Covariance Ratio Analysis of Active HIV-1 Reverse Transcriptase

James M. Seckler[1], Serdal Kirmizialtin[2], Patrick L Wintrode[3], Kenneth Johnson[2], Alan Grossfield[4]

University of Rochester Medical Center, Rochester, NY, USA

[1]Department of Biostatistics and Computational Biology, [2]Department of Biochemistry, University of Texas, Austin

[3]Department of Pharmaceutical Sciences, University of Maryland, Baltimore, [4]Department of Biophysics and Biochemistry

Abstract

HIV-1 reverse transcriptase (RT) is a major drug target for HIV treatment, and understanding its function and inhibition would significantly improve our ability to create new anti-HIV drugs. RT can perform DNA-polymerization from either a DNA or an RNA template, and possesses an RNase function. Elastic network modeling is a method to rapidly probe and compare protein dynamics. We have previously shown that combining elastic network modeling with hierarchical clustering of both structural and dynamics data elucidates RT functional states. Here we extend our method beyond X-ray crystallographic structural data, to structural data determined by short molecular dynamics trajectories of RT bound to a primer template and either the correct dNTP or a mismatched dNTP. This reveals that RT bound to a mismatched dNTP is capable of entering into a novel nonfunctional state after dNTP incorporation. In this state, the thumb subdomain experiences inhibited dynamics and the primer/template breaks contacts with the p51 subunit. The incorporation of the correct dNTP shields RT from this nonfunctional state, allowing polymerization to continue. In summary, surveying structural and dynamics changes that occur in molecular dynamics trajectories alongside X-ray crystallographic structural data provides novel insights into normal RT function.

Structure of HIV-1 RT

- p66 subunit
  - -Fingers subdomain (blue)
  - -Palm subdomain (red)
  - -Thumb subdomain (green)
  - -Connection domain (orange)
  - -RNase H domain (purple)
- DNA (beige)
- p51 subunit (grey)

Elastic Network Models

- All-atom molecular dynamics (MD) too slow
- Coarse-grained model, Cα resolution, fast
- "Beads on springs"
- Single harmonic potential:

$$U_{ij} = k(r_{ij})\left(r_{ij} - r_{ij}^{\circ}\right)^2$$

$$k(r_{ij}) = \begin{cases} 1 & : r_{ij} < r_c \\ 0 & : r_{ij} \geq r_c \end{cases}$$

-k is a uniform spring constant

-r_{ij}° minimum energy - starting structure

- Diagonalize hessian matrix

-Yields eigenpairs

·Eigenvalues describe frequency

·Low frequencies ⟶ collective dynamics

·Eigenvectors describe direction

- Covariance complement
- Compares ENM results of structure A and B

$$CC_{A,B} = \left[\frac{\sum_i^N (\lambda_i^A + \lambda_i^B) - \sum_i^N \sum_j^N \sqrt{\lambda_i^A \lambda_j^B}(\vec{\nu}_i^A \cdot \vec{\nu}_j^B)^2}{\sum_i^N (\lambda_i^A + \lambda_i^B)}\right]^{1/2}$$

-λ is the inverse of the eigenvalue

-ν is the eigenvector

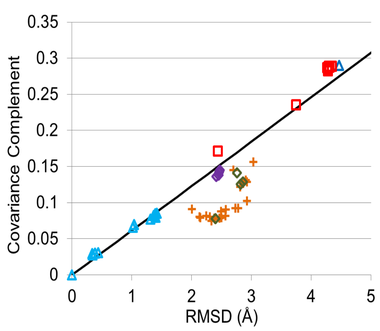
-Quantify similarity of modes

-Scales [0:1]

- 1 is completely orthogonal
- 0 is complete overlap

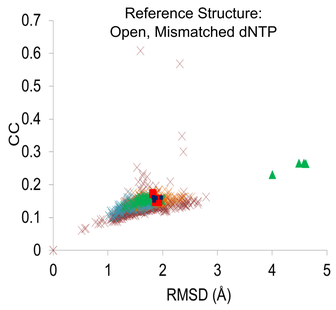
Hess, Phys Rev E (2000), 62, 8438–48

Covariance Ratio Analysis



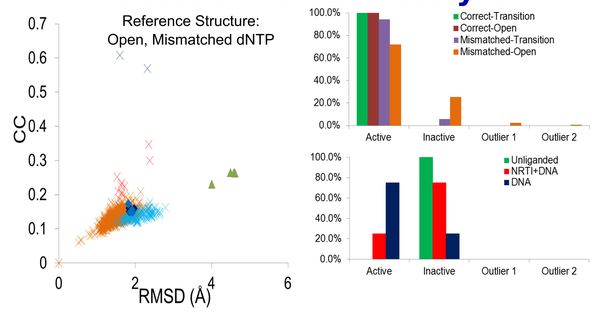
- Covariance ratio analysis (CRA)
- Linear relationship between like functional states
- DNA-bound (light blue △)
- RNA-bound (dark blue △)
- NNRTI-bound wild type or susceptible (orange +)
- NNRTI-bound hydrophobic core mutant (purple ◇)
- NNRTI-bound entry blocker mutant (green ◇)
- Unliganded (red □)
- Ratio of the RMSD to CC reference independent

Milestoning Simulation Data



- Short all-atom simulations
- Milestoning employs short MD simulations to determine kinetics
- Centered around steps of polymerization
- Four initial coordinate sets
  - Closed on dNTP (not shown)
  - Polymerization (not shown)
  - Transition
    - Correct dNTP (green X)
    - Mismatched dNTP (orange X)
  - Open
    - Correct dNTP (blue X)
    - Mismatched dNTP (brown X)
- Crystallographic data
  - DNA bound (blue □)
  - DNA + nRTI (red □)
  - Unliganded (green △)

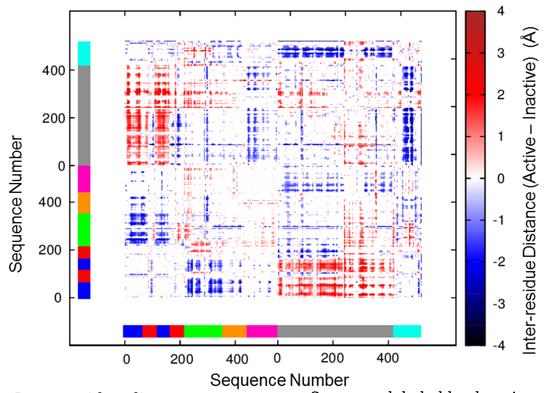
Covariance Ratio Analysis



- CRA reveals 4 clusters
- Active cluster (orange X)
- Inactive cluster (cyan X)
- DNA-bound (blue ◇)
- DNA+nRTI (lt blue ◇)
- Unliganded (green △)

- Active cluster
  - Most milestoning data
  - Majority of DNA bound
- Inactive cluster
  - Only mismatched dNTP milestoning data
  - Majority of DNA+nRTI
  - All unliganded

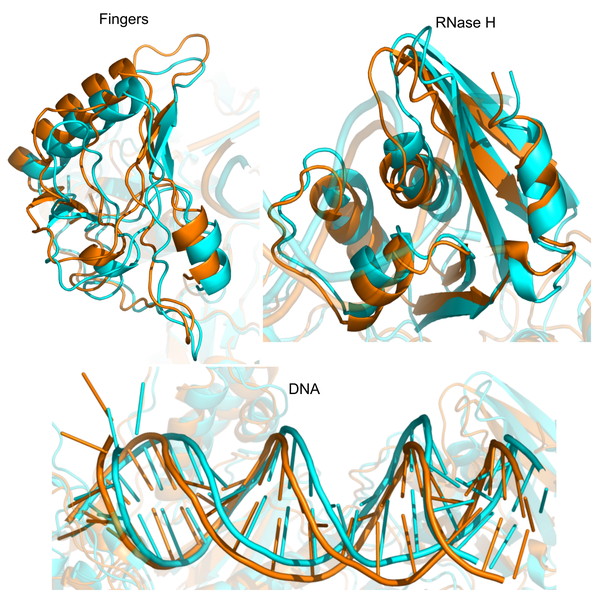
Changes in Inter-residue Distances Reveal Domain Reorientation



- Inter-residue distance
  - Welch's T-test of milestoning data
  - Significant differences colored
  - Active structures farther (red)
  - Inactive structures farther (blue)
- Domain reorientation
  - colored bands

- Sequence labeled by domain
  - Fingers subdomain (blue)
  - Palm subdomain (red)
  - Thumb subdomain (green)
  - Connection domain (orange)
  - RNase H domain (purple)
  - p51 subunit (gray)
  - DNA (cyan)

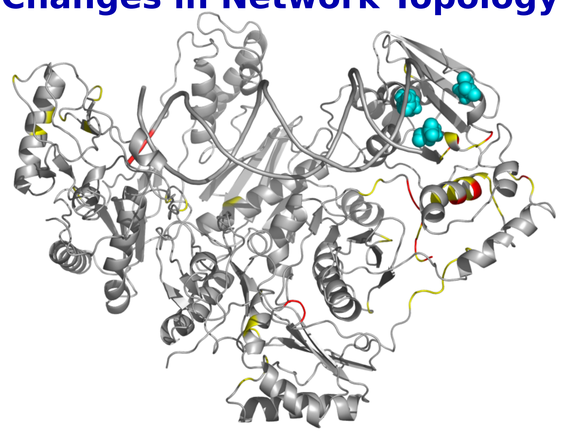
Structural Differences



- Active cluster (orange)
- Inactive cluster (cyan)
- Fingers rotated away from dNTP in inactive cluster
- RNase H rotated away from p51 subunit in inactive cluster
- DNA remains bound to RNase H

- DNA rotates away from p51 subunit
- DNA coordinates domain motion
- DNA responsible for the majority of change in contacts

Changes in Network Topology



- ENM insensitive to small changes in number of residue contacts
- Connectivity changes
  - Contact difference (Active - Inactive)
  - 30 or more contacts (red)
  - 15 to 30 (yellow)
- RNase H active site (cyan)

- Connectivity changes in p51 thumb subdomain, fingers subdomain, RNase H domains, and dNTP
- Subtle structural changes

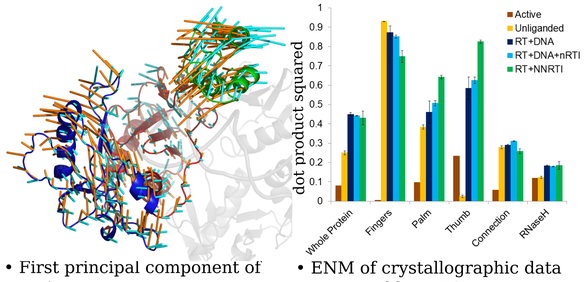
⬇

Large connectivity changes

⬇

Dynamics changes

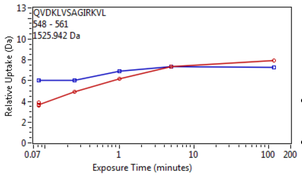
Difference in Dynamics



- First principal component of motion
  - Active cluster (orange)
  - Inactive cluster (cyan)
- RT polymerase domain
  - Fingers (blue)
  - Palm (red)
  - Thumb (green)

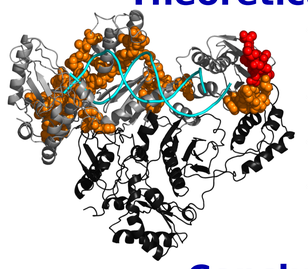
- ENM of crystallographic data comparable to PCA
- First principal component compared to lowest mode from ENM
- Active dynamics has poor agreement with ENM

Hydrogen/Deuterium Exchange Mass Spectrometry



- Hydrogen/deuterium exchange mass spectrometry (HXMS)
  - Labels exposed amide hydrogens
  - Secondary and tertiary structure stop exchange
- Probes changes in secondary structure
- Probes changes in local dynamics
- Results can be compared to ENM

HXMS Data Consistant With Theoretical Model



- Difference in deuterium uptake between RT with correct dNTP and mismatched dNTP incorporated
- Peptides show difference in amide exposure (orange)
- Peptide shows difference in structure (red)
- Backbone exposure change consistant with predictions

Conclusions

- Mismatched dNTP allows RT to access inactive state
- Inactive state dynamics mimic NNRTI inhibition
- Proofreading mechanism for preventing mistakes in polymerization
- Subtle changes in structure can lead to marked changes in dynamics

- CRA linear relationships maintained in molecular dynamics simulation
- Linear relationships allow for identification of different states, even with short MD trajectories
- Experiments consistant with mismatched dNTP bound RT has more mobile structure

Work done in LOOS (Lightweight Object Oriented Structure analysis library), an open source C++ library designed and maintained by the Grossfield lab. LOOS provides a concise, adaptable framework for designing analysis tools that interfaces with native file formats of most simulation packages.

http://loos.sourceforge.net