Evolution of Translation: Dynamics of Recognition in RNA:Protein Complexes

Zaida (Zan) Luthey-Schulten
Dept. Chemistry, Beckman Institute, Biophysics, Institute of Genomics Biology, & Physics

NIH Workshop 2009
On the Evolution of Translation in the Modern World

SepRS/CysRS Evolutionary History
Sethi, O’Donoghue, ZLS, PNAS 2005
O’Donoghue, Sethi, Woese, ZLS, PNAS 2005; Signaling Network
Sethi, Eargle, Black, ZLS, PNAS 2009

Dynamical Recognition Novel Amino Acids
Eargle, Sethi, Black, ZLS, JMB 2008
Evolutionary Analysis & Molecular Dynamics

MD performed with NAMD2 - System Setup

Simulation Parameters
- Minimization: 290,000 steps
- Production run: 108 ns
- Forcefield: CHARMM27(1)
- Time step: 1 fs
- VdW frequency: 2 fs
- VdW cutoff: 12 Å
- Switching distance: 10 Å
- Pair list distance: 14 Å
- Particle Mesh Ewald
- Full electrostatic update: 4 fs
- Ensemble: NPT
- Langevin temperature: 298.15 K
- Langevin pressure: 1 atm
- Periodic boundary conditions

Contents of System
- GluRS
- Glu-tRNA\textsuperscript{Glu}
- EF-Tu
- GTP
- Ions: Mg\textsuperscript{2+}, K\textsuperscript{+}
- H\textsubscript{2}O: ~27,000 molecules
- System: ~130,000 atoms

System Perturbations
- Deprotonation/protonation of charged Glu backbone amine
- Changing rotamer states of residues (2) involved in long-lived salt bridges

Charging tRNA through allosteric signaling

\[
\text{ATP + Glu + tRNA} \xrightarrow{\text{GluRS}} \text{Glu-tRNA + AMP + PP}_i
\]
How to Construct a Network

Nodes - defined at $C_\alpha$ (protein) and P (nucleotide) atoms
How to Construct a Network

Edges - connect nodes that are within a contact distance threshold for more than 75% of an MD trajectory
Communication between Identity Elements and Site of Chemistry
Different Paths can have Different Lengths

Path length = 10 (unweighted)

Information transfer? Weight contacts/links by correlations!
Correlations ($C_{ij}$) define signaling pathways (shortest) in GluRS:tRNA

- MD simulations run with NAMD
- NTP ensemble with explicit solvent
- Neutralized with Mg$^{2+}$ and K$^+$
- $C_{ij}$ values calculated over a 16-ns window

J. Eargle, A. Sethi, A. Black, Z. Luthey-Schulten, *PNAS* 2009
Mutation Changes Communication Pathways

\[
\frac{(k_{\text{cat}}/K_m)_{U13C}}{(k_{\text{cat}}/K_m)_{\text{wt}}} \approx 0.02 \quad \text{(experiment, Sekine, et al. JMB 1996)}
\]

\[
\frac{\exp(-\text{Sum of shortest path distances})_{U13\text{mutant}}}{\exp(-\text{Sum of shortest path distances})_{\text{wt}}} \approx 0.02 \quad \text{(Eargle, et al. PNAS 2009)}
\]
Change in the Network in Glu-tRNA Migration from GluRS to EF-Tu

>2-fold increase in \( k_{cat} \) for CysRS:tRNA:EF-Tu

8-fold increase in \( k_{cat} \) for LeuRS:tRNA:EF-Tu


Hausmann C, Praetorius-Ibba M, Ibba M. *NAR* 2007
Change in Protein:RNA Contacts During tRNA Migration

Protein contacts within 5 Å of the tRNA

blue - GluRS
red - EF-Tu
Change in Protein:RNA Contacts During tRNA Migration

purple - D arm community        yellow - T arm community
Evolution of Protein/RNA Interfaces: Dynamics of EF-Tu/tRNA Recognition

Role Modified Bases/Ions
Unnatural Amino Acids?

J. Eargle, et al. JMB 2008
tRNA\textsubscript{Cys} from E. coli with Modified Bases

- 4-thiouridine
- dihydouridine
- pseudouridine
- 2-methylthio-N6-isopentenyladenosine

**Acceptor Stem (AS)**

**D/GG Arm**

**T Arm/ Common**

**Anticodon (AC) Arm**
Interaction Energy at Interface

$\text{EF-Tu}^{T.th}/\text{RNA}^{Ecoli}$

Averaging over last 16 ns/20ns MD trajectories

J. Eargle, et.al. JMB 2008
EF-Tu Interface Conservation

Asp227
Deletion allows docking of asp-tRNA:Asn
Roy, Becker, Kern 2007

All three domains of life

• Residues colored by type.
• Positive: blue, Negative: red, Polar: green

Bacteria
Molecular Signatures in Ribosome Evolution

Ribosomal signatures: idiosyncrasies in ribosomal RNA and/or proteins characteristic of the domains of life

Comparing Ribosomal Structures

Ribosome large subunit showing ribosomal proteins L2 and L3
180,000 atoms in 4 rRNAs and 58 proteins

---

<table>
<thead>
<tr>
<th>Sequence Name</th>
<th>E. coli</th>
<th>H. marismortui</th>
</tr>
</thead>
<tbody>
<tr>
<td>23S rRNA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2aw4_B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s72_0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5S rRNA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2aw4_A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s72_9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribosomal Protein L2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2aw4_C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s72_A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribosomal Protein L3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2aw4_D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1s72_B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sequence and Structure Signatures in Ribosomal RNA

Ribosomal Signatures: Idiosyncrasies in rRNA and/or rproteins characteristic of the domains of life

69 (119) & 6 (14) in 16S (23S)

Roberts, Montoya, Sethi, Woese, ZLS, PNAS 2008 (in press)
Analysis of the ribosomal signatures in 90,000 environmental samples shows that no “gray” area exists: a ribosome is either bacterial or archaeal in nature.

It follows then that for the domains of life there is also no continuum. The bacteria and archaea are separated by an evolutionary saltation.
Domain Specific Proteins & rRNA Structure Signatures

Archaea

Late or Early Invention?

Effects on Translation?

Roberts, Montoya, Sethi, Woese
ZLS, PNAS 2008
Coevolution RNA-H16 and S4 rProtein Bacterial Insert

Using Mutual Information derived from Evolutionary Profiles
Max Entropy Cutoff for EP*

* E. Roberts, AS, JM, CW, ZLS PNAS (2009)

MI predicts signature in rRNA
Woese and Winker 1990
Protein Folding: hybrid MD
Charmm + Go Potentials

Energy function:

\[ E = E_{AA} + k \times E_{Go} \]
\[ E_{Go} = - \sum_{i}^{N_{C_{\alpha}}} \sum_{j \neq i, \pm 1, \pm 2}^{N_{C_{\alpha}}} \gamma_{ij} \times \exp \left[ -\frac{(r_{ij} - r_{ij}^{Nat})^2}{2 \left[ |i - j|^{0.15} \right]^2} \right] \]

Free energy profiles: umbrella sampling & WHAM

\[ E_i = E_{CHARMM} + V_i(Q) \]
\[ V_i(Q) = k_u(Q - Q_i)^4 \]
\[ \Delta FE(Q) = -k_BT \ln p_{est}^c(Q) \]

Mean first-passage time:

\[ \tau_{x_1 \rightarrow x_2} = \int_{x_1}^{x_2} dy \frac{\exp(\beta U(y))}{D(y)} \int_{0}^{y} \exp(-\beta U(z))dz \]

With T. Pogorelov, Biophys. J. 2004
Jonathan Montoya, 2007
R-protein S4 Docking
23S rRNA Sequence Signatures - Effects on Translation?

- Signatures are nucleotides that are conserved in and characteristic of a domain of life.
- 5% of 2900 nucleotides differentiate Bacteria from Archaea/Eukarya
- Signature A 2058 (*E. coli*) in exit channel is target of natural antibiotics like erythromycin
- In resistant bacterial strains, A2058DMA is post-translationally modified (riboswitch)
Evolutionary Analysis Tools in MultiSeq/VMD


Protein Sequence / Structure Data
Entire SwissProt Seqs

RNA Sequence / Structure Data
90,000 Envir. Seqs

Metadata Information, Clustal & Phylogenetic Trees, Signature Analysis

QR non-redundant seq/struct sets
Cluster analysis / Bioinformatics scripting

Blast & PsiBlast Sequence Editor

Tutorials
MultiSeq/AARS
EF-Tu/Ribosome
MultiSeq Tutorials