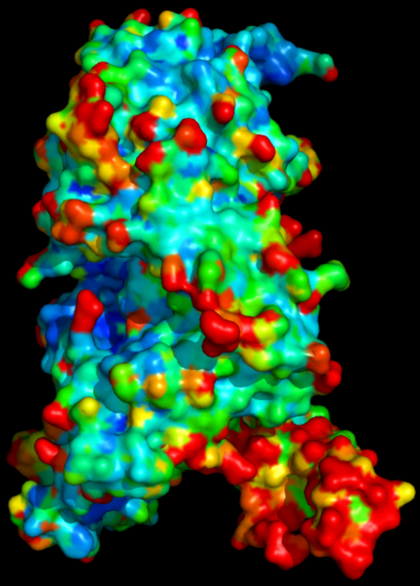
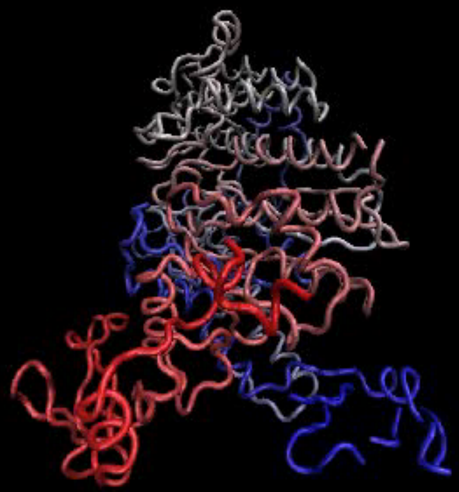
Modal Analysis of Myosin II and Identification of Functionally Important Sites

Onur Varol

M.Sc. Thesis Defense
25.07.2012

Jiggling and wiggling of atoms





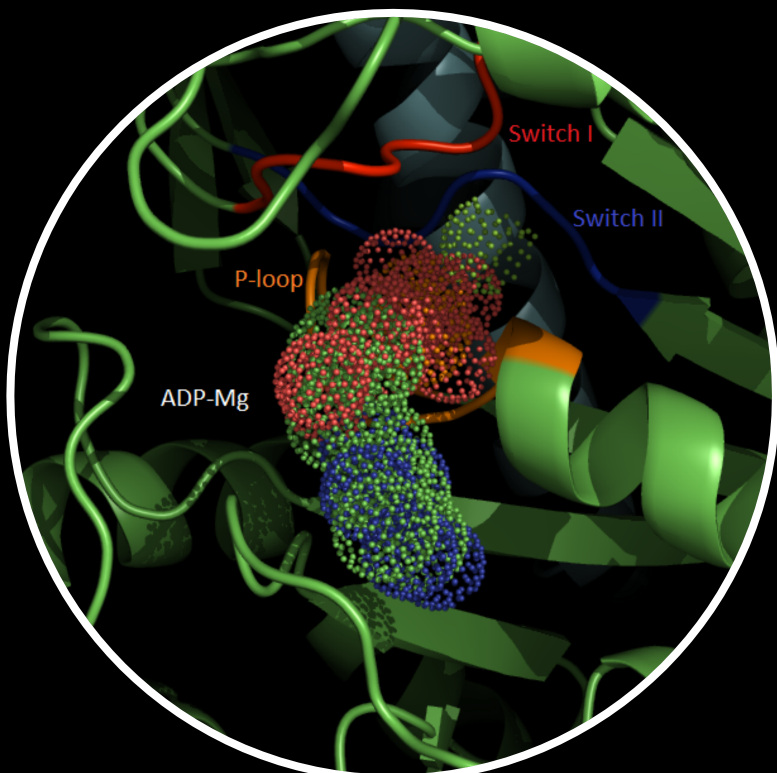
"Certainly no subject or field is making more progress on so many fronts at the present moment, than biology, and if we were to name the most powerful assumption of all, which leads one on and on in an attempt to understand life, it is that <u>all things are made of atoms, and that everything that living things do can be understood in terms of the jiggling and wiggling of atoms</u>."

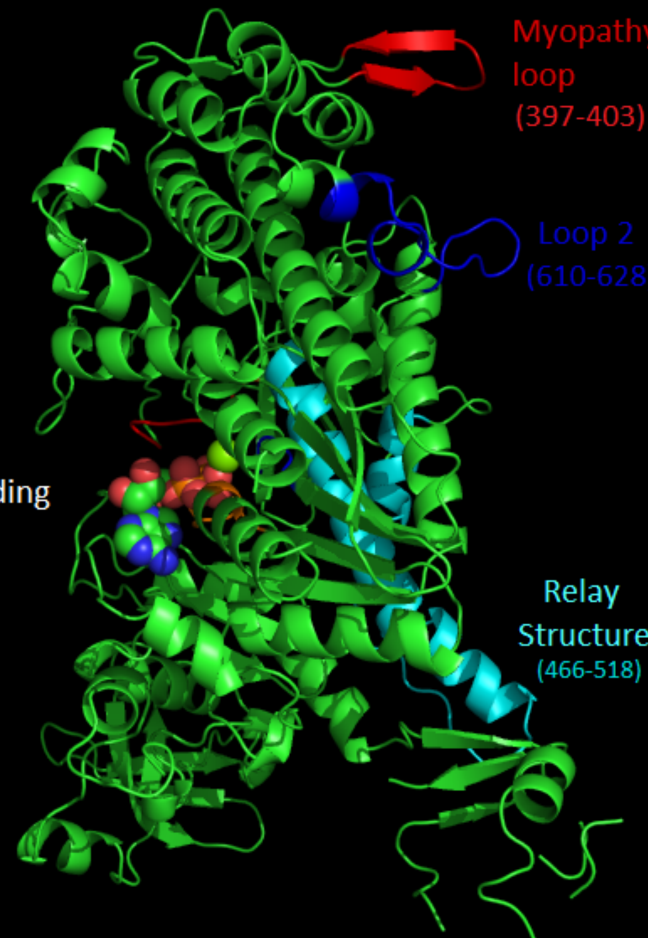
Richard Philips Feynman

Outline

- Myosin structure
- Allostery and myosin kinetic cycle
- Our approach
- Results

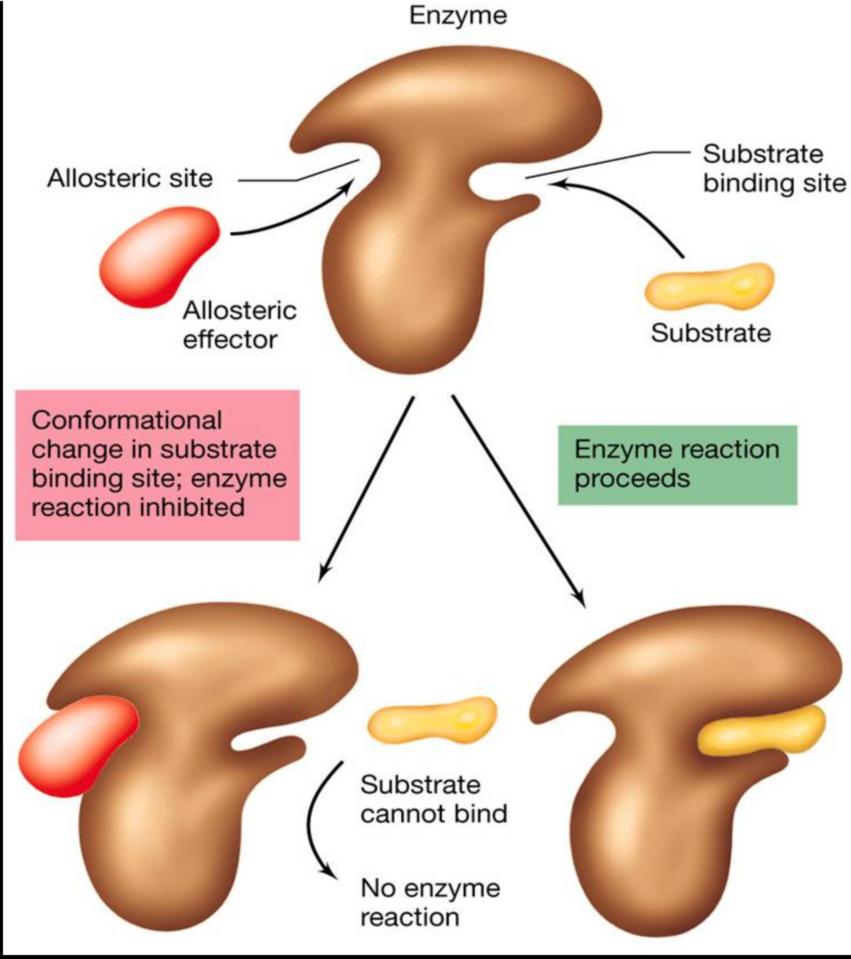
Myosin II


Switch I
Switch II
P-loop
ADP-Mg


Ligand-binding site
Myopathy loop (397-403)
Loop 2 (610-628)
Relay Structure (466-518)

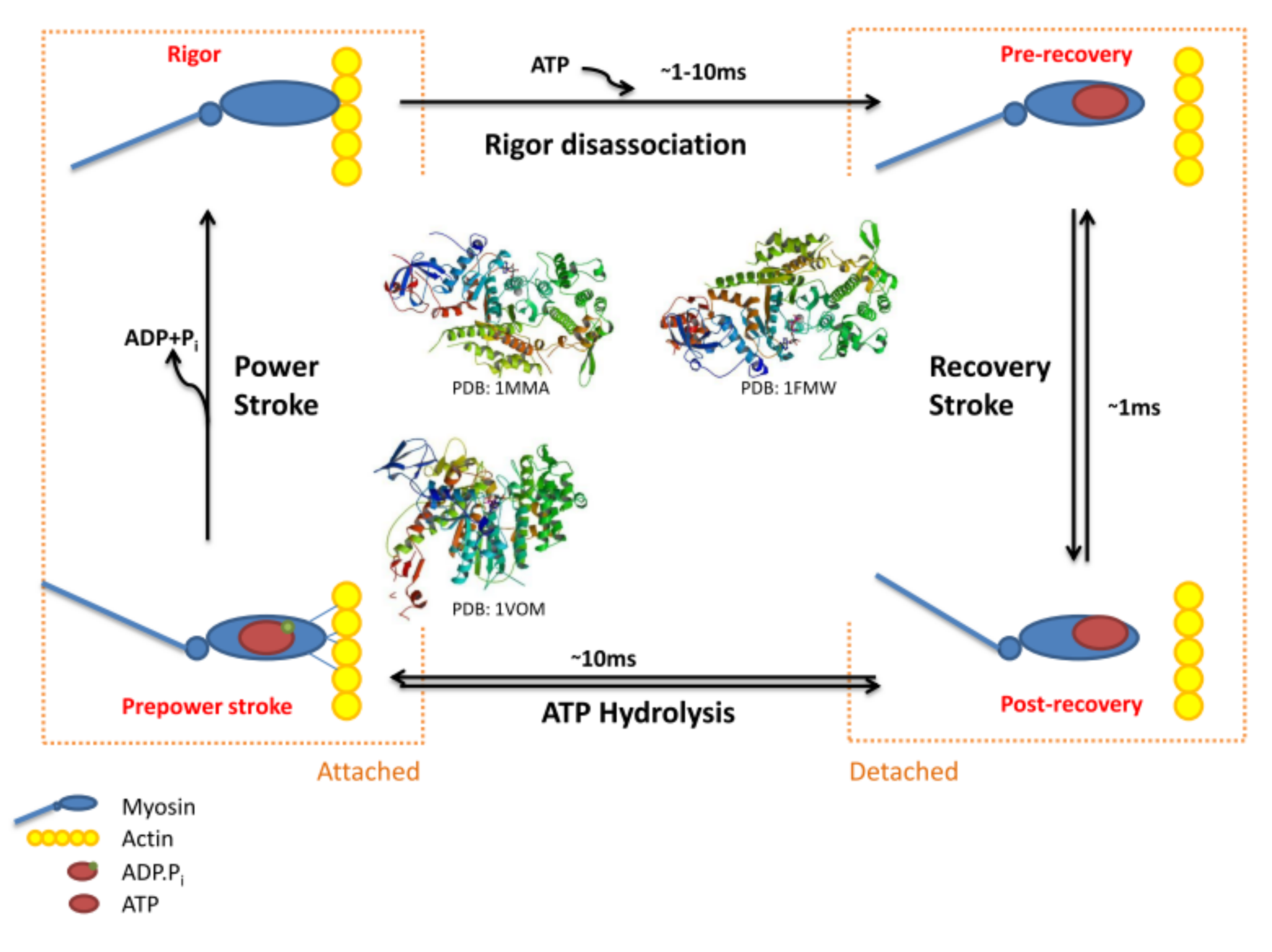
Allostery

- Allostery may cause conformational change on distant site
- Affinity of substrate binding changes
- Energy transfer from allosteric site to catalytic domain



http://202.204.115.67/jpkch/jpkch/2008/wswx/chapter%208.htm

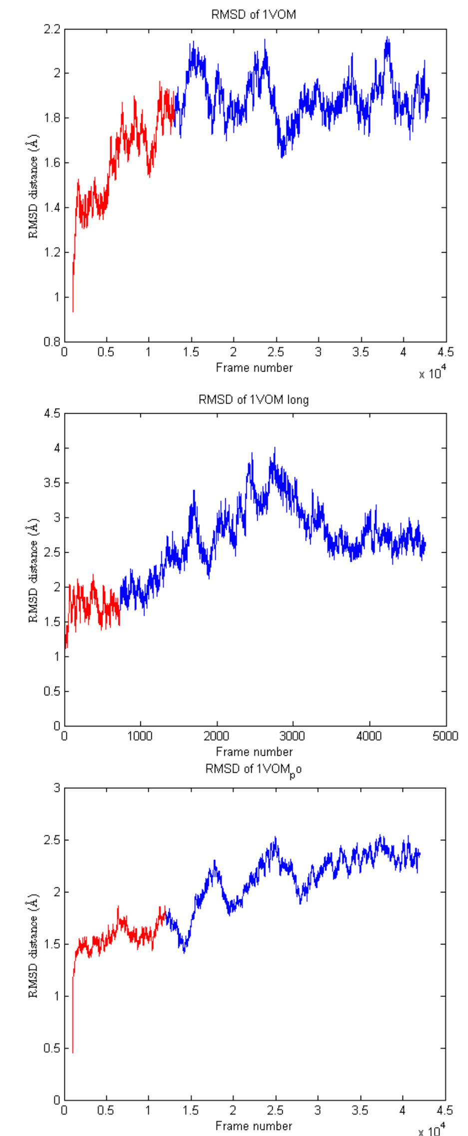
Myosin kinetic cycle



MD Simulation

- NAMD
- Charmm 27

- Langevin dynamics
- NPT ensemble
- Explicity solvent (15A cushion)
- 1fs time steps
  - 50fs frame rate 1VOM & 1VOM_po (~2ns)
  - 5000fs for 1VOM long (~20ns)



Computing resources used in this work were provided by the National Center for High Performance Computing of Turkey (UYBHM) under grant number 4001752012.

Our approach

$$\Delta \mathbf{R}_{3N_c,f} \xrightarrow[decomposition]{modal} \Delta \mathbf{r}_{3N_c-6,f}$$

$$\left.\begin{matrix} H_\nu \\ C_\nu \end{matrix}\right\} Hermite \begin{matrix} polynomial \\ constant \end{matrix} \; for\ rank\ \nu$$

f_0 distributions from normal distribution

f_1 distributions using normal distribution and Hermite polynomials and constants

PDF (probability density function) estimation for **MD**, f_0 and f_1 samples using KDE (Kernel Density Estimation) after <u>transform back into **coordinate space**</u>

Differences between distributions **MD**, F_0 and F_1 measured by KL metric

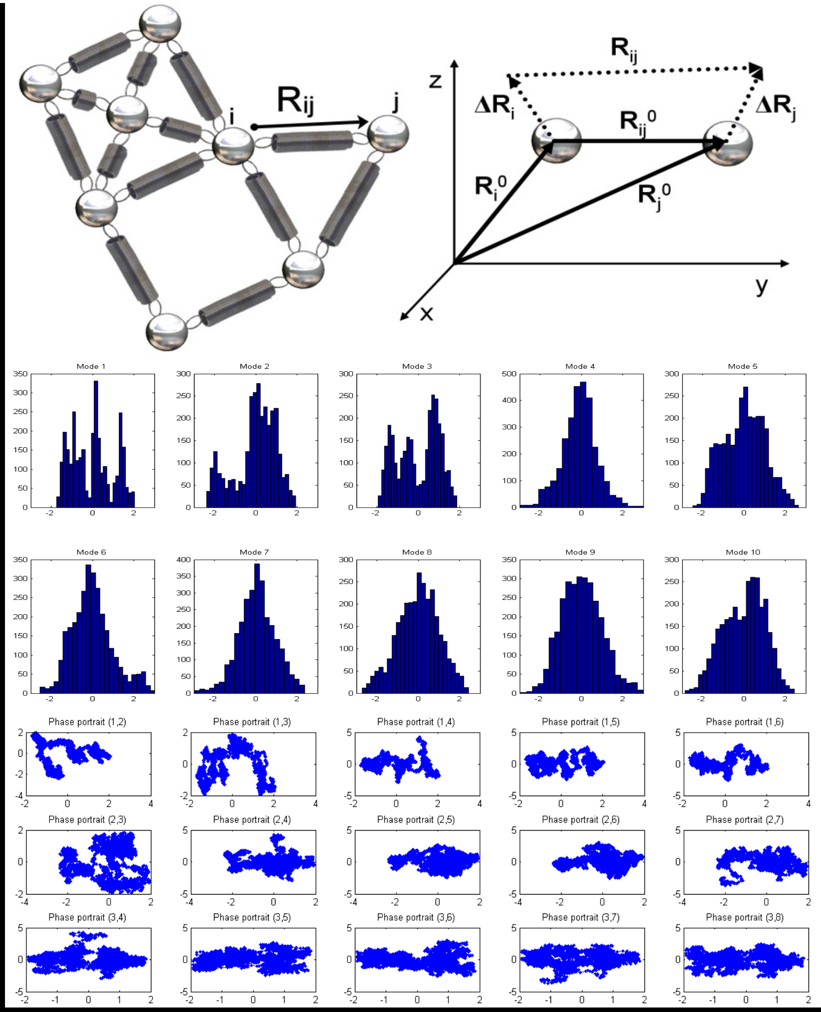
Modal decomposition

- Transform fluctuations into modal space
- Fluctuations can be sees as multivariate probability distribution function.

$$C = \left\langle \Delta\mathbf{R}\Delta\mathbf{R}^T \right\rangle$$

$$\left\langle \Delta\mathbf{R}\Delta\mathbf{R}^T \right\rangle^{-\frac{1}{2}} = diag\lambda^{-1/2}\mathbf{e}^T$$

$$\Delta\mathbf{r} = C^{-\frac{1}{2}}\Delta\mathbf{R}$$



Hermite expansion

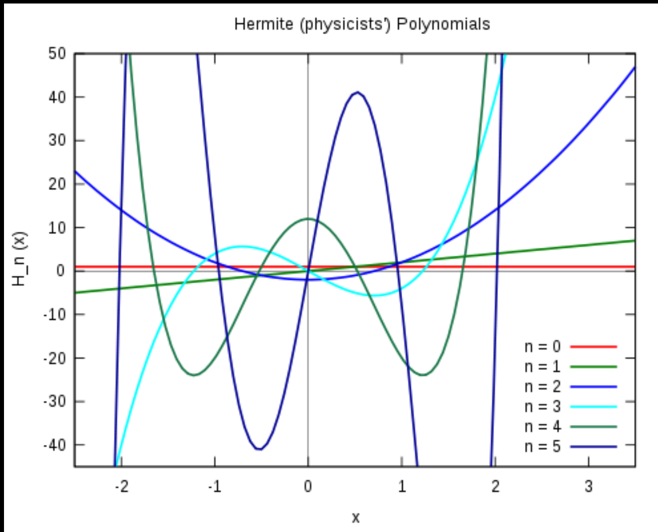
$$f(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2} [1 + \sum_i \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \langle H_\nu(\Delta r_i) \rangle H_\nu(\Delta r_i)$$

$$+ \sum_{i \neq j} \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \sum_{p=1}^{\nu-1} \binom{\nu}{p} \langle H_p(\Delta r_i) H_{\nu-p}(\Delta r_j) \rangle H_p(\Delta r_i) H_{\nu-p}(\Delta r_j) +$$

$$\sum_{i \neq j \neq k} \cdots]$$

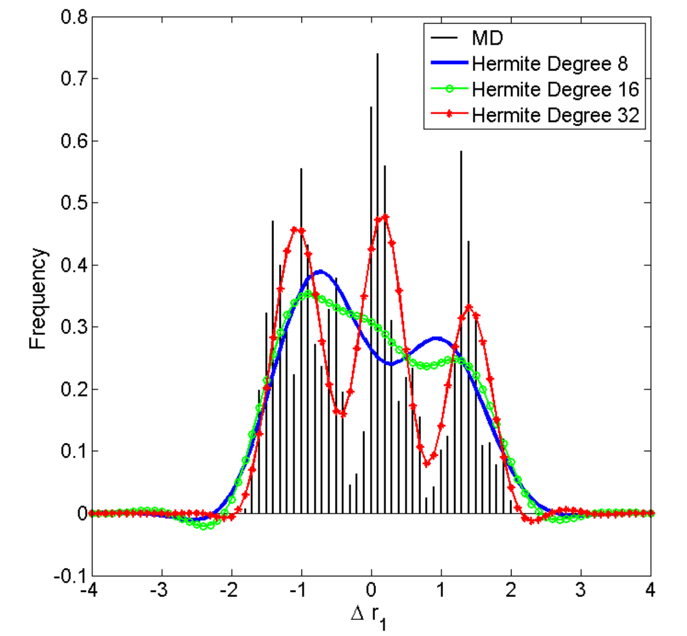
- Hermite polynomials are orthogonal w.r.t. weight function
- H_l hermite polynomial rank *l*
- Recursion equation

$$H_{n+1}(x) = x H_n(x) - H_n{}'(x)$$



Harmonic model & Anharmonic corrections

- Samples from f_o (harmonic)
- Samples f_1 (anharmonic model) obtained using f_o and first degree corrections
- Random sampling of f_1 by rejection sampling



$$f_0(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2}$$

$$f_1(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2}$$

$$\times \prod_i [1 + \sum_i \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \langle H_\nu(\Delta r_i) \rangle H_\nu(\Delta r_i)]$$

Kernel Density Estimation

- Samples in modal space transform back into real space
- KDE used in order to estimate pdf of the distributions

$$\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

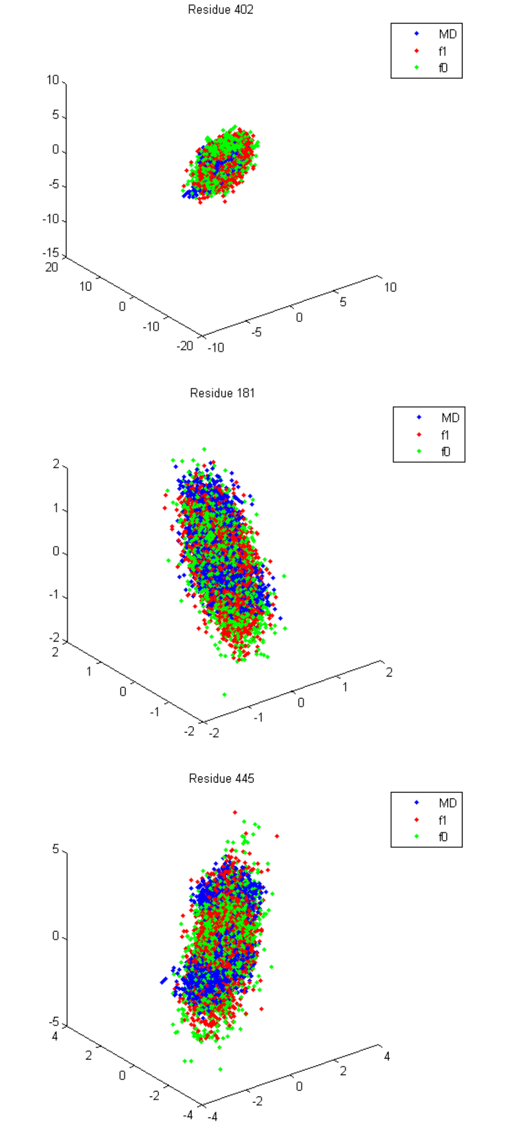
$$h = \left(\frac{4\hat{\sigma}^5}{3n}\right)^{1/5} \approx 1.06\hat{\sigma}n^{-1/5}$$

KL Divergence

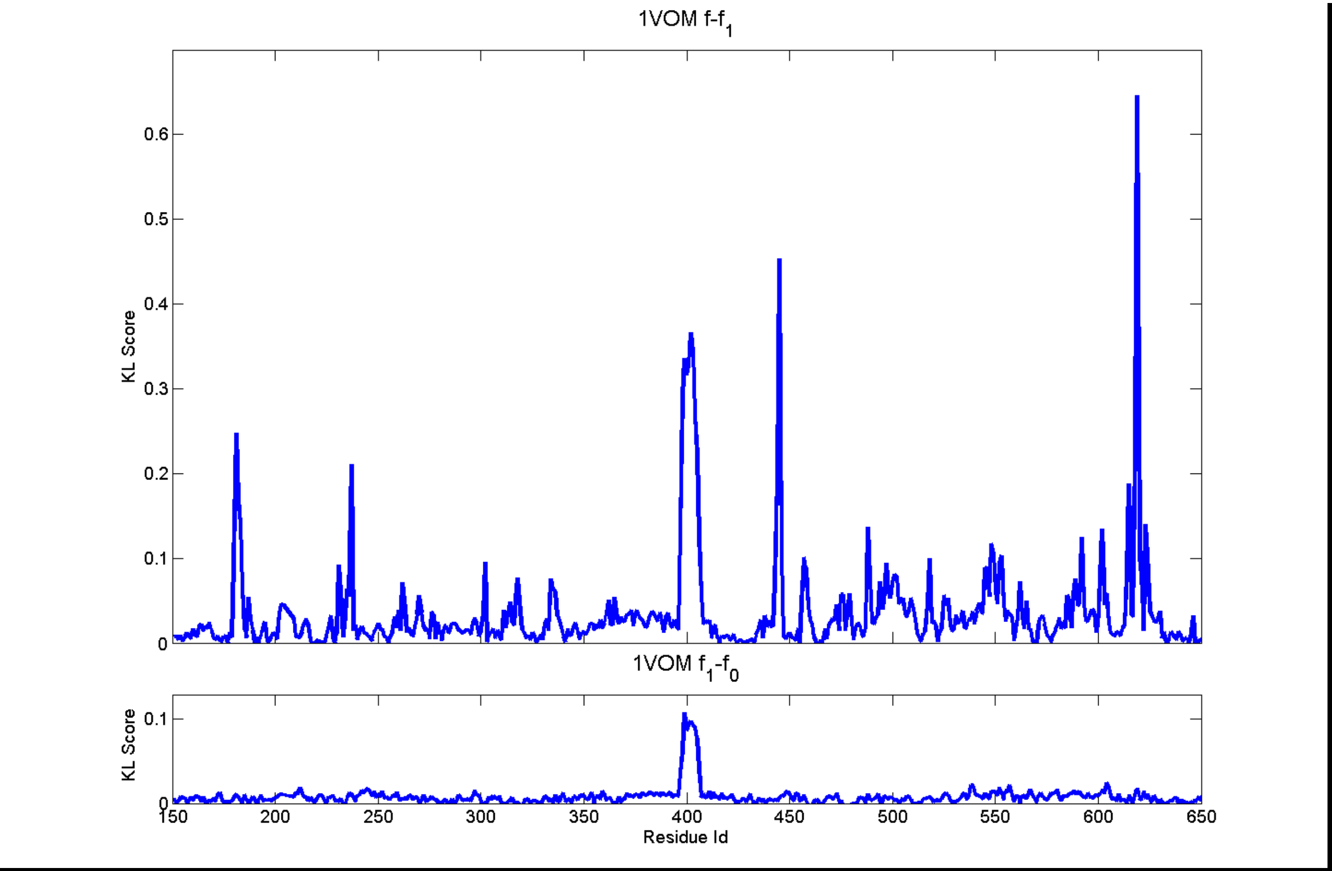
$$D_{KL}(P \parallel Q) = \sum_i P(i) ln \frac{P(i)}{Q(i)}$$

- Kullback-Leibler divergence measure of two pdf
- P is true data (MD), Q is model distributions

$$KL_{score} = KL_x + KL_y + KL_z$$

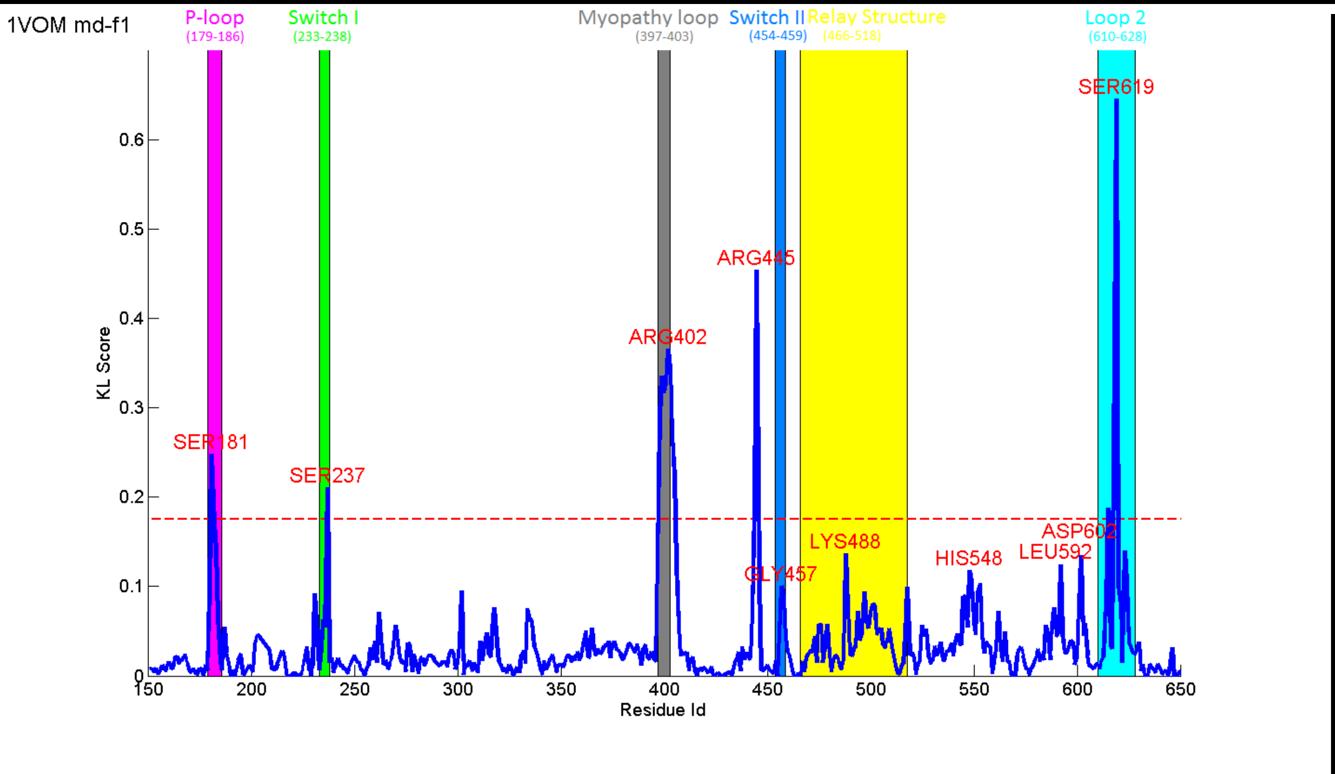


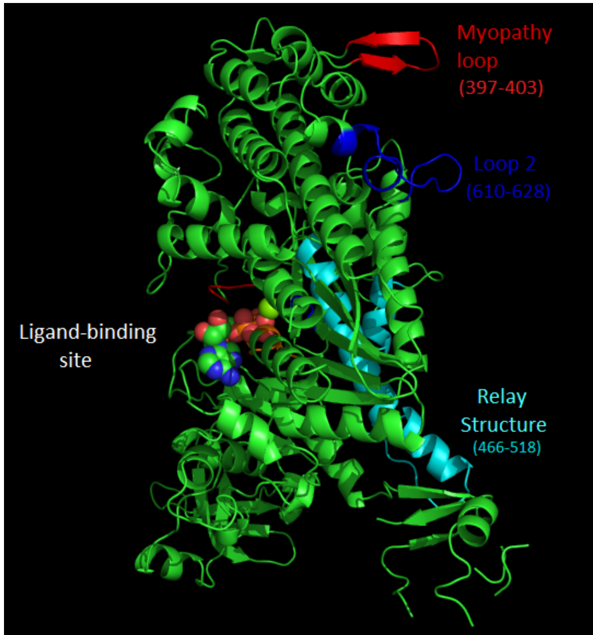
Anharmonicity vs. Mode-coupling

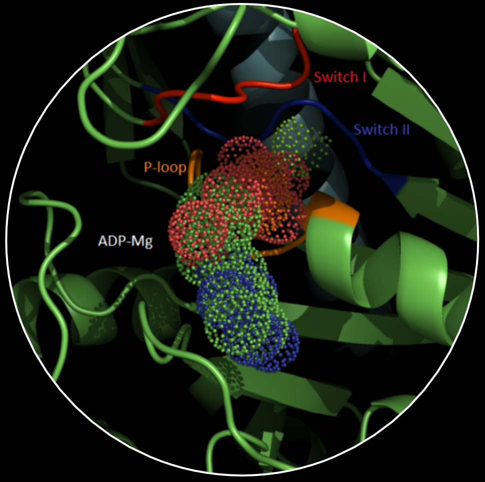


- Contribution of the mode-coupling is greater than anharmonicity which is also observed before in entropy change

Functionally Important Sites

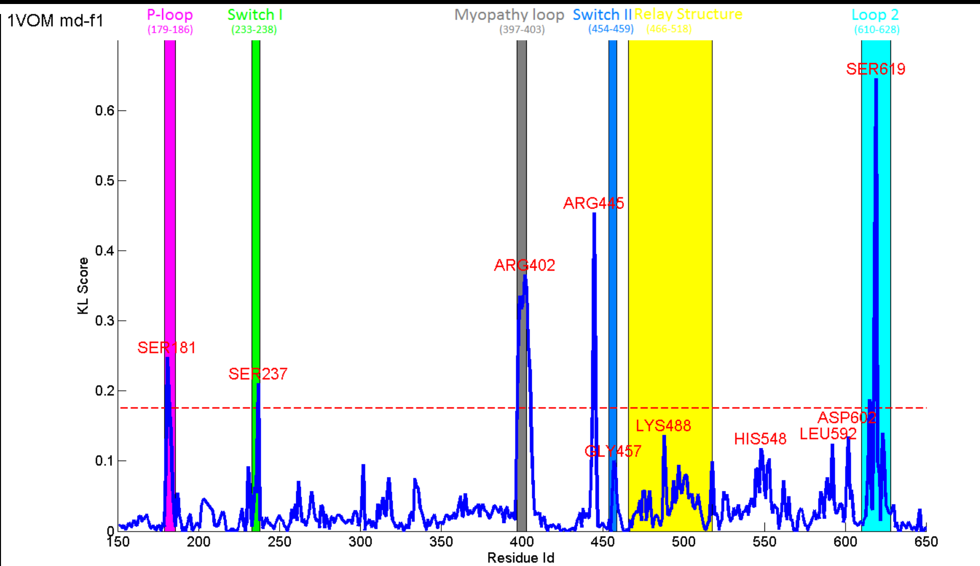


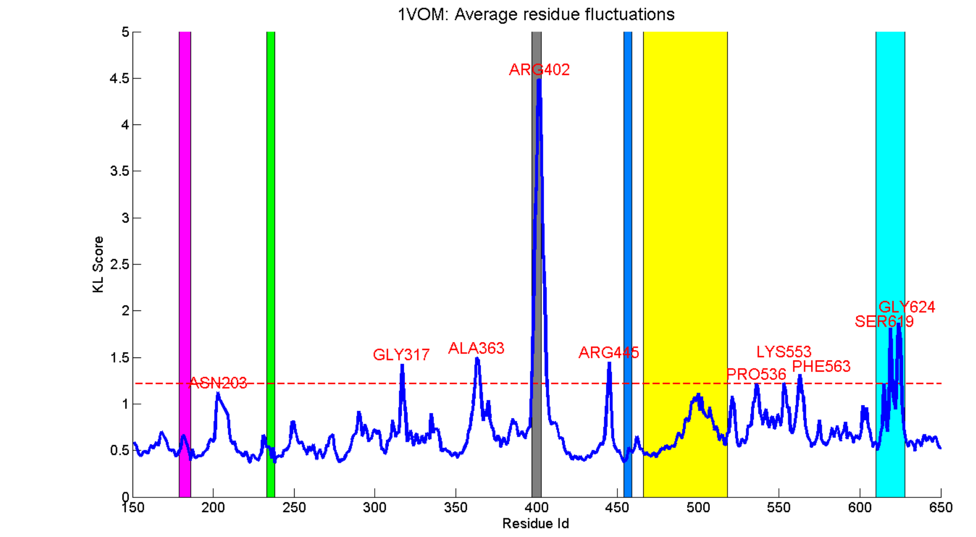




Compare with fluctuations

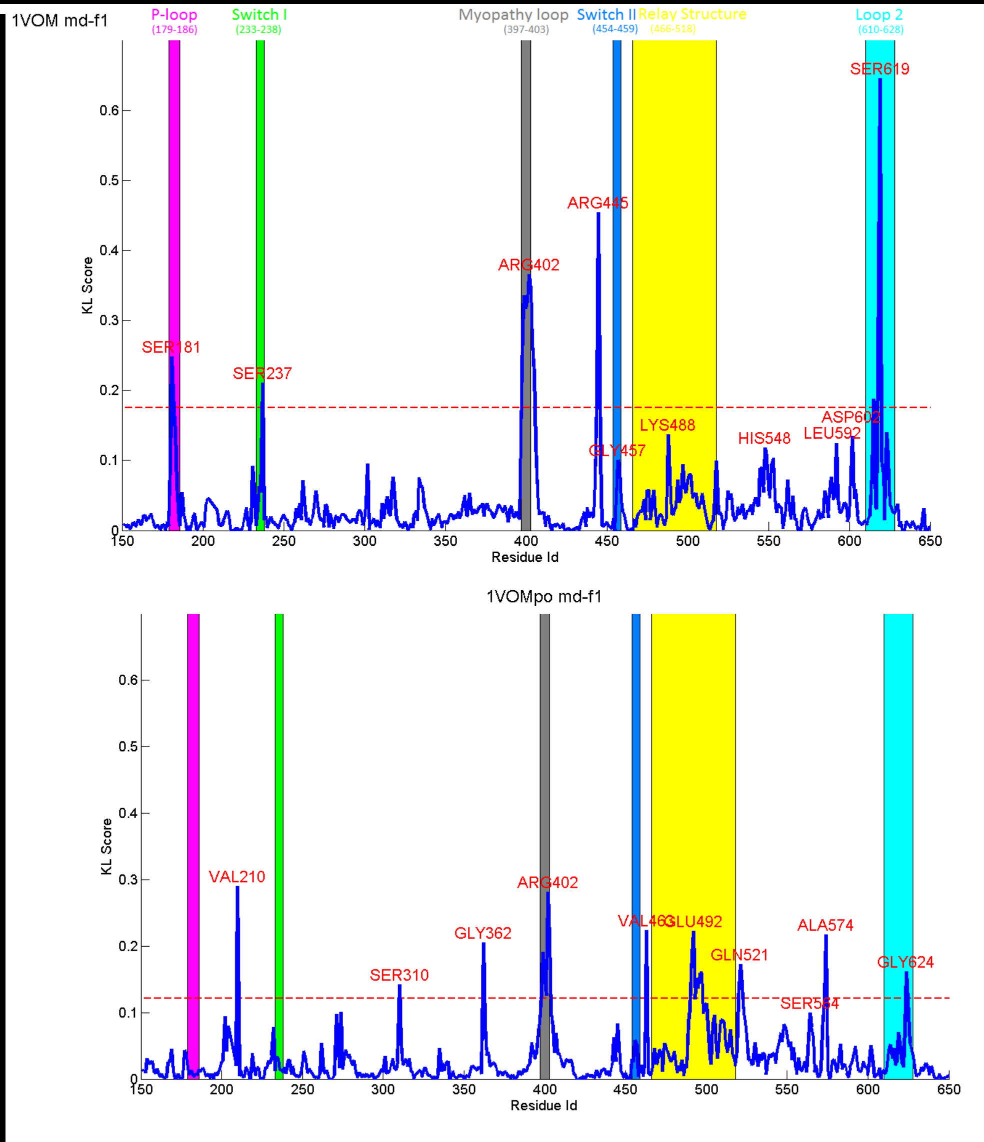
- Fluctuations inform about residue displacement
- Mostly on surface loop regions





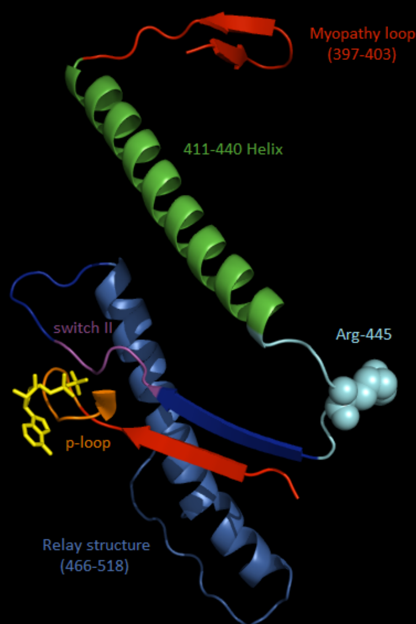
Ligand bound vs. unbound

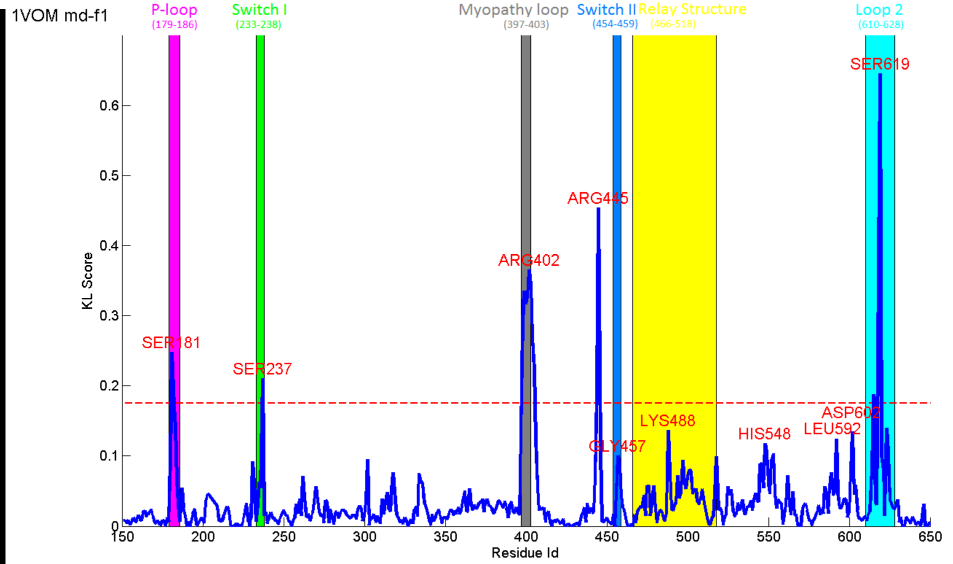
- Ligand bonding compress activation of other sites
- Ligand release reduces activation on actin-binding sites

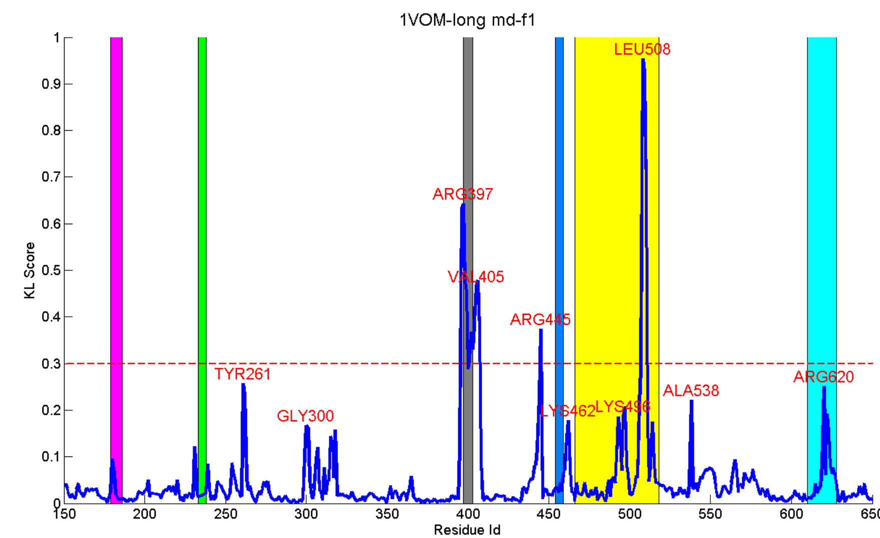


Longer simulation

- Long simulation shows characteristics of near-rigor state
- Central beta-sheets and relay structure drive transition



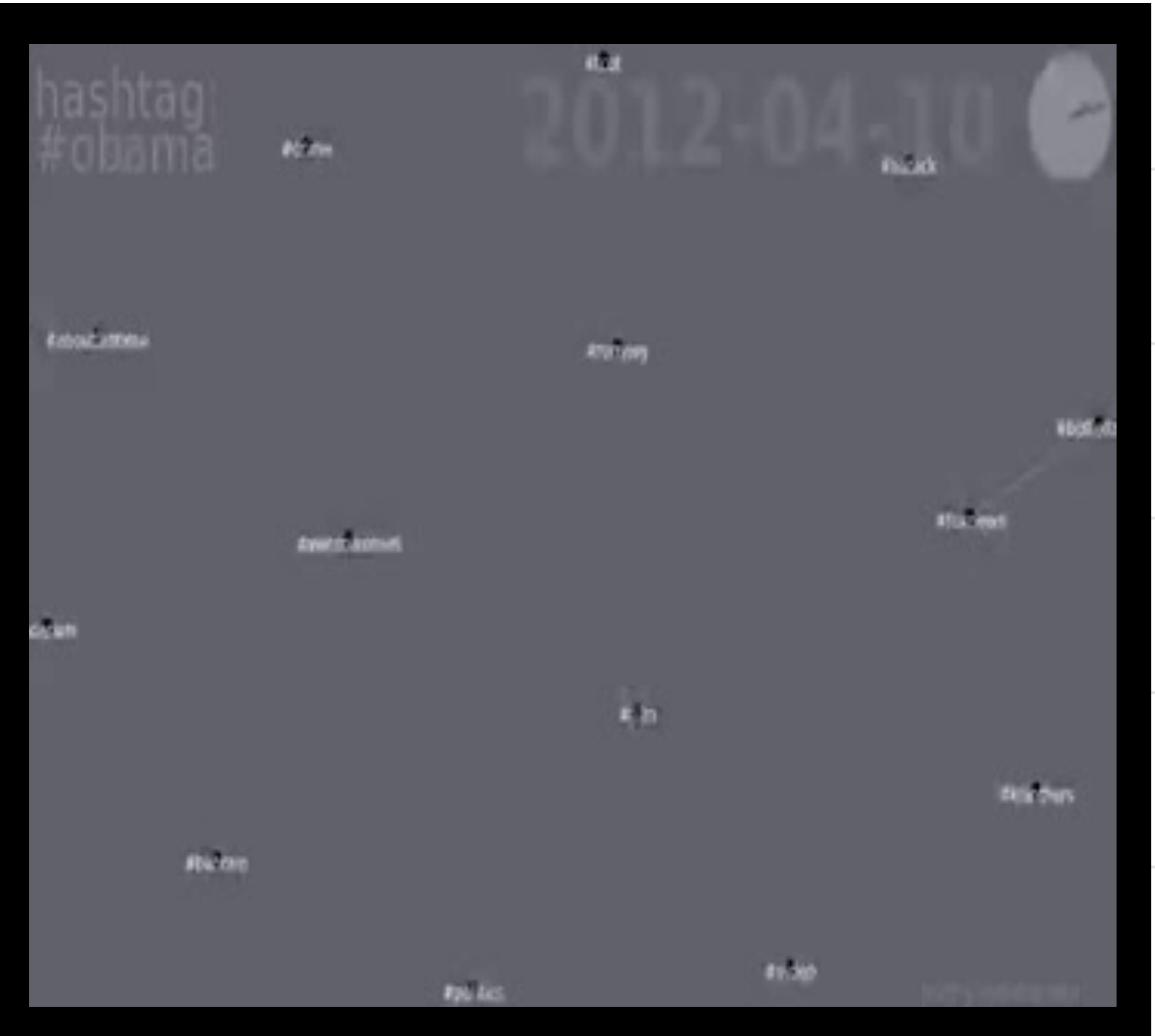


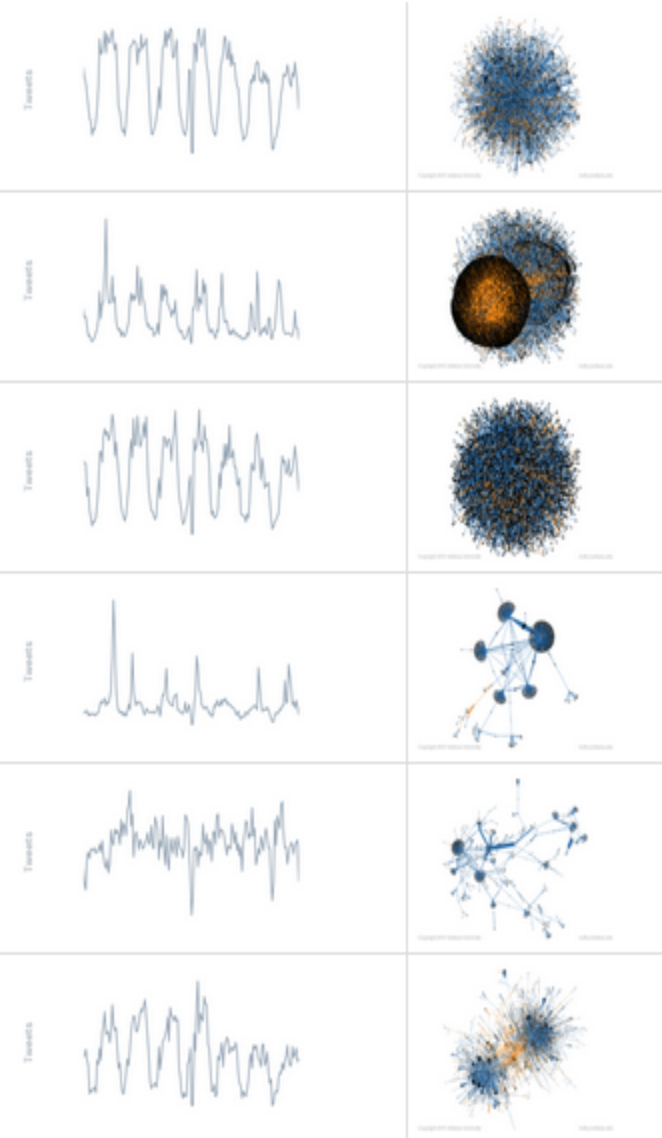


Conclusion

- Mode-coupling contribution highlight functionally important sites
- Ligand binding effect suppress nonspecific residues and points binding sites
- Importance of fluctuation analysis shown

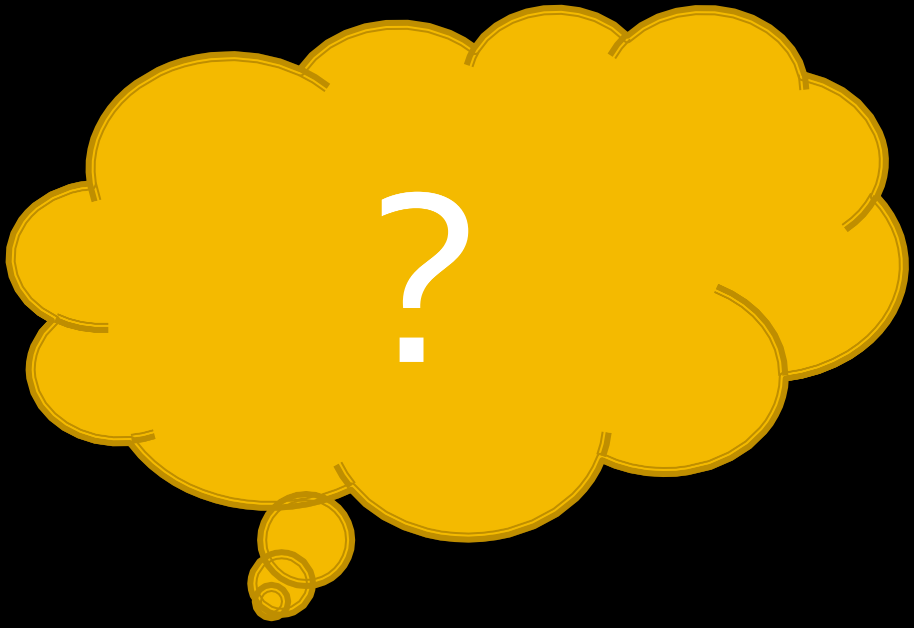
Interacting systems





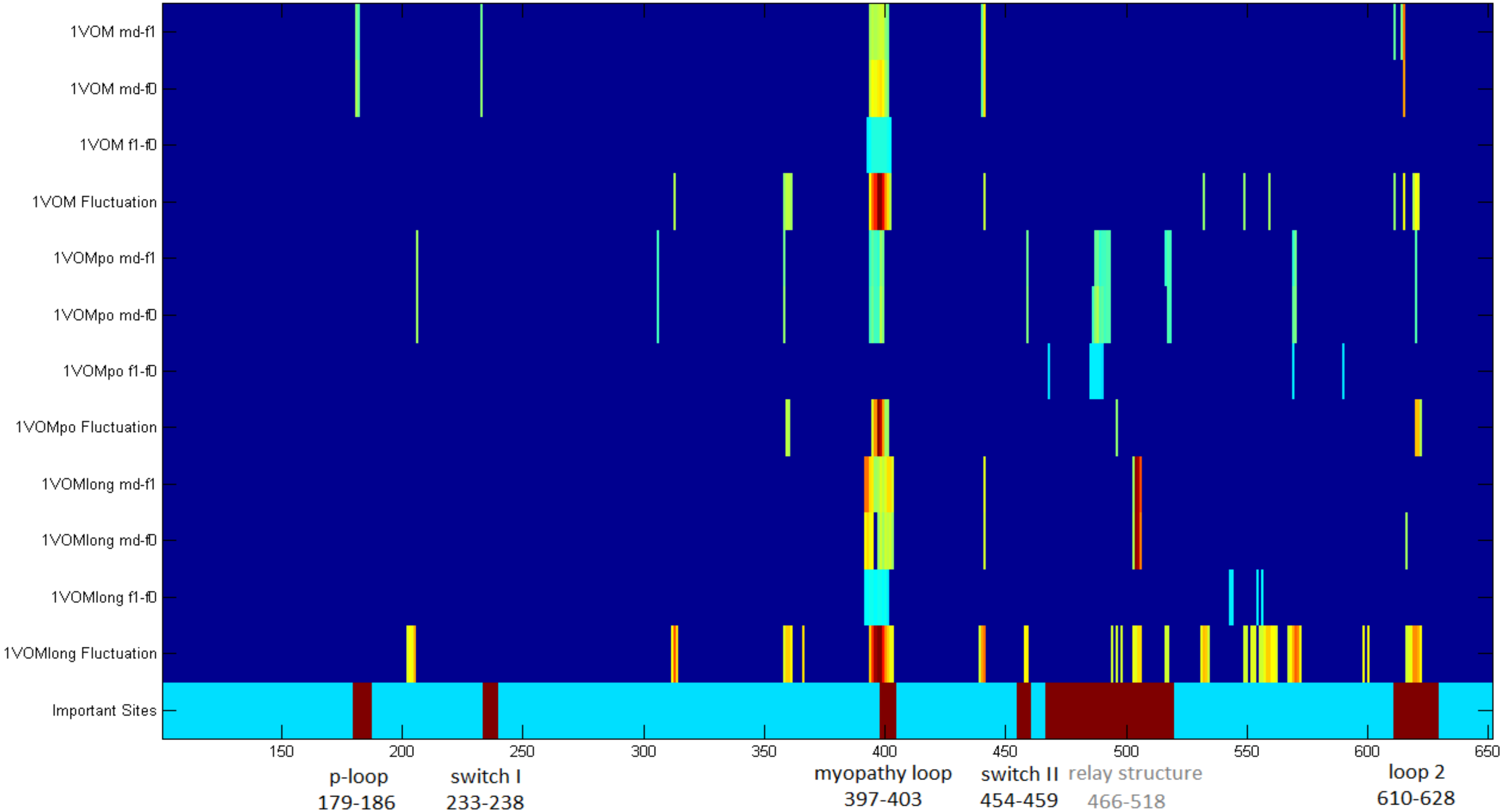
Thank you !!

Questions?



Backup slides

Precision table



Hermite

$$f(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^{3N}}} e^{-\frac{1}{2}\sum_{i=1}^{3N} \Delta r_i^2} \left[1 + \sum_{\nu=3}^{\infty} \mathbf{C}_\nu \cdot \mathbf{H}_\nu(\mathbf{\Delta r})\right]$$

$$H_\nu^{ij..k}(\mathbf{\Delta r}) = \frac{(-1)^\nu}{g(\mathbf{\Delta r})} \nabla^{ij..k} g(\mathbf{\Delta r})$$

Tensorial hermite polynomials can be obtained by successive differentiation using Rodrigues' formula

$$\mathbf{C}_\nu = \frac{1}{\nu!} \int_{-\infty}^{\infty} \mathbf{H}_\nu(\mathbf{x}) f(\mathbf{\Delta r})\, \mathbf{d\Delta r} = \langle \mathbf{H}_\nu(\mathbf{\Delta r}) \rangle / \nu!$$

Orthogonality relation

Value of tensor element does not depend on the order of indices due to commutativity of the gradient operator

$$\nabla_k \nabla_l - \nabla_l \nabla_k = 0$$

$$\mathbf{H}_\nu^{i_1 i_2 \ldots i_\nu}(\mathbf{\Delta r}) = \mathbf{H}_\nu^p(\Delta r_k, \Delta r_l)$$

where *p* is the number of indices equal to *k* (the remaining *v-p* indices equal to *l*)

Hermite-2

Covariance matrix in the normal basis is diagonal

$$\mathbf{H}_\nu^p(\mathbf{\Delta r}) = H_p(\Delta r_1) \times H_{\nu-p}(\Delta r_2)$$

$$f(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2} [1 + \sum_i \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \langle H_\nu(\Delta r_i) \rangle H_\nu(\Delta r_i)$$
$$+ \sum_{i\neq j} \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \sum_{p=1}^{\nu-1} \binom{\nu}{p} \langle H_p(\Delta r_i) H_{\nu-p}(\Delta r_j) \rangle H_p(\Delta r_i) H_{\nu-p}(\Delta r_j) +$$
$$\sum_{i\neq j\neq k} \cdots]$$

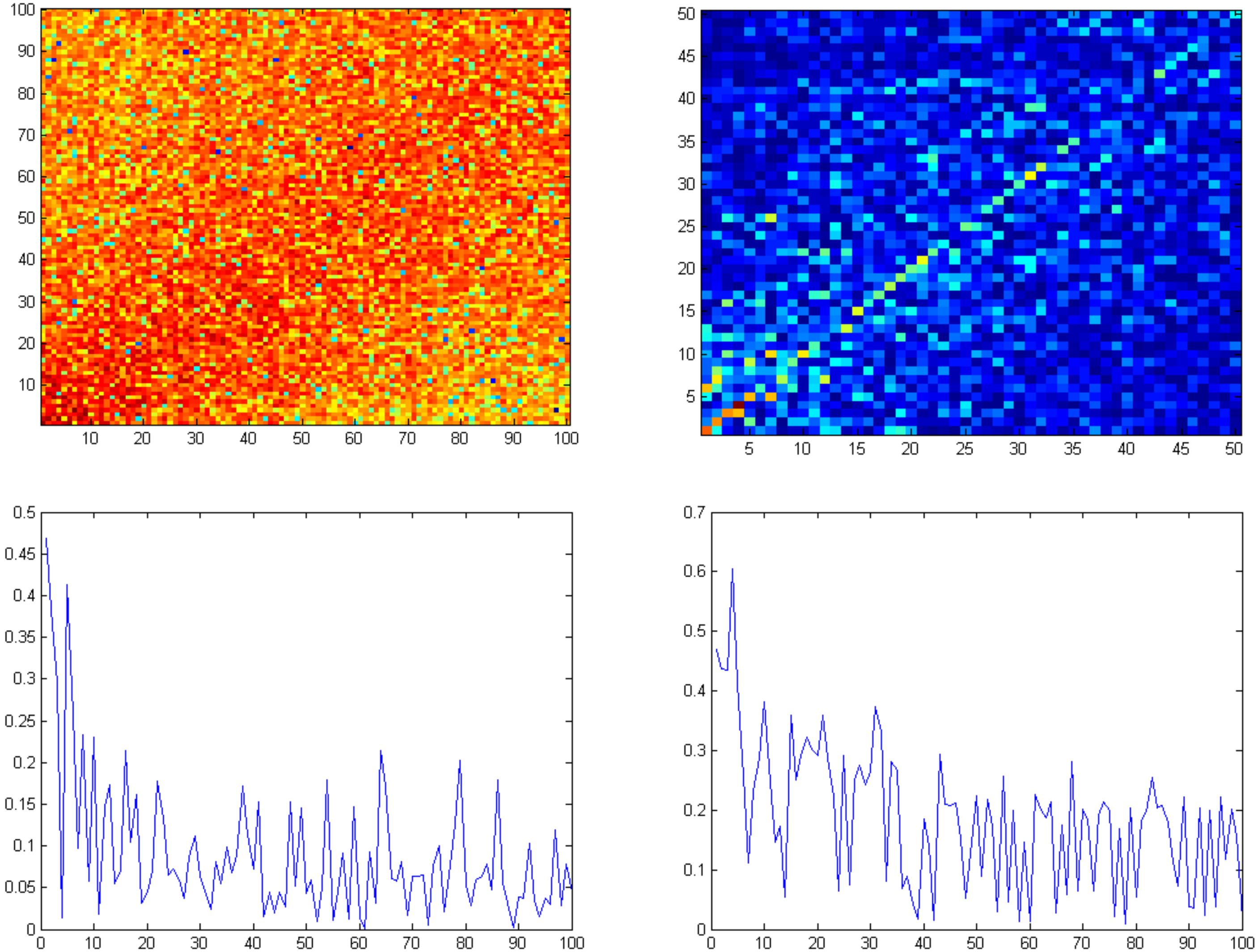
$$f_1(\mathbf{\Delta r}) = \frac{1}{\sqrt{(2\pi)^N}} e^{-\sum_i \Delta r_i^2/2}$$
$$\times \prod_i [1 + \sum_i \sum_{\nu=3}^{\infty} \frac{1}{\nu!} \langle H_\nu(\Delta r_i) \rangle H_\nu(\Delta r_i)]$$

The difference between full pdf and the approximation f1 is the mode-coupling correction such as $\langle H_p(\Delta r_i) H_{\nu-p}(\Delta r_j) \rangle - \langle H_p(\Delta r_i) \rangle \langle H_{\nu-p}(\Delta r_j) \rangle \neq 0$

$$f(\Delta r_i) \equiv \int_0^\infty \prod_{j\neq i} d\Delta r_j \; f(\mathbf{\Delta r})$$

Marginal distributions are transparent to such Corrections as a merit of the orthogonality relation.

Long-short simulation consistency



Experimentally verified residues

233 236 237 238 403 405 454 456 457 458 459 464 465 467 470 472 473 475 481
482 487 494 499 501 506 531 536 538 548 562 586 624 664 680 691 692 740 746