Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

A survey of coarse-grained methods for modeling protein conformational transitions

Wenjun Zheng and Han Wen

The decryption of transient structural changes during protein conformational transitions is essential to a detailed understanding of protein functions. To this end, coarse-grained protein structural models have proven valuable by enabling cost-effective simulation/analysis of protein conformational transitions which are too slow for all-atom molecular dynamics simulation. Here we survey state-of-the-art coarse-grained methods for protein conformational transition modeling developed in the past decade, with focus on those available online to public. We highlight the similarities and differences between these methods, and illustrate their usage in case of the T-to-R″ transition of chaperonin GroEL. This survey aims to provide researchers with a useful guide to the available tools for modeling protein conformational transitions.

Address
Physics Department, University at Buffalo, Buffalo, NY 14260, United States

Corresponding author: Zheng, Wenjun (wjzheng@buffalo.edu)

Current Opinion in Structural Biology 2017, **42**:24–30

This review comes from a themed issue on **Proteins: bridging theory and experiment**

Edited by **Igor N Berezovsky** and **Ugo Bastolla**

http://dx.doi.org/10.1016/j.sbi.2016.10.008

Introduction

The molecular functions of many protein complexes hinge upon their ability to undergo conformational transitions between various functional states via coordinated motions of functional domains. These conformational transitions span a wide range of time scales (from microseconds to seconds) and system sizes (from small single-domain proteins to large multi-domain/protein complexes). It has been a holy grail of molecular biophysics and structural biology to probe transient conformational changes between known structural states at high spatiotemporal resolution. This is an extremely challenging task because: first, these transient changes are often rare events that require exceedingly long time for computer simulation; second, the short-lived intermediates visited during conformational transitions are only sparsely populated, and thus very difficult to probe experimentally. Therefore, despite fast growth in computing technology, single-molecule biophysical tools [1,2], and time-resolved structural biology techniques [3–5], it remains very challenging to computationally simulate or experimentally probe the transient structural changes and intermediates in protein conformational transitions.

Various structural analysis tools are available for elucidating key features of protein conformational changes. Given multiple experimental structures of a protein in different functional states, one can readily identify dynamic domains undergoing rigid-body motions [6,7], and visualize a movie of physically feasible conformational path between two given end-point structures using various morphing techniques. These morphing methods are based on linear/nonlinear geometric interpolation of Cartesian or internal coordinates [8–10] or rigid-body rotations [11] or inter-residue distances [12,13,14•], so they tend to yield highly concerted domain motions as contrary to sequential domain motions. Despite their extensive usage, these analysis tools do not provide structural insights to the transient intermediates of protein conformational transitions and cannot realistically predict how individual domains move in a sequential or concerted fashion.

To directly explore protein conformational transitions at atomic resolution, molecular dynamics (MD) simulation [15] is the method of choice that uses molecular mechanics force fields to realistically simulate protein dynamics under physiological conditions. Nevertheless, all-atom MD simulations are computationally expensive and often limited to a short time scale up to hundreds of nanoseconds, although much longer (microsecond-millisecond) simulations have been achieved using massively parallelized or specially-designed supercomputers [16].

To efficiently probe long-time protein conformational transitions, a variety of coarse-grained (CG) models [17,18] have been developed based on simplified protein structural representations and potential energy functions at the expense of losing all-atom details. These simplifications have led to significantly reduced system size and computing cost for energy/force calculations. Notably, certain simplified energy functions (such as a harmonic potential) enables analytical solutions to the equation of motion and simulation-free analysis of protein dynamics. Consequently, CG models have been widely used to efficiently simulate/analyze

protein conformational transitions that are not accessible to all-atom MD simulation. Among various CG models, the Go model [19], constructed based on the native residue-residue contacts of a folded protein structure, has been used for protein folding/unfolding simulation for decades, and for exploring conformational transitions between two known protein structures [20,21]. Another popular CG model is the elastic network model (ENM), also known as the anisotropic network model (ANM) [22–24], which represents a protein structure as a network of C_α atoms with nearby ones connected by springs with a uniform force constant or distance-dependent force constant [25]. The ENM is routinely used to perform CG normal mode analysis (NMA) that solves a handful of low-frequency normal modes to describe collective domain motions involved in various large conformational changes as observed between different protein structures [26] (also see reviews [27,28]). Indeed, the ENM has formed the basis of many methods for modeling protein conformational transitions, which either use the ENM potential to construct and sample a multi-well energy landscape [29], or use the ENM-based normal modes to guide transition path generation [30] toward the target conformation.

In this review, we will survey recent development of CG methods for modeling protein conformational transitions based on relevant literature published in the past decade. We will focus on methods that use a CG physical potential function to generate a transition path from a given beginning structure to a given end structure. We will not consider transitions involving protein folding/unfolding, or unbiased CG simulations of a transition starting from an initial structure. We will place emphasis on those methods which are available online as public webservers or downloadable programs. Due to limited space and non-exhaustive literature search, we apologize for not covering some relevant methods in this short review.

Review of available conformational transition modeling methods (in the chronological order of publication)

MinActionPath (MAP) [31•]

This method defines two potential functions of Cα-only ENM for the beginning and the end structures which are expanded using the harmonic approximation ($E(X)\approx E(X_0) + (X^T - X_0^T)H(X - X_0)/2$), where E is the ENM potential energy, X/X_0 represents the Cartesian coordinates of the present/initial structure, H is the ENM Hessian matrix). Then it solves the Langevin equation analytically using the Onsager and Machlup action minimization formalism on each side of the transition. The crossover between the two ENM potentials is found numerically using an iterative approach, producing the most probable trajectory with transition state and energy determined as well. The MAP-produced transition path is reversible (i.e., unchanged after switching the beginning and the end structures). It exhibits less non-linearity than the other ENM-potential-based methods (see Figure 1(b)), which can be attributed to the use of harmonic approximation resulting in loss of higher-order terms in the potential energy functions.

mixed ENM (mENM) [32•]

This method solves the saddle points of a double-well potential function constructed from two ENM potential functions based at the beginning/end conformations of a transition using an exponential mixing scheme [20]. The saddle point equation is solved analytically as a linear equation after using the harmonic approximation to expand both ENM potentials to the second order (see above). The idea of mixing two single-well potentials into a double-well potential for conformational sampling was introduced in early studies of conformational transitions [20,29]. Similar to MAP, the mENM-predicted path is also reversible, and it exhibits greater non-linearity than MAP, but less non-linearity compared to those ENM-potential-based methods that do not use the harmonic approximation (see Figure 1(b)).

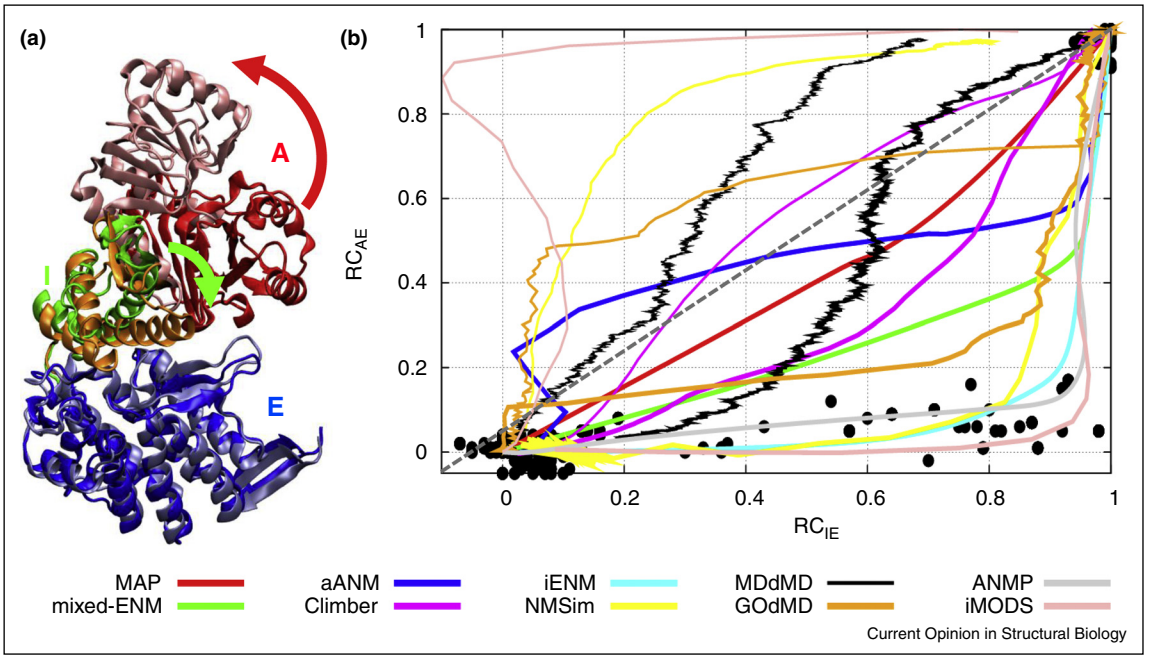
adaptive ANM (aANM) [33•]

This method starts simultaneously from both the beginning and the end structures by recruiting small subsets of ENM-based normal modes to create a series of intermediate conformations via an adaptive ANM methodology until the two intermediates merged within a predefined root mean square deviation (RMSD). The aANM parameters allow users to adjust the balance between two competing requirements for transition path generation: first, to minimize the deformation energy by following the lowest-frequency modes which often cause large deviation from the targeted direction; second, to minimize the length of transition path by moving along the targeted direction. The optimal balance between these two competing factors is likely protein-dependent, which complicates the use of aANM as a general method for modeling any given protein conformational transition. Like MAP and mENM, the aANM-generated paths are reversible. An aANM path generally exhibits a three-stage transition which initially follows the low-frequency modes of the beginning structure, and then enters an energy-barrier region before subsequently following the low-frequency modes of the end structure (see Figure 1(b)).

Climber [14•]

In this nonlinear morphing method, the inter-residue distances of a beginning conformation are pulled towards the distances in the end conformation using a set of harmonic restraints that are added to the internal energy function, which is minimized iteratively. The force constant of the restraint energy is self-adjusted to maintain a roughly constant speed in approaching the end conformation. Like other geometric interpolation methods, the goal of Climber is to generate physically feasible paths that go around (not over) high-energy barriers. Because

Figure 1



(a) Comparison of the T-state and the R″-state conformations of GroEL superimposed along the E domain: the E, I, and A domain in the T state are colored blue, green, and red; the E, I, and A domain in the R″ state are colored iceblue, orange, and pink; the motions of the I/A domains are marked by curved arrows colored in green/red. **(b)** Analysis of transition paths for the T-to-R″ transition of GroEL: two reaction coordinates (RC_{IE} and RC_{AE}) are calculated for intermediate conformations along the transition paths generated by MAP, mixed-ENM, aANM, Climber, iENM, NMSim, MDdMD, GOdMD, ANMP, and iMODS. The diagonal dash line corresponds to a linear interpolation path. RC_{IE} (RC_{AE}) quantifies the progress of conformational change of the I domain (the A domain) relative to the E domain — it is 0 for the T-state conformation and 1 for the R″-state conformation. The backward R″-to-T paths are also shown for Climber, NMSim, MDdMD, GOdMD, and iMODS (using thicker lines than the forward paths). The RC data points for 53 experimental structures of GroEL are shown as black dots (PDB ids: 1aon, 1dk7, 1dkd, 1fy9, 1fya, 1gr5, 1grl, 1gru, 1jon, 1kid, 1kp8, 1la1, 1mnf, 1oel, 1pcq, 1pf9, 1ss8, 1svt, 1sx3, 1sx4, 1xck, 2c7c, 2c7d, 2c7e, 2cgt, 2eu1, 2nwc, 2yey, 2ynj, 3c9v, 3cau, 3e76, 3qou, 3vz6, 3vz7, 3vz8, 3wvl, 3zpz, 3zq0, 3zq1, 4aaq, 4aar, 4aas, 4aau, 4ab2, 4ab3, 4hel, 4ki8, 4pkn, 4pko, 4wgl, and 4wsc).

this method pulls all distant pairs of Cα atoms, it can drive large concerted motions between distant domains like in linear interpolation. This is distinct from the ENM-potential-based methods which only pull those residue pairs that are in contact in either the beginning or the end structure. Therefore, a Climber-produced path exhibits intermediate non-linearity between a linear-interpolation path and the ENM-potential-based paths (see Figure 1(b)). The Climber-generated paths are not reversible.

interpolated-ENM (iENM) [34•]

This method is an improved version of the mixed ENM formulation [32•]. The key idea is to accurately solve the saddle points of a general double-well potential function constructed from two ENM potential functions based at the beginning and the end conformations of a transition. These saddle points form a minimal-energy path which is independent of the ENM and mixing parameters. Besides its generality, iENM has improved the accuracy of mixed ENM by accurately solving the saddle point equation without using the harmonic approximation, preserving covalent bonding between neighboring residues, and penalizing residue-residue collisions during a transition. This protocol is efficient thanks to the use of a sparse linear equation solver in place of a more expensive eigensolver. Similar to MAP and mENM, an iENM-predicted path is reversible. It exhibits more pronounced non-linearity than MAP and mENM (see Figure 1(b)).

NMSim [35•]

This method first decomposes a protein structure into rigid clusters and flexible regions using the FIRST algorithm [7], then solves normal modes using an ENM-based iterative rigid cluster NMA [36], and uses the low-frequency modes to guide constrained geometric simulations by biasing backbone motions toward a target conformation and side chain motions toward favorable rotamer states. Like other normal-modes-following methods, the NMSim-generated paths are biased by the normal modes of the beginning structure (not the end structure) and therefore not reversible (see Figure 1(b)).

MDdMD [37•]

This method constructs a transition path by using discrete molecular dynamics [38] and biasing techniques includ-

ing normal-modes-based essential dynamics and Maxwell–Demon sampling techniques. The MDdMD-generated paths are stochastic and not reversible, which differ significantly from the deterministic and reversible paths generated by the ENM-potential-based methods (MAP, mENM, iENM, etc, see Figure 1(b)).

GOdMD [39•]

Similar to MDdMD, this method uses discrete molecular dynamics [38] to sample protein conformational space. It combines a multi-minima Go-like potential energy function with enhanced sampling strategies such as normal-modes-biased metadynamics, Maxwell Demon molecular dynamics, and normal-modes-based essential dynamics. The GOdMD-generated paths are not reversible, and similar to other normal-modes-following paths (see Figure 1(b)) in being biased by the low-frequency modes of the beginning structure.

ANMPathway (ANMP) [40•]

This method is based on a two-state potential combining two ENM potentials of the beginning/end structures. It first locates the minimum energy structure on the cusp hypersurface as the transition state, and then follows the steepest descent trajectories from the transition state on each side of the cusp hypersurface. This method corresponds to a limiting case of the Plastic Network Model [29] with zero mixing between the two ENM potentials. An ANMP-predicted path is reversible, and exhibits pronounced non-linearity similar to the iENM path (see Figure 1(b)). Unlike iENM, an ANMP-produced path depends the location of the transition state (i.e., the saddle point) which is sensitive to the ENM parameters and energy offset between the two ENM potentials.

iMODS [41•]

This method uses NMA in internal (dihedral) coordinates to solve normal modes that capture collective motions while implicitly maintaining stereochemistry. To simulate a conformational transition, the beginning structure is iteratively deformed along the lowest modes while the RMSD to the end structure is minimized. The iMODS-generated paths are not reversible, and qualitatively similar to other normal-modes-following paths (see Figure 1(b)).

Review of other conformational transition modeling methods

Other normal-modes-following methods have been developed. The ANM Monte Carlo algorithm [42] generates a targeted path between two conformations, where the collective modes from ANM are used for iterative deformation, and the conformational energy of the deformed structure is minimized via a Monte Carlo algorithm. The coarse-grained virtual atom molecular mechanics algorithm [43] yields a transition path between two given structures by moving each structure toward the other in iterations of moves directed along the normal mode of greatest engagement with its target structure. Al-Bluwi *et al.* used a tripeptide-based-ENM to perform NMA and predict several collective modes, which are linearly combined for conformational exploration performed by a robot motion planning algorithm [44].

A geometric targeting method was developed for generating stereochemically acceptable transition paths in proteins [45]. It gradually changes the system's RMSD relative to the target structure while enforcing a set of geometric constraints. The generated paths are geometrically plausible by maintaining good covalent bond distances and angles, keeping backbone dihedral angles in allowed Ramachandran regions, avoiding eclipsed side-chain torsion angles, avoiding non-bonded overlap, and maintaining a set of hydrogen bonds and hydrophobic contacts.

A method named Path Similarity Analysis was developed to quantify the similarity and difference between two transition paths, and it was applied to compare a number of protein transition path generating algorithms including those based on molecular dynamics and ENM [46].

A hybrid elastic-network Brownian dynamics simulation method was proposed by extracting the transition routes from principal component analysis of structurally-rich ensembles and coarse-grained simulations [47]. This method was used to explore the conformational landscapes of five well-studied proteins, and predict the structures of intermediates along the paths.

Analysis of transition paths for the T-to-R″ transition of GroEL

The quality of conformational transition modeling can be assessed using various criteria: First, does the transition path preserve the chemical structures with minimal geometrical distortions or steric clashes? Second, does the transition path pass known intermediate conformations? Third, does the transition path correctly predict the order of domain motions during a transition? Here we will focus on the third criterion.

To quantify the domain motional order, we introduced the following reaction coordinate (RC) for an intermediate conformation of a given domain S: $RC_S = (\delta X_S \bullet \delta X_{S,obs})/\left|\delta X_{S,obs}\right|^2$, where δX_S is the displacement vector of S from the beginning conformation of a transition to a given intermediate conformation, and $\delta X_{S,obs}$ is the observed displacement of S from the beginning conformation to the end conformation of a transition. RC_S measures the motional progress of S in the direction of a transition. $RC_S = 0$ (1) at the beginning (end) of a transition. For two different domains (named S_1 and S_2) in an intermediate conformation, if

$RC_{S_1} > RC_{S_2}$, then S_1's movement precedes S_2's movement.

To illustrate and compare the above-reviewed methods, we chose as a test case the well-studied T-to-R″ transition in *Escherichia Coli*. chaperonin GroEL, which features large motions between three domains: a large twisting and upward displacement of the apical (A) domain and a downward movement of the intermediate (I) domain relative to the equatorial (E) domain [48] (see Figure 1(a)). We have applied the above methods to generate a forward transition path from the T-state conformation (PDB id: 1AON, chain H) to the R″-state conformation (PDB id: 1AON, chain A) of a GroEL subunit, together with a backward R″-to-T transition path. Then we have calculated the reaction coordinates (RC_{IE} and RC_{AE}) for the I & E domains (residues 2-135, 136-191, 373-410, and 411-525) and the A & E domains (residues 2-135, 192-372, and 411-525) to quantify the motional order between the I domain and the A domain relative to the E domain during the T-to-R″ transition. For validation, we calculated and plotted the RC data points for 53 experimental structures of *E. Coli*. GroEL (see Figure 1(b)). The main differences between these structures are in their nucleotide state and whether they are bound to the co-chaperonin GroES. They can be grouped into various states including the apo T state, the nucleotide-bound R state, and the nucleotide-GroES-bound R″ state. The R-state structures correspond to structural intermediates of the T-to-R″ transition.

The results are summarized as follows:

The ENM-potential-based methods (MAP, mENM, iENM, and ANMP) predict qualitatively similar and reversible paths, suggesting an early I-domain motion followed by a later A-domain motion in the T-to-R″ transition, and a reversed order for the backward R″-to-T transition (see Figure 1(b)). This is consistent with the distribution of experimental structures (see Figure 1(b)). However, they exhibit different level of non-linearity (i.e., deviation from the linear-interpolation path), with MAP showing minimal non-linearity while iENM and ANMP showing maximal non-linearity (see Figure 1(b)).

The ENM-modes-following methods (GOdMD, NMSim, and iMODS) predict qualitatively similar and irreversible paths. The forward and backward paths both suggest an early A-domain motion followed by a later I-domain motion (see Figure 1(b)). Only the backward path is consistent with the distribution of experimental structures (see Figure 1(b)).

Unlike the other ENM-based methods, the aANM-generated T-to-R″ path (using the optimal parameter $F_{min} = 0.5$[33•]) predicts a more complex three-stage T-to-R″ transition with the A domain moving first followed by the I domain motion and then the A domain motion again. At higher F_{min}, aANM is expected to behave more like a linear interpolation [33•].

MDdMD and Climber predict irreversible paths that deviate less from the linear-interpolation path than the other methods (except MAP, see Figure 1(b)). They suggest roughly concerted motions of the A domain and the I domain during the T-to-R″ and R″-to-T transitions.

Conclusion

In sum, we have reviewed ten CG methods for modeling protein conformational transitions which differ in reversibility and non-linearity. For those methods that generate irreversible paths, users should be careful in choosing which of the two end-point structures to start the simulation. In particular, when using those normal-modes-following methods, one should start from the structure whose low-frequency modes best capture the targeted conformational changes (such as the R″-state conformation of GroEL in the above test case). If the goal is to predict a sequence of domain motions during a transition, one should use those methods capable of producing highly non-linear paths (such as iENM and ANMP). To ensure the robustness of transition path modeling, it is recommended that multiple methods are utilized to make consensus-based predictions.

We caution that some complex issues may arise in the study of transition paths, including partial unfolding and existence of multiple/irreversible paths. If one of the two end-point structures is partially disordered, then those ENM-based methods may not be appropriate and other models that allow unfolding (such as the Go model) should be used. Such complex issues should be better addressed in future development of transition path modeling methods.

Although structural interpolation has been widely used by structural biologists for visualizing conformational changes between different structures, less effort has been made to obtain mechanistic insights from the predicted structural intermediates (e.g., predicting the sequence of domain motions [49]). The methods reviewed here promise to boost such effort in the future. We urge all structural biologists to make full use of these methods and provide feedback on whether they work or do not work. This will enable synergistic efforts for future development, refinement, and validation of these useful methods.

Acknowledgement

We thank funding support from American Heart Association (#14GRNT18980033) and National Science Foundation (#0952736).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest

1. Colomb W, Sarkar SK: **Extracting physics of life at the molecular level: a review of single-molecule data analyses**. *Phys Life Rev* 2015, **13**:107-137.

2. Ha T, Tinnefeld P: **Photophysics of fluorescent probes for single-molecule biophysics and super-resolution imaging**. *Annu Rev Phys Chem* 2012, **63**:595-617.

3. Schotte F, Lim M, Jackson TA, Smirnov AV, Soman J, Olson JS, Phillips GN Jr, Wulff M, Anfinrud PA: **Watching a protein as it functions with 150-ps time-resolved X-ray crystallography**. *Science* 2003, **300**:1944-1947.

4. Levantino M, Yorke BA, Monteiro DC, Cammarata M, Pearson AR: **Using synchrotrons and XFELs for time-resolved X-ray crystallography and solution scattering experiments on biomolecules**. *Curr Opin Struct Biol* 2015, **35**:41-48.

5. Kupitz C, Basu S, Grotjohann I, Fromme R, Zatsepin NA, Rendek KN, Hunter MS, Shoeman RL, White TA, Wang D *et al.*: **Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser**. *Nature* 2014, **513**:261-265.

6. Hayward S, Lee RA: **Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50**. *J Mol Graph Model* 2002, **21**:181-183.

7. Jacobs DJ, Rader AJ, Kuhn LA, Thorpe MF: **Protein flexibility predictions using graph theory**. *Proteins* 2001, **44**:150-165.

8. Krebs WG, Gerstein M: **The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework**. *Nucleic Acids Res* 2000, **28**:1665-1675.

9. Flores S, Echols N, Milburn D, Hespenheide B, Keating K, Lu J, Wells S, Yu EZ, Thorpe M, Gerstein M: **The database of macromolecular motions: new features added at the decade mark**. *Nucleic Acids Res* 2006, **34**:D296-D301.

10. Kleywegt GJ: **Use of non-crystallographic symmetry in protein structure refinement**. *Acta Crystallogr D Biol Crystallogr* 1996, **52**:842-857.

11. Ye Y, Godzik A: **FATCAT: a web server for flexible structure comparison and structure similarity searching**. *Nucleic Acids Res* 2004, **32**:W582-W585.

12. Kim MK, Jernigan RL, Chirikjian GS: **Efficient generation of feasible pathways for protein conformational transitions**. *Biophys J* 2002, **83**:1620-1630.

13. Seo S, Kim MK: **KOSMOS: a universal morph server for nucleic acids, proteins and their complexes**. *Nucleic Acids Res* 2012, **40**:W531-W536.

14. • Weiss DR, Levitt M: **Can morphing methods predict intermediate structures?** *J Mol Biol* 2009, **385**:665-674.
This study presented a nonlinear morphing method named Climber in comparison with other morphing methods, and it is available for download at https://simtk.org/projects/climber.

15. Karplus M, McCammon JA: **Molecular dynamics simulations of biomolecules**. *Nat Struct Biol* 2002, **9**:646-652.

16. Shaw DE, Grossman JP, Bank JA, Batson B, Butts JA, Chao JC, Deneroff MM, Dror RO, Even A, Fenton CH *et al.*: **Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer**. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis.*; *New Orleans, Louisana: IEEE Press; 2014:*41-53 http://dx.doi.org/10.1109/sc.2014.9.

17. Tozzini V: **Coarse-grained models for proteins**. *Curr Opin Struct Biol* 2005, **15**:144-150.

18. Kmiecik S, Gront D, Kolinski M, Wieteska L, Dawid AE, Kolinski A: **Coarse-grained protein models and their applications**. *Chem Rev* 2016, **116**:7898-7936.

19. Go N: **Theoretical studies of protein folding**. *Annu Rev Biophys Bioeng* 1983, **12**:183-210.

20. Best RB, Chen YG, Hummer G: **Slow protein conformational dynamics from multiple experimental structures: the helix/sheet transition of arc repressor**. *Structure* 2005, **13**:1755-1763.

21. Koga N, Takada S: **Folding-based molecular simulations reveal mechanisms of the rotary motor F1-ATPase**. *Proc Natl Acad Sci U S A* 2006, **103**:5367-5372.

22. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I: **Anisotropy of fluctuation dynamics of proteins with an elastic network model**. *Biophys J* 2001, **80**:505-515.

23. Tama F, Sanejouand YH: **Conformational change of proteins arising from normal mode calculations**. *Protein Eng* 2001, **14**:1-6.

24. Zheng W, Doniach S: **A comparative study of motor-protein motions by using a simple elastic-network model**. *Proc Natl Acad Sci U S A* 2003, **100**:13253-13258.

25. Yang L, Song G, Jernigan RL: **Protein elastic network models and the ranges of cooperativity**. *Proc Natl Acad Sci U S A* 2009, **106**:12347-12352.

26. Krebs WG, Alexandrov V, Wilson CA, Echols N, Yu H, Gerstein M: **Normal mode analysis of macromolecular motions in a database framework: developing mode concentration as a useful classifying statistic**. *Proteins* 2002, **48**:682-695.

27. Bahar I, Rader AJ: **Coarse-grained normal mode analysis in structural biology**. *Curr Opin Struct Biol* 2005, **15**:586-592.

28. Tama F, Brooks CL: **Symmetry, form, and shape: guiding principles for robustness in macromolecular machines**. *Annu Rev Biophys Biomol Struct* 2006, **35**:115-133.

29. Maragakis P, Karplus M: **Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase**. *J Mol Biol* 2005, **352**:807-822.

30. Miyashita O, Onuchic JN, Wolynes PG: **Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins**. *Proc Natl Acad Sci U S A* 2003, **100**:12570-12575.

31. • Franklin J, Koehl P, Doniach S, Delarue M: **MinActionPath: maximum likelihood trajectory for large-scale structural transitions in a coarse-grained locally harmonic energy landscape**. *Nucleic Acids Research* 2007, **35**:W477-W482.
This study presented a transition path modeling method based on ENM and action minimization, and it is available at the following webserver: http://lorentz.dynstr.pasteur.fr/joel/index.php.

32. • Zheng W, Brooks BR, Hummer G: **Protein conformational transitions explored by mixed elastic network models**. *Proteins* 2007, **69**:43-57.
This study presented a transition path modeling method based on the approximate solution of saddle points for a mixed ENM potential, and it is available at the following webserver: http://enm.lobos.nih.gov/start_path.html.

33. • Yang Z, Majek P, Bahar I: **Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL**. *PLoS Computat Biol* 2009, **5**.
This study presented a transition path modeling method that recruits a subset of normal modes to create a series of intermediate conformations via an adaptive ANM method, and it is available for download at: https://simtk.org/frs/index.php?group_id=531.

34. • Tekpinar M, Zheng W: **Predicting order of conformational changes during protein conformational transitions using an interpolated elastic network model**. *Proteins-Struct Funct Bioinf* 2010, **78**:2469-2481.
This study presented a transition path modeling method based on the accurate solution of saddle points for a general double-well potential, and it is available at the following webserver: http://enm.lobos.nih.gov/start_ienm.html.

35. • Krueger DM, Ahmed A, Gohlke H: **NMSim Web Server: integrated approach for normal mode-based geometric simulations of biologically relevant conformational transitions in proteins**. *Nucleic Acids Res* 2012, **40**:W310-W316.
This study presented a transition path modeling method based on rigid cluster decomposition, rigid cluster normal mode analysis, and constrained geometric simulation, and it is available at the following webserver: http://www.nmsim.de.

36. Ahmed A, Gohlke H: **Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory**. *Proteins* 2006, **63**:1038-1051.

37. • Sfriso P, Emperador A, Orellana L, Hospital A, Gelpi JL, Orozco M: **Finding conformational transition pathways from discrete molecular dynamics simulations**. *J Chem Theory Comput* 2012, **8**:4707-4718.
This study presented a transition path modeling method based on discrete molecular dynamics, essential dynamics, and Maxwell-Demon sampling, and it is available at the following webserver: http://mmb.irbbarcelona.org/MDdMD/

38. Proctor EA, Dokholyan NV: **Applications of discrete molecular dynamics in biology and medicine**. *Curr Opin Struct Biol* 2016, **37**:9-13.

39. • Sfriso P, Hospital A, Emperador A, Orozco M: **Exploration of conformational transition pathways from coarse-grained simulations**. *Bioinformatics* 2013, **29**:1980-1986.
This study presented a transition path modeling method based on discrete molecular dynamics, Go-like potential energy function, metadynamics, essential dynamics, and Maxwell-Demon sampling, and it is available at the following webserver: http://mmb.irbbarcelona.org/GOdMD.

40. • Das A, Gur M, Cheng MH, Jo S, Bahar I, Roux B: **Exploring the conformational transitions of biomolecular systems using a simple two-state anisotropic network model**. *PLoS Comput Biol* 2014, **10** e1003521.
This study presented a transition path modeling method based on the steepest descent trajectories initiated from a transition state located at the cusp of two ENM potentials, and it is available at the following webserver: http://anmpathway.lcrc.anl.gov/anmpathway.cgi/.

41. • Ramon Lopez-Blanco J, Aliaga JI, Quintana-Orti ES, Chacon P: **iMODS: internal coordinates normal mode analysis server**. *Nucleic Acids Res* 2014, **42**:W271-W276.
This study presented a transition path modeling method based on normal mode analysis of internal coordinates, and it is available at the following webserver: http://imods.chaconlab.org.

42. Kantarci-Carsibasi N, Haliloglu T, Doruker P: **Conformational transition pathways explored by Monte Carlo simulation integrated with collective modes**. *Biophys J* 2008, **95**:5862-5873.

43. Korkut A, Hendrickson WA: **Computation of conformational transitions in proteins by virtual atom molecular mechanics as validated in application to adenylate kinase**. *Proc Natl Acad Sci U S A* 2009, **106**:15673-15678.

44. Al-Bluwi I, Vaisset M, Simeon T, Cortes J: **Modeling protein conformational transitions by a combination of coarse-grained normal mode analysis and robotics-inspired methods**. *Bmc Struct Biol* 2013:13.

45. Farrell DW, Speranskiy K, Thorpe MF: **Generating stereochemically acceptable protein pathways**. *Proteins* 2010, **78**:2908-2921.

46. Seyler SL, Kumar A, Thorpe MF, Beckstein O: **Path similarity analysis: a method for quantifying macromolecular pathways**. *PLoS Comput Biol* 2015, **11** e1004568.

47. Orellana L, Yoluk O, Carrillo O, Orozco M, Lindahl E: **Prediction and validation of protein intermediate states from structurally rich ensembles and coarse-grained simulations**. *Nat Commun* 2016, **7**:12575.

48. Xu Z, Horwich AL, Sigler PB: **The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex**. *Nature* 1997, **388**:741-750.

49. Zheng W, Qin F: **A combined coarse-grained and all-atom simulation of TRPV1 channel gating and heat activation**. *J Gen Physiol* 2015, **145**:443-456.