-tRNA synthetases, domain organization, primary structure, electrostatic potential, Poisson-Boltzmann theory

SUMMARY

Motivation: Unlike many other families of enzymes, which catalyze the same overall reaction, aminoacyl-tRNA synthetases (aaRSs) are extremely heterogeneous in terms of primary sequence and subunit organization. For the most part aaRSs are negatively charged at physiological conditions, as are tRNA substrates. What are the driving forces that ensure an attraction between like-charged macromolecules? As may be inferred from multiple sequence alignments (MSA), concentration of the invariant charged residues in structural domains doesn’t correlate with contribution of the domains to formation of the electrostatic field at long distances.

Results: In aaRSs family the subset of evolutionary non-conserved charged residues generates long-range electrostatic potential (EP) similar to the native one. We evaluate contribution of individual structural domains to the EP generated by native (NS), conservative (CS) and non-conservative subsets (NCS) of charged residues. For monomeric IleRS and heterodimeric PheRS we further analyzed the interplay between the domain functionality and their role in the field formation at long distances.

INTRODUCTION

AaRSs are of primary importance in the transformation of the genetic information from mRNA into polypeptide chain covalently attaching appropriate amino acids to the corresponding nucleic acid adaptor molecules – tRNA via the two-step aminoacylation reaction. The attachment of the correct amino acid to a tRNA is the crucial step determining the accuracy of protein biosynthesis. AaRSs exist as monomers, α2-dimers or tetramers of α4 and (αβ)2 types. Previously (Tworowski, Safro, 2003; Tworowski et al., 2005), we evaluated the contribution of electrostatic interactions to formation of aaRS-tRNA encounter complexes. It has been shown that 3D-isopotential surfaces (IPS) generated by monomeric, dimeric and heterotetrameric synthetases at 0.01kT/e contour level reveal the presence of large positive patches (“blue spaces”), one for each tRNA substrate molecule. It is apparent that this characteristic landscape of aaRS’s electrostatic potential is triggered by specific distribution of the charged residues along sequences and, thus in space.

A challenging problem is to identify the charge distribution along the polypeptide chain that substantially affects the topology of aaRS’s ES field. Based on MSA, we subdivided the entire pool of aaRS’s charged residues into three subsets: native (NS),
conservative (CS) and non-conservative (NCS). It is of interest that the total charge of CS is close to zero, whereas that of NCS is similar to the aaRS’s net charge. According to Smoluchowski’s theory of bimolecular association, the capture distance for two macromolecules is a sum of reactants hydrodynamic radii (Berg, von Hippel, 1985). Thus, each aaRS can be conceived as a reactive sphere (RES) built around the geometric center of enzyme with radius equal to the capture distance (R_{RES}). Our calculations made apparent the resemblance of shape and topography of positive patches at ±0.01 kT/e formed by NS and NCS, and essential differences in landscapes generated by the CSs.

**METHODS**

**Multiple sequence alignment (MAA).** ClustalW program (www.ebi.ac.uk/clustalw) with BLOSUM62 matrix and gap penalty of –12 was used. The sequences of different bacterial organisms were included in the MSA for each tested aaRS system. The sequences were derived from the Swiss-Prot Database (www.ebi.ac.uk/swissprot) and the Protein Data Bank (PDB) (www.rcsb.org/pdb). “Conservative subset” (CS) of charged residues includes strictly conserved residues as well as all meaningful substitutions (Asp ↔ Glu or Arg ↔ Lys) identified by MSA; “non-conservative subset” (NCS) consists of non-conserved charged residues.

**Statistical analysis of charge distribution among aaRSs’ domains.** Domains’ “charging” (D_{chrg}^i) and fractions of conserved (F_{CS}^i) and non-conserved (F_{NCS}^i) charged residues associated with ith domain are presented by:

\[
D_{chrg}^i = \frac{N_{chrg}^i}{N_{chrg}}, \quad F_{CS}^i = \frac{N_{CS}^i}{N_{CS}}, \quad F_{NCS}^i = \frac{N_{NCS}^i}{N_{NCS}}
\]

where \(N_{CS}^i\), \(N_{NCS}^i\) and \(N_{chrg}^i\) are the numbers of conserved, non-conserved and all charged residues of the ith domain, respectively; \(N_{CS}\), \(N_{NCS}\) and \(N_{chrg}\) are the conserved, non-conserved and total number of charged residues in aaRS’s polypeptide chain.

**Calculation of electrostatic potentials.** Electrostatic potentials were calculated at each grid point on the reactive encounter sphere (RES), built around the geometric center of molecule, by using Poisson-Boltzmann equation implemented in Delphi4.0 (Rocchia et al., 2004). The standard ionization states of the charged residues at physiological pH 7 were applied, i.e. neutral form for the His and ionized forms for Asp, Glu, Arg and Lys. To model charge distribution of native subsets, the formal charges were assigned to all charged residues of the protein. To evaluate the contribution of individual domain to the electrostatic potential the original PDB-files were modified by switching off the charged residues of the domain.

**Similarity analysis for electrostatic potentials.** Electrostatic potentials \(\phi\) for in silico modified aaRSs were calculated at the same grid points on RES, as native ones. For each pair of potentials (\(\phi_{mod}\) and \(\phi_{native}\)) in the native “blue space” area (i.e. at the points \(N_{\phi (+)}\)), the Hodgkin similarity index (SI) was calculated (Blomberg et al., 1999):

\[
SI = \frac{2(\phi_{mod}, \phi_{native})}{(\phi_{mod}, \phi_{mod}) + (\phi_{native}, \phi_{native})}
\]
\[
(\vec{\phi}_{\text{mod}} \cdot \vec{\phi}_{\text{native}}) = \sum_{x,y,z} \phi_{\text{mod}}(x, y, z)\phi_{\text{native}}(x, y, z)
\]  

(5)

Here \((\vec{\phi}_{\text{mod}} \cdot \vec{\phi}_{\text{native}}), \ (\vec{\phi}_{\text{mod}} \cdot \vec{\phi}_{\text{mod}})\) and \((\vec{\phi}_{\text{native}} \cdot \vec{\phi}_{\text{native}})\) denote scalar product calculated for all points within the region of positive patch; \(x, y, z\) are Cartesian coordinates of the grid points on RES.

RESULTS AND DISCUSSION: A CASE STUDY

By way of illustration we selected two types of aaRSs with different subunit organization: monomeric IleRS \([\alpha; \text{PDB code 1qu2}]\), and hetero-tetrameric PheRS \([\alpha\beta]; \text{PDB code 1eiy}]\).

The Hodgkin index was used as a measure of similarity between the native electrostatic potential and those produced by different subsets of charged residues. SI falls in the range from 1 to -1. The SI close to 1 indicates a high degree of similarity, whereas 0 and -1 correspond to fully uncorrelated and anti-correlated potentials, respectively. As it follows from our results SI reaches its peak \(~1\) when electrostatic potential on RES is generated by NCS. In contrast, the electrostatic potential generated by CSs uncorrelated (SI \(~0\)) with those of native set and NCS subset.

Analysis of CS’s residues, distributed among aaRS’s domains, reveals that larger portion of conserved charged residues is concentrated in catalytic domains (Fig. 1а, c). The Rossmann fold of IleRS contains 64 % of CS’s residues while the catalytic domain of PheRS 46 %. However there are structural domains such as N-term, Cp2, Zn-binding of IleRS and B4 domain of PheRS in which conserved charged residues were not found. This suggests dissimilar distribution of CS’s residues among different aaRS’s domains.

The intriguing result is that, regardless domain’s charging, the switching-off charged residues from certain domains has no significant impact on the distribution of electrostatic potential on RES. It is of interest that contribution of some domains to the EP on RES remains small, even though the concentration of non-conserved charged residues \((F_i^{\text{NCS}})\) there is relatively high. Thus high degree of similarity to native EP is observed for IleRS, when contribution of non-conserved charged residues from the N-terminal, Rossmann-fold, Cp2, Helical, C-term junction or Zn-binding domains is alternately excluded. This is evidenced by proximity of SI values to 1 (Fig. 1b). In case of PheRS, when charged residues are switched off within the domains B1, B3-B8 or catalytic module, the resulting electrostatic potentials are remain unchanged in compare to the native one (Fig. 1d). The domains that are significant for positive patches formation usually not involved in aminoacylation reaction and demonstrate low SI values. Some of them interact with tRNA, whereas function of others hasn’t been detected yet.

The Cp1 domain of IleRS that may be considered as a “crucial” for characteristic positive patch formation (see Fig. 1b) is associated with additional proofreading activity of the enzyme and contains a distinct active site where misactivated aminoacyl-adenylate or misaminoacylated tRNA are hydrolyzed (Silvian et al., 1999).

In PheRS isolated from \(Thermus thermophilus\), the coiled-coil domain of \(\alpha\)-subunit (CC*) is characterized by SI close to 0 (Mosyak et al., 1995). Two positive patches on RES, corresponding to two cognate tRNA interacting with PheRS becomes distorted and vanishingly small when charged residues of coiled-coil are switched-off. Functions of B2 domain that also plays a significant (albeit less pronounced) role in EP formation are not immediately evident from structures of the various functional complexes (Goldgur et al., 1997). A similar observation hold true for other aaRSs. Thus, domains that contribute significantly to the “blue space” formation very often involved in alternative activities of the
aaRSs. It is notable that structural domains of aaRSs playing a ‘crucial’ role in tRNA-protein recognition at long distances have a relatively low concentration of conserved charged residues. Therefore, NCS arranged in these domains can be treated as positive electrostatic determinant favoring the attraction and navigation of tRNA to its binding area.

It is possible to speculate that aaRSs family has acquired these domains at later stages of evolution.

Figure 1. Distribution of charged residues and SI values in IleRS (a, b) and PheRS (c, d) domains.
REFERENCES


