Studying the Unfolding Kinetics of Proteins under Pressure Using Long Molecular Dynamics Simulation Runs

Oswaldo Chara · José Raúl Grigera ·
Andrés N. McCarthy

Received: 8 November 2007 / Accepted: 22 April 2008 /
Published online: 1 July 2008
© Springer Science + Business Media B.V. 2008

Abstract The usefulness of computational methods such as molecular dynamics simulation has been extensively established for studying systems in equilibrium. Nevertheless, its application to complex non-equilibrium biological processes such as protein unfolding has been generally regarded as producing results which cannot be interpreted straightforwardly. In the present study, we present results for the kinetics of unfolding of apomyoglobin, based on the analysis of long simulation runs of this protein in solution at 3 kbar (1 atm = 1.01325, bar = 101 325 Pa). We hereby demonstrate that the analysis of the data collected within a simulated time span of 0.18 μ s suffices for producing results, which coincide remarkably with the available unfolding kinetics experimental data. This not only validates molecular dynamics simulation as a valuable alternative for studying non-equilibrium processes, but also enables a detailed analysis of the actual structural mechanism which underlies the unfolding process of proteins under elusive denaturing conditions such as high pressure.

Keywords Molecular dynamics · Unfolding kinetics · Apomyoglobin · Two-state model · Mathematical modelling

O. Chara · J. R. Grigera · A. N. McCarthy (✉)
Instituto de Física de Líquidos y Sistemas Biológicos,
CCT La Plata CONICET, UNLP, CIC, La Plata, Argentina
e-mail: amccarthy@iflysib.unlp.edu.ar

O. Chara
Departamento de Fisiología y Biofísica, Facultad de Medicina,
Universidad de Buenos Aires, Buenos Aires, Argentina

J. R. Grigera · A. N. McCarthy
Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas,
Universidad Nacional de La Plata, 49-789, cc 565,
B1900BTE, La Plata, Argentina

1 Introduction

Molecular dynamics has been used in the past to extract kinetic information regarding a variety of processes. Some recent applications include, among others, studies related to water penetrating into proteins [1], the formation and disintegration of spherical micelles [2], water movements in carbon nanotubes [3] and protein folding [4, 5].

The usefulness of computational methods such as molecular dynamics simulations has been extensively established for studying systems in equilibrium. Nevertheless, its application to complex non-equilibrium biological processes has been generally regarded as producing results which cannot be interpreted straightforwardly. A clear example is the non-equilibrium phenomenon of the protein unfolding problem, which constitutes one of the most interesting non-equilibrium phenomena in biology. While the process of protein folding is crucial for achieving native conformation of proteins, protein unfolding is a key step in several cellular processes such as protein degradation, where susceptibility to unfolding of substrate proteins becomes essential, as well as translocation across some membranes [6]. Although protein unfolding has previously been studied using molecular dynamics [7], certain aspects of the process remain yet unsolved. In fact, the times involved in experiments suggest that protein unfolding is of the order of 10^{-1} ms [8]. In contrast, molecular dynamics simulations of proteins usually do not exceed the order of a few hundred nanoseconds [9]. On the other hand, it is important to have experimental information regarding the unfolding process that the studied protein undergoes. In this sense, there is considerable information related to the unfolding of apomyoglobin, from acid denaturation [10] to pressure denaturation [11]. Apomyoglobin is produced by removing the prosthetic (non-proteic heme) group from myoglobin. This molecule is a small, all- α -helical, globular protein, which constitutes an ideal model for the study of protein folding (Fig. 1).

It was previously reported that molecular dynamics simulations of apomyoglobin in solution effectively show the beginning of the denaturation process under a pressure of 3 kbar (1 atm = 1.01325 bar = 101 325 Pa) within a time window of 60 ns [9]. However, this time span is several orders of magnitude shorter than the experimentally determined time for the unfolding of apomyoglobin.

Thus, expansion of this time window is important to gain insight into the dynamics of the pressure denaturation of apomyoglobin. Hence, in this work we have utilized 182-ns-long simulations of apomyoglobin in solution at 1 bar and 3 kbar. Such data have allowed us to follow in great detail various aspects of the unfolding process of apomyoglobin under pressure.

The process of protein folding/unfolding may be followed by a number of structural parameters such as (among others) the protein's radius of gyration or solvent accessible surface (SAS) [9, 12–15], all of which are related to (and give a global account of) the general tertiary structure of the protein.

It has been recently proposed that the main driving force for protein denaturation under pressure is the loss of the hydrophobic effect (due to the drastic changes in the structure of the water bulk), which in turn results in an increase in exposure to solvent of the hydrophobic residues, thus causing the exposure of the mainly hydrophobic core, ultimately causing a loss of the general native state tertiary structure. Hence, we have chosen SAS as the parameter to track the unfolding process under pressure.

We have analysed these long simulation runs assuming the simplest mathematical model to explain the results: a two-state model (folded–unfolded) with first-order kinetic

Fig. 1 3D representation of the solvent accessible surface (SAS) of apomyoglobin at: **a)** 1 bar; and **b)** 3 kbar, **c)** fitted ribbon structures for apomyoglobin at 1 bar (white) and 3 kbar (RMSD grey-coded)

behaviour. For this, software dedicated to simulate mathematical model's behaviour as well as its fit to the simulated data extracted from molecular dynamics was explicitly developed.

We hereby demonstrate that the analysis of the data collected within a simulated time span of 0.18 μ s suffices to produce results that coincide remarkably with the available unfolding kinetics experimental data [8]. This not only validates molecular dynamics simulation as a valuable alternative for studying non-equilibrium processes, but also enables a detailed analysis of the actual structural mechanism that underlies the unfolding process of proteins under elusive denaturing conditions such as high pressure.

2 Computational Methods

2.1 Molecular Dynamics Simulations

We carried out molecular dynamics (MD) simulations using the GROMACS 3.2.1 package [16–18] in which the equations of motion are solved using a leap-frog integration step, with

a step size of 2 fs. We used an all-atoms force field [19–24] for the minimization process and for all the MD simulation steps and kept all protein bond lengths constrained using the Linear Constraint Solver algorithm [25]. Water molecules were constrained using the SETTLE algorithm [26]. For the calculation of electrostatic forces we applied the Reaction Field method. Lennard–Jones interactions were calculated within a cutoff radius of 1.4 nm.

For all the simulation runs we have used a Xeon based, dual-processor cluster, running under GNU/Linux, and, for all plots and graphics, MS Windows or GNU/Linux, using the reference Visual Molecular Dynamics package, Swiss PDB Viewer or XGrace software.

As starting configurations we have used the high- and low-pressure crystal structures *Sperm Whale myoglobin* (PDB codes 1JP8 and 1VXD, respectively) readily available at the Protein Data Bank, and generated the topology using the PDB2GMX tool, with standard pH 5.0 amino acid protonation states and removing all non-protein residues. The SPC/E [27] water model was used for both high- and low-pressure systems. The starting system consisted of a cubic simulation box of $X = 7.19617$ nm, $Y = 7.19617$ nm and $Z = 7.19617$ nm, with a total volume of $372,653$ nm³, containing one apomyoglobin molecule and 11,663 water molecules.

This system was energy-minimised by first using the steepest descent method, converging to machine precision. Second, the conjugated gradient method was applied, converging in less than 20 cycles. After this, the solvent was allowed to relax, running a short 10-ps MD simulation with position restraints applied on all the protein atoms.

At the start the system was weakly coupled to a thermal and hydrostatic bath in order to work in the isothermal–isobaric ensemble [28] at $T = 300$ K and $P = 1$ bar. Subsequently, the whole system was allowed to relax (restraints were removed from all atoms) during 200 ps, coupled to the same weak thermal and hydrostatic bath.

At this point we divided the system in two. The first was slowly submitted to increasing hydrostatic pressure, while the second was kept at atmospheric pressure.

The pressure in the first system was elevated during a total time of 2 ns, using a low relaxation constant value for the pressure bath ($\tau_P = 400$ ps). The second system was allowed to run for an equivalent 2-ns MD simulation period at 1 bar.

Having reached this point we allowed both systems to evolve during 8 ns each, in order to produce comparably equilibrated starting points for the low- and high-pressure systems.

Once this was achieved, we performed a 180-ns-long MD simulation for both the 1 bar and 3 kbar systems. All of the subsequent analysis corresponds to the data produced by these 180-ns-long MD simulations.

All plots were calculated using the corresponding GROMACS package programs. The Connolly SAS was calculated using the `g_sas` program [29] with a probe radius of 1.4 Å.

2.2 Mathematical Modelling and Fitting

A two-state model was developed and simulations were performed using custom-made software developed in Fortran 95 code. The procedure results in a prediction of the time course of solvent-accessed surface (SAS) according to the two-state model (see “[Results and Discussion](#)”). The software allows model-dependent fitting over the results extracted from the molecular dynamics trajectories to obtain the best-fitted parameter values of the model. The software and the source code are available.

3 Results and Discussion

Molecular dynamics simulations of apomyoglobin in solution were carried out at 1 bar and 3 kbar during 182 ns. The structures evaluated at the end of the trajectories show clearly the structural differences between them (Fig. 1c, white versus grey, corresponding to the protein at 1 bar and 3 kbar, respectively).

To study the putative unfolding kinetics of apomyoglobin at 1 bar and 3 kbar, the SAS was calculated for each simulation. To do that, the program `g_sas` from the GROMACS package was employed. Representations of these areas showed that at 3 kbar an unfolding process occurred, which was absent at 1 bar (Fig. 1a, b). Figure 2 shows the time evolution of SAS in both conditions. The results for the apomyoglobin SAS at 1 bar evinces no significant variation from the initial value (Fig. 2, dashed black curve). On the contrary, at 3 kbar, there is a clear monotonic behaviour (Fig. 2, dashed grey curve) in which $\frac{dSAS}{dt} > 0$.

Fig. 2 Nanosecond timescale representation of the SAS data for apomyoglobin at: 1 bar (dashed black line) and 3 kbar (dashed grey line), as well as the average for 1 bar (continuous black line) and the nonlinear regression for 3 kbar (continuous grey line). *Inset:* Microsecond timescale representation of the SAS-extrapolated nonlinear regression for 3 kbar (continuous grey curve), as well as the average for 1 bar (continuous black line). Experiments show values for kinetic constants of the folding–unfolding equilibrium process in the sub-millisecond time scale [8]. The model considered here bridges the gap between simulation and experiment, showing extrapolated unfolding times on the order of microseconds ($t_{90} = 4.0 \mu s$)

To interpret the previous results, a two-state model was proposed. This model can be described by the following differential equation:

$$\frac{dSAS}{dt} = k(SAS_{max} - SAS), \quad (1)$$

where SAS_{max} is the maximal surface exposed to the solvent and k is a kinetic constant. This equation shows that the surface begins to grow from a minimum value and maximum rate of growth to finally achieve, asymptotically, the maximum value. This last value is the equilibrium point for the system at 3 kbar. By solving this equation, the following expression is obtained:

$$SAS(t) = SAS_{max} - (SAS_{max} - SAS_{min})e^{-kt}, \quad (2)$$

where SAS_{min} is hence the minimal solvent accessible surface and is given by the initial condition.

Equation 2 was fitted to simulated results of apomyoglobin at 3 kbar (Fig. 2, continuous grey curve) using a program implemented in Fortran 95 language (see “[Computational Methods](#)”, code available).

We have performed a free fit of the two-state model, which implies that the behaviour of the simulated data is fit to first-order kinetics without the application of any kind of restriction on the values of SAS for the folded or unfolded states, or for the kinetic unfolding constant (SAS_{min} , SAS_{max} and k). Furthermore, these three values arise as a result of the best fit possible for the collected simulation data to the first-order model.

The fitted curve for apomyoglobin SAS at 3 kbar as well as its mean value at 1 bar can be observed in Fig. 2 (continuous grey and black curves, respectively). The values for the parameters of Eq. 2, as obtained from the free fitting, are shown in Table 1.

To interpret these results, we have calculated the maximum solvent accessible surface for apomyoglobin, without violating the restrictions imposed by topology. This process was performed using the Swiss PDB program. Thus, we have calculated the SAS for the “stretched” structure, obtaining $SAS_s = 201.16 \text{ nm}^2$ (where subscript s means “stretched”).

This last value of SAS_s can be interpreted as the theoretical maximum SAS (i.e. apomyoglobin may only further increase its SAS exposure, breaking covalent bonds). Thus, the actual SAS of the protein can never be higher than this value. For the protein to reach this value implies that all amino acids must remain entirely exposed to the solvent (no atom contacts between either neighbouring or distant residues). Such a situation results in a very restricted accessible conformational space for the protein, which is expected to be very unfavourable, due to the reduction in conformational entropy. On an interesting note, the fitted value for SAS_{max} results well within the mathematical limit of SAS_s . Nevertheless, other than offering an extrapolated value for unfolded SAS (SAS_{max}), it is not the intention

Table 1 Fitting results

Parameters	Values
$S_{min} \text{ (nm}^2\text{)}$	162.40
$S_{max} \text{ (nm}^2\text{)}$	80.10
$k \text{ (ns}^{-1}\text{)}$	$4.10 \cdot 10^{-4}$
$t_{90} \text{ (}\mu\text{s)}$	4.0

Parameter values resulting from the free-fitting procedure of (2) to the apomyoglobin SAS time evolution which was, in turn, calculated from the complete 182-ns-long molecular dynamics trajectories of this protein at 3 kbar.

of the present study to model the fully unfolded state for apomyoglobin. The two-state model fitted to our presented molecular dynamics simulation predicts an unfolding kinetics which reveals a maximum SAS ($SAS_{\max}$) which is, reasonably, roughly halfway between the SAS for the folded conformation ($SAS_{\min}$) and the mathematical maximum SAS value (SAS_s).

On the other hand, the kinetic constant allows us to solve (2) to obtain the time in which 90% of the SAS is achieved. This value (t_{90}) represents the time in which the unfolding is practically accomplished. It is interesting to note that the value for t_{90} results in the order of microseconds. Two very important conclusions may be derived from this observation. First, this value gives us the extrapolated time for which we should extend our simulation to observe completeness in the unfolding of apomyoglobin, which immediately puts such study beyond our present computational capabilities. Nevertheless, it may also be readily stated that the simulated data collected for what is only a small fraction of the total process suffices to extrapolate measurable macroscopic parameters, such as the unfolding time, with a remarkable accordance with the experimental values [8].

The two-state model allows us to make a prediction for the unfolding of apomyoglobin in solution at 3 kbar. Figure 2, inset, shows a microsecond time scale of an extrapolation of the SAS at 3 kbar and 1 bar for the protein under study. In this figure we can observe the extrapolation of the SAS behaviour at both pressures. Experiments show values for kinetic constants of the folding–unfolding equilibrium process (extrapolated for conditions of zero denaturant concentration) on the sub-millisecond time scale [8]. The model considered here bridges the gap between simulation and experiment, showing that both perspectives are not only in good agreement, but are also highly complementary.

4 Conclusions

In this paper we reported a prediction for the kinetic of unfolding of apomyoglobin under pressure. This prediction is in good agreement with previous experimental reports showing a notably comparable kinetic constant of unfolding for sperm whale apomyoglobin [8]. This work shows a combination of molecular dynamics simulations and mathematical modelling, attempting to open a new perspective on the usefulness of these computational methods.

References

1. Damjanović, A., García-Moreno, B., Lattman, E.E., García, A.E.: Molecular dynamics study of water penetration in staphylococcal nuclease. *Proteins* **60**, 433–449 (2005)
2. Mohan, G., Kopelevich, D.I.: A multiscale model for kinetics of formation and disintegration of spherical micelles. *J. Chem. Phys.* **128**, 044905 (2008)
3. Sriraman, S., Kevrekidis, I.G., Hummer, G.: Coarse nonlinear dynamics and metastability of filling-emptying transitions: water in carbon nanotubes. *Phys. Rev. Lett.* **95**, 130603 (2005)
4. Seibert, M.M., Patriksson, A., Hess, B., van der Spoel, D.: Reproducible polypeptide folding and structure prediction using molecular dynamics simulations. *J. Mol. Biol.* **354**, 173–83 (2005)
5. van der Spoel, D., Seibert, M.M.: Protein folding kinetics and thermodynamics from atomistic simulations. *Phys. Rev. Lett.* **96**, 238102 (2006)
6. Prakash, S., Matouschek, A.: Protein unfolding in the cell. *Trends Biochem. Sci.* **29**, 593–600 (2004)
7. Scheraga, H.A., Khalili, M., Liwo, A.: Protein-folding dynamics: overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **58**, 57–83 (2007)
8. Jamin, M., Yeh, S., Rousseau, D.L., Baldwin, R.L.: Submillisecond unfolding kinetics of apomyoglobin and its pH 4 intermediate. *J. Mol. Biol.* **292**, 731–740 (1999)

9. McCarthy, A.N., Grigera, J.R.: Pressure denaturation of apomyoglobin: a molecular dynamics simulation study. *Biochim. Biophys. Acta.* **1764**, 506–515 (2006)
10. Eliezer, D., Yao, J., Dyson, H.J., Wright, P.E.: Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* **5**, 148–155 (1998)
11. Kitahara, R., Yamada, H., Akasaka, K., Wright, P.E.: High pressure NMR reveals that apomyoglobin is an equilibrium mixture from the native to the unfolded. *J. Mol. Biol.* **320**, 311–319 (2002)
12. Creamer, T.P., Srinivasan, R., Rose, G.D.: Modeling unfolded states of proteins and peptides. II. Backbone solvent accessibility. *Biochemistry* **36**, 2832–2835 (1997)
13. Geierhaas, C.D., Nickson, A.A., Lindorff-Larsen, K., Clarke, J., Vendruscolo, M.: BPPred: a Web-based computational tool for predicting biophysical parameters of proteins. *Protein Sci.* **16**, 125–134 (2007)
14. Myers, J.K., Pace, C.N., Scholtz, J.M.: Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148 (1995)
15. Schellman, J.A.: Solvent denaturation. *Biopolymers* **17**, 1305–1322 (1978)
16. Berendsen, H.J.C., van der Spoel, D., van Drunen, R.: GROMACS: a message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995)
17. Lindahl, E., Hess, B., van der Spoel, D.: Gromacs 3.0: a package for molecular simulation and trajectory analysis. *J. Molecular Modeling* **7**, 306–317 (2001)
18. van Gunsteren, W.F., Berendsen, H.J.C.: Computer simulation of molecular dynamics: methodology, applications, and perspectives in chemistry. *Angew. Chem. Int. Ed. Engl.* **29**, 992–1023 (1990)
19. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L.: Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983)
20. Liu, H., Müller-Plathe, F., van Gunsteren, W.F.: A force field for liquid dimethyl sulfoxide and liquid properties of liquid dimethyl sulfoxide calculated using molecular dynamics simulation. *J. Am. Chem. Soc.* **117**, 4363–4366 (1995)
21. Mark, A.E., van Helden, S.P., Smith, P.E., Janssen, L.H.M., van Gunsteren, W.F.: Convergence properties of free energy calculations: cyclodextrin complexes as a case study. *J. Am. Chem. Soc.* **116**, 6293–6302 (1994)
22. van Buuren, A.R., Berendsen, H.J.C.: Molecular dynamics simulation of the stability of a 22 residue alpha-helix in water and 30% trifluoroethanol. *Biopolymers* **33**, 1159–1166 (1993)
23. van Buuren, A.R., Marrink, S.J., Berendsen, H.J.C.: A molecular dynamics study of the decane/water interface. *J. Phys. Chem.* **97**, 9206–9212 (1993)
24. van Gunsteren, W.F., Berendsen, H.J.C.: Gromos-87 manual. Biomos BV Nijenborgh, vol. 4, p. 9747. AG Groningen, The Netherlands (1987)
25. Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M.: LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997)
26. Miyamoto, S., Kollman, P.A.: SETTLE An analytical version of the SHAKE and RATTLE algorithms for rigid water models. *J. Comput. Chem.* **13**, 952–962 (1992)
27. Berendsen, H.J.C., Grigera, J.R., Straatsma, T.P.: The missing term in effective pair potentials. *J. Chem. Phys.* **91**, 6269–6271 (1987)
28. Berendsen, H.J.C., Postma, J.P.M., DiNola, A., Haak, J.R.: Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984)
29. van der Spoel, D., Lindahl, E., Hess, B., van Buuren, A.R., Apol, E., Meulenhoff, P.J., Tieleman, D.P., Sijbers, A.L.T.M., Feenstra, K.A., van Drunen, R., Berendsen, H.J.C.: Gromacs user manual version 3.2, www.gromacs.org (2004)