Singular Value Decomposition of Protein Conformational Motions: Application to HIV-1 Protease

Miguel L. Teodoro[1] **George N. Phillips Jr.**[1] **Lydia E. Kavraki**[2]
mteodoro@rice.edu georgep@rice.edu kavraki@rice.edu

[1] W. M. Keck Center for Computational Biology and Department of Biochemistry and Cell Biology, Rice University, 6100 Main St - MS150, Houston, Texas 77005-1892, USA

[2] Department of Computer Science, Rice University, 6100 Main St - MS132, Houston, Texas 77005-1892, USA

Keywords: protein flexibility, drug design, molecular dynamics, singular value decomposition, HIV-1 protease

1 Introduction

Protein conformational motions play a critical role in biochemical catalysis. Through anharmonic conformational deformations during the course of its function a protein can change the chemical environment of its reactive site in order to bind substrates and catalyze reactions. An illustrative example of these changes is the opening and closing of the binding site of HIV-1 protease (HIV Pr) that results from conformational changes in the flaps region covering the reactive site. HIV Pr is a protease which plays a vital role in the maturation of the HIV-1 virus by targeting amino acid sequences in the gag and gag-pol polyproteins. Cleavage of these polyproteins produces proteins that contribute to the structure of the virion, RNA packaging, and condensation of the nucleoprotein core. Due to its large medical importance this protein has been studied extensively by experimental as well as computational methods and has been a successful target for rational drug design.

In this work we analyzed a molecular dynamics trajectory (1.4 ns) of a fully hydrated HIV Pr homodimer using the Singular Value Decomposition (SVD) method. Given the large number of degrees of freedom in a protein, the best way to analyze conformational motions is to obtain a representation for the dynamics trajectory in a lower dimensional basis. One method to achieve this simplification is to use the SVD method to perform a change of basis from individual cartesian degrees of freedom for each atom in the protein to a basis formed by an equal number of collective modes of motion. Using this formulation the conformational space accessible to a protein is separated into two subspaces: (1) an essential subspace containing only a few degrees of freedom which correspond to major modes of anharmonic motion and describe most of the positional fluctuations; and (2) a nonessential subspace consisting of constrained harmonic motions [1]. By using only the major degrees of freedom in the essential subspace of the protein, we were able to clearly identify stable protein conformational substates. The method presented here can be generalized to other biomolecular systems and used not only as a novel method in identifying new target conformations for drug design but also to simplify the analysis and understanding of important structural motions.

2 Method

Molecular dynamics simulations were performed using the the program NAMD [3]. The simulations were started from the crystal structure of the MVT-101 bound conformation (PDB entry 4hvp) without the presence of ligand. This system was inserted into a box of waters and the simulation was run for 1.4 ns using current state of the art methodology. The time dependent coordinates of the Cα atoms were used to construct a displacement matrix which is constructed by the column-wise concatenation of atomic displacement vectors, for each time sample during the molecular dynamics run. The SVD calculations were performed as described previously [2]. The left singular vectors obtained from the decomposition correspond to modes of collective motion whose displacements are directly proportional to the value of the corresponding singular values. The calculated right singular vectors correspond to the projection of the original trajectory in the new basis.

3 Results and Discussion

During the simulation we observed that starting from the closed conformation reported in the X-ray structure we were able to observe an opening of the binding site which we attribute to the absence of ligand in our simulations. The opening of the binding site proceeded through a small rearrangement of the protein core and a large rearrangement of the flaps region covering the binding site.

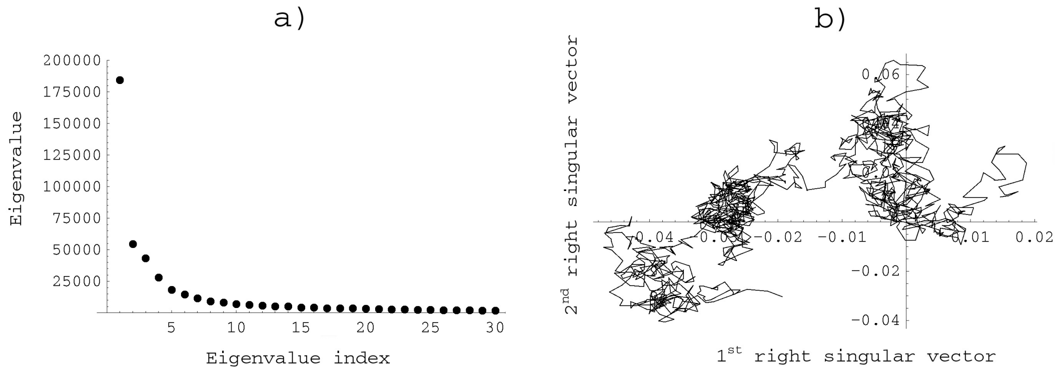

Figure 1: a) Rank ordered eigenvalue spectrum resulting from the SVD analysis of the dynamics trajectory. b) Configuration space projection of the entire simulation along the two left singular vectors corresponding to the two largest eigenvalues shows three distinct conformational substates.

In Figure 1-a) we can observe that the eigenvalue spectrum for the SVD decomposition is clearly dominated by a small number of left singular vectors. The largest eigenvalue accounts for 34.7% of the cumulative eigenvalue sum and the first 20 eigenvalues (out of a total of 597) account for 80%. The left singular vector, corresponding to the largest eigenvalue, indicates the direction of motion in the essential subspace for which the fluctuations in atomic positions are largest. When this collective mode of motion in the essential subspace is mapped to the protein structure, it is clearly observed that the mode corresponds to the opening and closing of the flaps covering the binding site.

Figure 1-b) shows that three conformational substates can be observed during the dynamics simulation. We propose that the opening of the binding site of HIV Pr occurs by a series of steps in which the protein jumps rapidly between stable intermediate conformations until it reaches the open conformation. The intermediate conformational substates should be useful as extra targets in rational drug design methods for the development of improved inhibitors of HIV Pr.

References

[1] Amadei, A., Linssen, A.B.M. and Berendsen, H.J.C., Essential dynamics of proteins, *Proteins: Structure, Function, and Genetics*, 17:412–425, 1993.

[2] Andrews, B.K., Romo, T., Clarage, J.B., Pettitt, B.M. and Phillips, G.N., Characterizing global substates of myoglobin, *Structure*, 6(5):587–594, 1998.

[3] Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. and Schulten, K., NAMD2: Greater scalability for parallel molecular dynamics, *Journal of Computational Physics*, 151:283-312, 1999.