Journal of the Mechanics and Physics of Solids 83 (2015) 36–47

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

journal homepage: www.elsevier.com/locate/jmps

Geometric analysis characterizes molecular rigidity in generic and non-generic protein configurations

Dominik Budday [a], Sigrid Leyendecker [a], Henry van den Bedem [b,*]

[a] Chair of Applied Dynamics, University of Erlangen-Nuremberg, Haberstrasse 1, 91058 Erlangen, Germany
[b] Division of Biosciences, SLAC National Accelerator Laboratory, Stanford University, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form
8 June 2015
Accepted 13 June 2015
Available online 18 June 2015

Keywords:
Protein
Rigidity
Nullspace
Singular configuration
Protein collective motions
Conformational sampling

ABSTRACT

Proteins operate and interact with partners by dynamically exchanging between functional substates of a conformational ensemble on a rugged free energy landscape. Understanding how these substates are linked by coordinated, collective motions requires exploring a high-dimensional space, which remains a tremendous challenge. While molecular dynamics simulations can provide atomically detailed insight into the dynamics, computational demands to adequately sample conformational ensembles of large biomolecules and their complexes often require tremendous resources. Kinematic models can provide high-level insights into conformational ensembles and molecular rigidity beyond the reach of molecular dynamics by reducing the dimensionality of the search space. Here, we model a protein as a kinematic linkage and present a new geometric method to characterize molecular rigidity from the constraint manifold Q and its tangent space $\mathcal{T}_{\mathbf{q}}Q$ at the current configuration $\mathbf{q}$. In contrast to methods based on combinatorial constraint counting, our method is valid for both generic and non-generic, e.g., singular configurations. Importantly, our geometric approach provides an explicit basis for collective motions along floppy modes, resulting in an efficient procedure to probe conformational space. An atomically detailed structural characterization of coordinated, collective motions would allow us to engineer or allosterically modulate biomolecules by selectively stabilizing conformations that enhance or inhibit function with broad implications for human health.

1. Introduction

A protein is a linear sequence of amino acids or residues, synthesized into a polypeptide chain by the ribosome (Fig. 1a, b). The function of a protein is largely dictated by its folded, three-dimensional structure, which determines its ability to bind to other molecules, such as small ligands, other proteins, or nucleic acids (Donald, 2011). Advances in imaging technology such as X-ray crystallography, nuclear magnetic resonance spectroscopy or cryo-electron microscopy increasingly enable rapid characterization of biological macromolecules in atomic detail. The Protein Data Bank (PDB), an international repository of the three-dimensional coordinates of protein, RNA, and DNA, now contains over 100,000 structures (Berman

* Corresponding author. Fax: +1 650 926 3293.
E-mail addresses: dominik.budday@fau.de (D. Budday), sigrid.leyendecker@ltd.uni-erlangen.de (S. Leyendecker), vdbedem@stanford.edu (H. van den Bedem).
URL: https://people.stanford.edu/vdbedem/ (H. van den Bedem).

http://dx.doi.org/10.1016/j.jmps.2015.06.006
0022-5096/

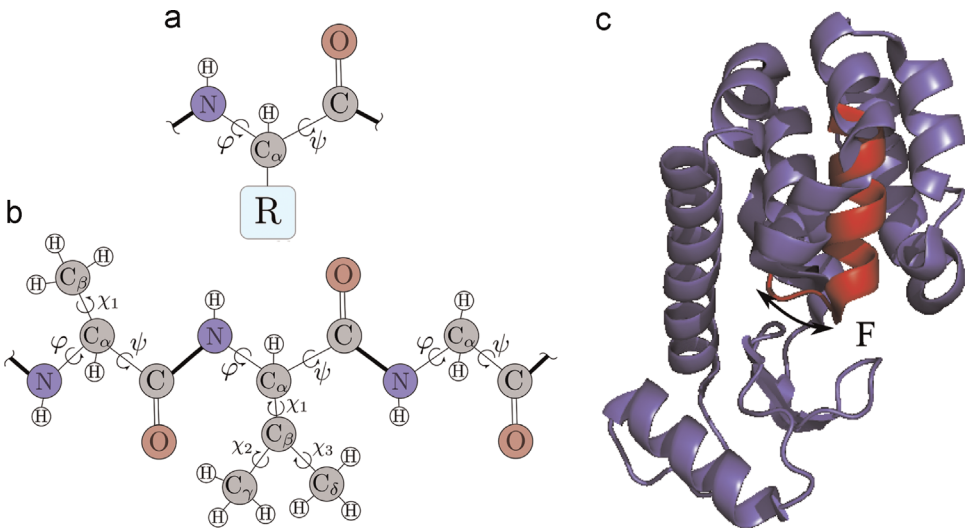

Fig. 1. A protein is a polypeptide chain folded into a three-dimensional shape. (a) An amino acid or residue consists of a fixed backbone, and one of 20 side-chains, indicated by R, covalently bound to the C_α backbone atom. (b) A polypeptide chain is a kinematic linkage, with groups of atoms as rigid bodies and covalent, rotatable bonds as joints with a revolute degree of freedom (φ, ψ, χ). (c) A mutant T4 lysozyme can exchange between a low energy ground state, and a sparsely populated excited state (Bouvignies et al., 2011). Helix F collectively rearranges between the ground state (blue) and the excited state (red). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

et al., 2000). However, proteins fluctuate between conformational substates spanning a wide range of spatiotemporal scales to perform their cellular function and engage with partners (Fig. 1c). These motions range from pico-second timescale atomic vibrations to diffusive, collective motion at millisecond or longer timescales often associated with biological activity (van den Bedem and Fraser, 2015). Despite enormous advances in experimental techniques, we cannot directly observe biomolecular, spatiotemporal ensembles. Characterizing these exchanges and understanding how different parts of proteins are dynamically coupled through collective motions can tremendously benefit human health: It would allow us to engineer or allosterically modulate biomolecules by selectively stabilizing conformations that enhance or inhibit function.

Experimental techniques have significant potential to uncover a molecular basis for protein conformational dynamics. While X-ray crystallography experiments mostly yield a single, low-energy ground state of the molecule, nuclear magnetic resonance relaxation dispersion experiments can provide insight into functionally relevant excited states, but lack a structural basis for collective motions. Computationally integrating these data sources has proved challenging (Fenwick et al., 2014; Fonseca et al., 2014). Molecular dynamics simulations can yield atomically detailed trajectories, but rely on imperfect force-fields and often demand specialized hardware (Hein et al., 2005) and algorithms to examine long, biologically relevant time scales or larger molecules (Klepeis et al., 2009). By contrast, non-deterministic conformational sampling-based algorithms, such as kinematics-based methods, can provide high-level insights into conformational ensembles at spatiotemporal scales beyond the reach of molecular dynamics simulations (Fonseca et al., 2014, 2015; Pachov et al., 2015).

Kinematics-based methods exploit that the linear, branched topology of a biomolecule closely resembles kinematic truss structures. These methods represent a protein or nucleic acid as a kinematic linkage with groups of atoms as rigid bodies and covalent, rotatable bonds as joints with a revolute degree of freedom (Fig. 1a, b). Hydrogen bonds and other non-covalent interactions are encoded as holonomic constraints, resulting in nested, interdependent cycles that require coordinated changes of the degrees of freedom, effectively reducing the dimensionality of configuration space. The remaining motions are known as floppy modes and yield collective motion of the degrees of freedom in a lower-dimensional constraint manifold Q (Burdick, 1989; Thorpe et al., 2001; van den Bedem et al., 2005; Yao et al., 2012). The constraints reduce conformational flexibility or can even completely rigidify larger substructures of biomolecules by merging rigid bodies through rotationally locked degrees of freedom or hydrogen bonds. Configuration space, i.e., the set of all degrees of freedom, is sometimes denoted as conformation space when applied to proteins.

In generic, e.g., non-singular configurations rigidity is a topological property, which is characterized completely by combinatorial, explicit constraint counting using an exact, graph theoretical 'pebble game' algorithm (Jacobs and Hendrickson, 1997; Jacobs and Thorpe, 1995). However, the pebble game fails to recognize additional flexibility resulting from special geometries like singular or symmetric configurations. In these non-generic situations, rigidity is a geometric property that cannot be characterized by combinatorial methods. While singularities form a non-dense subset of configuration space (Hendrickson, 1992), biomolecules could exploit specific characteristics of non-genericity such as increased instantaneous mobility (Zlatanov et al., 1994), a change of motion pattern (Wenger and Chablat, 1998) or large motions along emerging hinge axes to control accessibility of substates. Many biomolecules possess structural symmetries that allow geometrically concerted motions (Schulze et al., 2014; Jagodzinski et al., 2013).