PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

The maturation of HIV-1 protease precursor studied by discrete molecular dynamics

Sachie Kimura,[1] Martina Caldarini,[1] Ricardo A. Broglia,[1,2] Nikolay V. Dokholyan,[3] and Guido Tiana[1]*

[1] Department of Physics and INFN, Università degli Studi di Milano, Milano 20133, Italy

[2] The Niels Bohr Institute, University of Copenhagen, Copenhagen 2100, Denmark

[3] Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, North Carolina 27599

ABSTRACT

The equilibrium properties of a HIV-1-protease precursor are studied by means of an efficient molecular dynamics scheme, which allows for the simulation of the folding of the protein monomers and their dimerization into an active form and compare them with those of the mature protein. The results of the model provide, with atomic detail, an overall account of several experimental findings, including the NMR conformation of the mature dimer, the calorimetric properties of the system, the effects of the precursor tail on the dimerization constant, the secondary chemical shifts of the monomer, and the paramagnetic relaxation enhancement data associated with the conformations of the precursor. It is found that although the mature protein can dimerize in a unique, single way, the precursor populates several dimeric conformations in which monomers are always native-like, but their binding can be non-native.

Proteins 2014; 82:633–639.

Key words: MD simulations; equilibrium properties; computational model; protein dimerization; non-native interactions.

INTRODUCTION

A key step in the replication of the HIV-1 virus is the maturation of its aspartyl protease (PR) into an active dimeric form. The protease is expressed by the virus as part of a longer polyprotein, where it can fold into a monomeric, inactive conformation.[1–3] The occasional formation of transient active dimers can lead to the cleavage of the protease itself, and eventually of the other viral proteins into their mature form. The blocking of such a transient binding has been suggested as a strategy to inhibit viral replication.[4] Central to the quest of making operative such expectation is to have a detailed description of the maturation of the protease with atomic detail.

The minimal system that mimics the properties of the HIV-1 protease precursor was obtained building a construct (^{SFNF}PR) that includes the protease and four residues of its N-terminal flanking sequence in the precursor.[5] Biochemical studies[1] of the inactive D25N variant of ^{SFNF}PR indicate that the monomeric structure is identical to that found in the mature dimer, except for the active site (residues 25–27) and residues 3–6 at the N terminus. On the other hand, the dimerization constant increases from the value of ∼10 nM in the case of the mature dimer to ∼500 μM in the case of $^{SFNF}PR_{D25N}$, making it largely monomeric under conditions that mimic cellular environment.

The goal of this study is to understand the role of the precursor-flanking sequence at atomic detail, making use of a computational model of the system and advanced computational algorithms. A thorough sampling of the conformational space of the protease is particularly trying because of the size of the system, which involves two chains of 103 residues each. So far, similar task was carried out successfully with structure-based Go models,[2–8] which cannot be used in the present context because the precursor-flanking sequence is unstructured. Portable models, either with reduced degrees of freedom[1,9] or all-atom in explicit solvent,[10] were used to study conformational changes involving only a small part of the whole

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Guido Tiana, Department of Physics, University of Milano, via Celoria 16, Milano 20133, Italy. E-mail: tiana@mi.infn.it

Received 19 June 2013; Revised 6 September 2013; Accepted 26 September 2013
Published online 9 October 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24440

conformational space. In fact, the present tools do not allow to use an explicit solvent description of the system and standard molecular dynamics (MD) to study the whole conformational space of such a large system as the HIV-1-PR precursor. To date, the largest system studied by MD in explicit solvent to obtain equilibrium properties has been a monomeric 36-residue protein.[2,3,11]

To overcome this problem, we made use of an implicit-solvent model of the protein, Medusa force field, and of discrete MD simulations. Medusa force field has proven efficient in folding small proteins to their crystallographic native state *ab initio*,[4,12] to reproduce equilibrium and kinetic properties of proteins up to 100 residues,[5,13] and to predict the thermodynamic effect of mutations on a large set of proteins.[14] Discrete MD simulations[15] is an event-driven algorithm that allows to use time steps larger than traditional MD, its speed scaling as $N \log N$, making it particularly attractive to study large molecular systems. The parallel version of discrete MD coupled with a replica-exchange algorithm allows for an efficient study of the equilibrium properties of large molecular systems.[16]

The system under study displays an additional degree of complexity with respect to that of proteins studied in the past with the same model, namely, its dimeric character. In fact, in the present case, the model has to reproduce two related but distinct phenomena: the folding of the two polypeptidic chains and the recognition of the two structured (or partially structured) molecules. The reason for which one attempts at describing such a forbidding system is to be found in the need of devising new strategies to fight AIDS, let alone of designing inhibitors which do not lead to escape mutants. Furthermore, the case of HIV-1 PR is a challenging test to evaluate the performance of the model.

As a first step in the comprehension of the maturation of the HIV-1 PR, we will validate the ability of the model to reproduce available experimental data, such as the structure of the mature dimer, the thermodynamics of the system, and the NMR data describing the transient dimeric precursor. In particular, paramagnetic relaxation enhancement (PRE) data obtained for ^{SFNF}PR[17] provide a sensitive atomic-scale test for the model. Moreover, we shall look for equilibrium structural features that can hardly be obtained with available experimental techniques.

METHODS

Discrete MD algorithm is an event-driven simulation method using a discrete potential energy, relying on the calculation of a sequence of atomic collisions.[11,12,16] The all-atom protein model interacts through the implicit-solvent force-field Medusa, which takes into account van der Waals interaction, hydrogen bonds, electrostatics, dihedral and angular interactions, and, in an effective fashion, the interactions mediated by the solvent. We make use of the extended version of the force field, improved from the point of view of electrostatic interactions and sequence-dependent local backbone interactions.[12,13] The force field is defined except for a multiplicative constant, and consequently energy units are arbitrary.

The replica-exchange sampling scheme is used to significantly enhance conformational sampling by simulating several replicas of the system at different temperatures. For each system, we first carried out preliminary simulations using eight replicas, from $T = 0.5$ to $T = 0.65$ (having here set the Boltzmann constant to 1, thus setting temperatures in energy units), and then 20 replicas, from $T = 0.2$ to $T = 0.65$ (to 0.70 in the case of the monomer). Preliminary simulations started from the native conformation, thermalized at each temperature for 10^5 steps, whereas in the 20-replica simulation, each replica started from the last conformation of the preliminary simulation carried out at the closer temperature. Attempts to exchange neighboring replicas are carried out every 1000 steps. Eventually, we joined the two sets of simulations of each system, extracting the average energies by means of a weighted histogram algorithm.[13,18]

The side length of the simulation box with periodic boundary condition is chosen to be 8 and 16 nm in the case of the monomer and of the mature dimer/^{SFNF}PR, respectively. Analysis of the trajectories has been carried out with the help of the Gromacs tools.[19] The cluster analysis was performed with the algorithm described in Ref. 20 making use of a cutoff of 0.7 nm.

RESULTS

One of the main advantages of describing the dynamics of such a large system as the ^{SFNF}PR dimer by discrete MD is the possibility of simulating its equilibrium properties. After ~3 months of simulation on 20 processors, we obtained a good approximation of its equilibrium state (cf. Supporting Information Figs. S1 and S2).

As a first control, we have to check that the mature HIV-1-PR dimer folds into the experimental native conformation. For this purpose, the minimum-energy conformations found in the simulation were compared with the minimized NMR structure,[21] calculating the associated RMSD on all atoms of the protein. The comparison is displayed in Figure 1, and the resulting RMSD is 0.46 nm, which decreases to 0.41 nm if the termini and the flap regions are not taken into account. These values are comparable with the results of popular tools used for *ab initio* protein structure prediction.[22] No conformations with comparable energies but markedly different from the experimental native conformation were found (see Supporting Information Fig. S3).

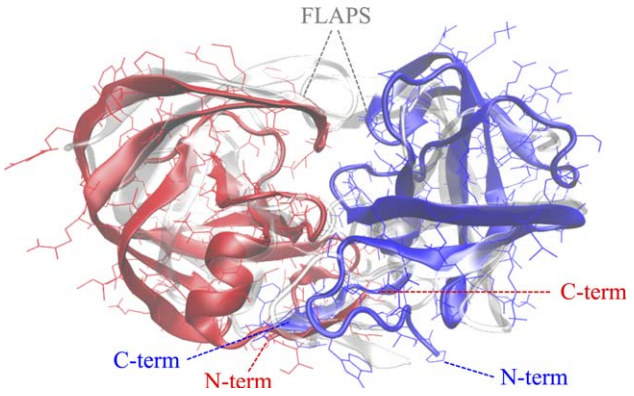

Figure 1

The minimum-energy conformation of the mature dimer, the corresponding monomers being colored in red and blue, superimposed to the average NMR conformation, displayed as a transparent cartoon. The RMSD between simulated and NMR conformations is 0.46 nm, the main differences being found at the termini and at the flap region. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

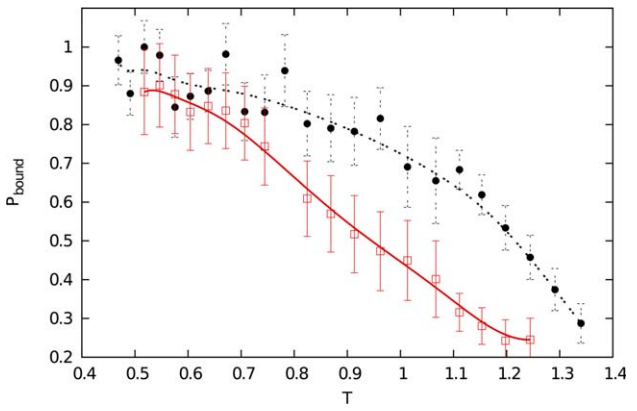

Figure 3

The probability that the two monomers in the dimer (solid black circles) and in the precursor (open red squares) are bound, defined as the fractional number of contacts between CA atoms of the two monomers (the error bars indicate the standard deviation of such fractional number). A contact between CA atoms is considered formed if they are closer than 0.5 nm. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The replica-exchange simulations described above can also provide all equilibrium properties of the protease. Figure 2 displays the heat capacity C_v of the monomer, of the dimer, and of the precursor protein, calculated with a weighted histogram method.[18] All these systems display a single major peak corresponding to the unfolding of the monomer (cf. Supporting Information Fig. S4), dimerization increasing the area under the single peak displayed by the C_v, and thus the two-state character of the unfolding transition of the monomer. Although the monomeric protein is stable at room temperature, its thermodynamics seems quite far from being two-state character. These results are qualitatively in agreement with the results of differential scanning calorimetry described in Ref. 23, although for different variants of the protease and under different solution conditions. In Figure 2 and in the following figures, the temperature scale is set in units of room temperature, determined as that at which the unfolding free energy of the monomer is equal to its experimental value (cf. Supporting Information Fig. S5 and Ref. 24). It should be noted that, as usually happens in simplified models, the scaling between model and real temperature is expected to be only approximately linear when moving away from room temperature, and consequently, the absolute value of the unfolding temperature of the protease may be unphysical.[25]

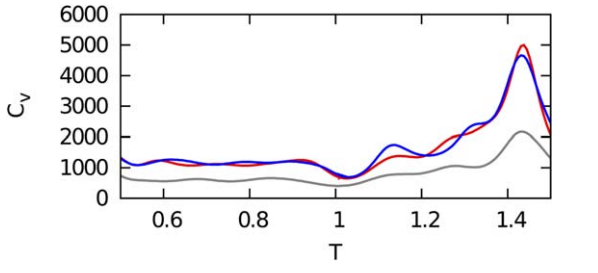

Figure 2

Heat capacity of monomer (gray curve), mature dimer (red curve), and ^{SFNF}PR (blue curve) as a function of temperature, in units of room temperature (see text and Supporting Information Fig. S5). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The simulations were carried out in a cubic box of side 16 nm, resulting in a protein concentration [P] = 400 μ*M*. The binding probability p_{bound} between the two monomers of the dimer and of the precursor are displayed in Figure 3 as a function of temperature. The dimerization constant at room temperature results to be 93 μ*M* for the dimer and 488 μ*M* for the precursor.

$$k_D = [P]\frac{1-p_{\text{bound}}}{p_{\text{bound}}}.$$

Consequently, the precursor dimer results to be considerably less stable than the mature dimer. This is in agreement with data available in the literature, which describe an increase of k_D from 5 n*M* of the wild-type mature dimer to 680 n*M* of the TFP-p6^{pol} precursor at pH 5,[26] from 0.5 μ*M* of the D25N mature dimer to >500 μ*M* of ^{SFNF}PR$_{D25N}$ precursor at pH 5,[1] and from 10 n*M* of the D25N mature dimer to 3–6 m*M* of the ^{SFNF}PR$_{D25N}$ precursor at pH 5.8.[17]

The average structures at biological temperature of the monomer alone, of the monomer when bound into the dimer, of the precursor in monomeric state, and of the

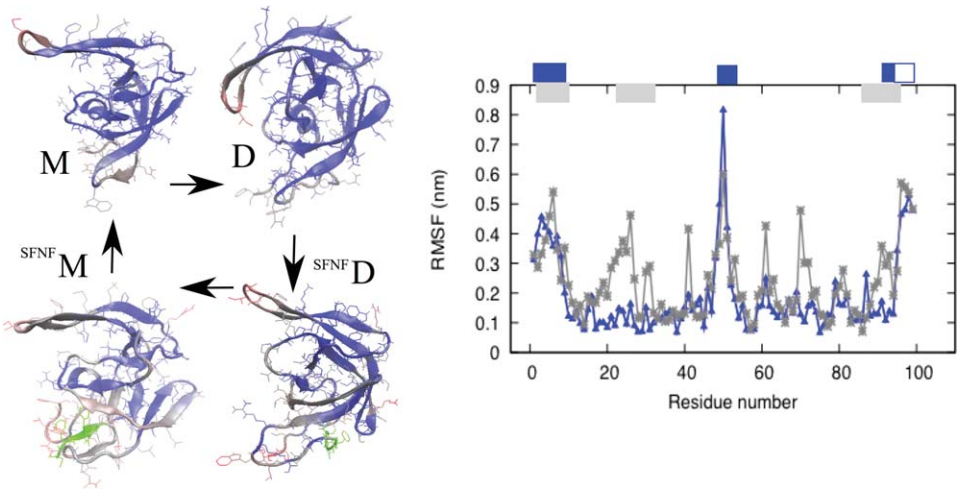

Figure 4

On the left panel, the conformational difference between the average structure of the monomer by itself (M), in the dimer (D), within the monomeric precursor (^{SFNF}M), and within the dimeric precursor (^{SFNF}D) is shown. The colors indicate the deviation of the position of the monomer corresponding to a structure to which it is connected by an arrow (blue: small deviations; red: large deviations). The green segment is the N-terminal tail indicating the precursor. On the right panel, the average deviation of the position of residues in conformations M and D (blue line) and in conformations ^{SFNF}M and M (gray line) is shown. The gray and blue bars on the top of the plot indicate the difference in chemical shifts for the same molecules observed in Refs. 5 and 4, respectively. The white square indicates that the monomeric protease was deprived of residues 96–99.

monomer of the precursor in dimeric state are displayed in the left panel of Figure 4. Their RMSDs to the minimized NMR conformation of the monomer are 0.50, 0.38, 0.41, and 0.6 nm, respectively (0.31, 0.32, 0.30, and 0.48 nm not counting the flaps and the termini). This agrees with the NMR data suggesting that both the monomeric and the precursor forms of the protease display an overall native structure.[1,5]

The structural differences between monomer and dimer and between monomeric precursor and mature monomer are displayed in the right panel of Figure 4, where they are compared with the secondary chemical shifts of CA and N atoms reported in Refs. 1 and 5. The model predicts correctly large deviations between the monomer and the dimer in the termini and in the flap region and between the monomeric precursor and the mature monomer in the termini and in the regions 22–32. The model also predicts deviations of single residues (41, 61, and 70) and of the flap residues 49–51 not observed experimentally. The comparison among the four forms of the protease (see also Supporting Information Fig. S6) suggests that all of them are different in the termini (residues 1–15 and 95–99) and in the flap region (residues 44–56), whereas the monomeric precursor displays larger differences involving also the regions 18–30 participating to the hydrophobic core.

The width of the thermal fluctuations of the residues of the protease is displayed in Figure 5 (see also Supporting Information Fig. S7). As concluded in Ref. 5 from the absence of NOE signals in the termini and the flap of the Δ96–99 monomeric variant, these regions are highly mobile. The simulations suggest a large amount of conformational freedom in these regions also in the dimer and in the dimeric precursor. Concerning the

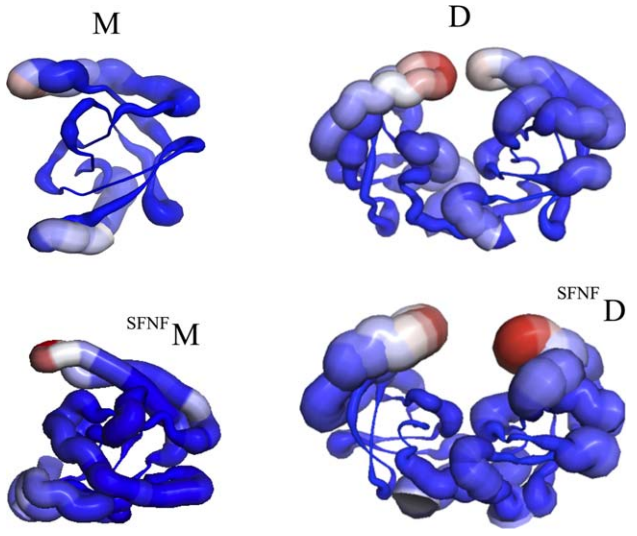

Figure 5

A cartoon representation of the thermal fluctuations of monomer (M), dimer (D), and monomeric precursor (^{SFNF}M) and dimeric precursor (^{SFNF}D) at room temperature. The thicker the cartoon (emphasized in red), the larger are the fluctuations around the average structure. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

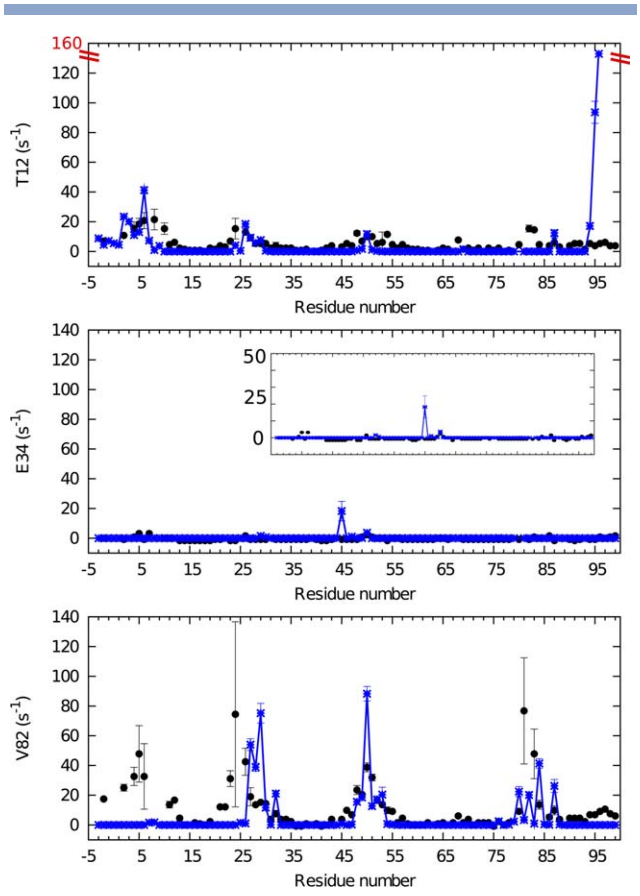

Figure 6

The experimental values obtained by PRE in Ref. 17 between T12, E34, and V82 of a subunit and the core residues (10–94) of the other subunit in ^{SFNF}PR (solid black circles) when compared with the value of $<r^{-6}>$ calculated from the corresponding simulation at room temperature, normalized accordingly (blue symbols). The error bars represent the standard deviation of the signal. Inset, a zoom of the E34 data. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

flaps, the fluctuations observed in the simulations of the mature dimer are in agreement with the high-temperature factors observed in the crystallographic structure of the protease.[27]

An interesting experiment to characterize the precursor is reported in Ref. [17], where PRE data are reported concerning long-distance information between residues T12, E34, and V82 of a monomer of the precursor and the core residues (10–95) of the other monomer. These data are in good agreement with the average value of $1/r^6$, normalized accordingly, displayed in Figure 6. As discussed in Ref. [17], the data concerning residue E34 represent a negative control, which we are also able to satisfy. The agreement of the data concerning residues T12 and V82 is definitely not trivial to obtain, as the same calculation carried out in the simulation of the mature dimer displays a rather different pattern (see Supporting Information Fig. S8).

From the analysis of the simulated system, one can get insight into the mechanism associated with the dimerization of the precursor. Within this context, we focus our attention on the terminal segments, which embody most of the dimerization interface. Figure 7 displays the distribution of RMSD of segments 1–5, which appears bimodal, whereas that calculated in the mature protease displays a single peak. This observation suggests that in the precursor, the N terminus acts as a switch, sharply discriminating the native from non-native conformations.

The conformations of the precursor and of the mature dimer at room temperature have been further studied by cluster analysis. Although the mature dimer displays few, highly populated clusters, essentially accounting for thermal fluctuations around the native conformation, the precursor displays a broad distribution composed of several clusters (see Fig. 7; cf. also Supporting Information Fig. S9). The most populated of them is native-like, with the precursor tails not interfering with the native binding between the monomers (see upper-right structure in Fig. 7). The other clusters still display native-like monomers, but with various mutual arrangement and displaying the precursor tail interacting with the termini of the protein. The lower-right panel of Figure 7 shows the second cluster as an example.

To get a further insight into the set of conformations that the precursor can populate, we show in the upper-left panels of Figure 7 the distribution of RMSD of the N-terminal segments 1–5 to its native conformation for the precursor (^{SFNF}D) and, as reference, for the mature dimer (D). Unlike the mature dimer, the N-terminus of the precursor displays a bimodal distribution, suggesting two possible states. The probability of the more non-native-like state increases with the temperature, suggesting that it is entropically stabilized, and thus is more disordered. In summary, the precursor can bind the two native monomers in a complicated set of different mutual positions, according to the different ways the precursor tail interacts with the terminals of the protein.

DISCUSSION

The maturation of the HIV-1 protease is particularly rich of features, as it involves time and length scales that span many orders of magnitude. Even restricting to a single step of the maturation process, that is, the transient dimerization of the precursor protease, one is confronted with a variety of biophysical phenomena: the folding of the monomeric protein, the molecular recognition and binding, and the dynamics of the unstructured terminal. Experimental techniques, typically involving NMR, can be used at profit to obtain many valuable information on such a process, but are not able to provide a full atomistic picture of the system.

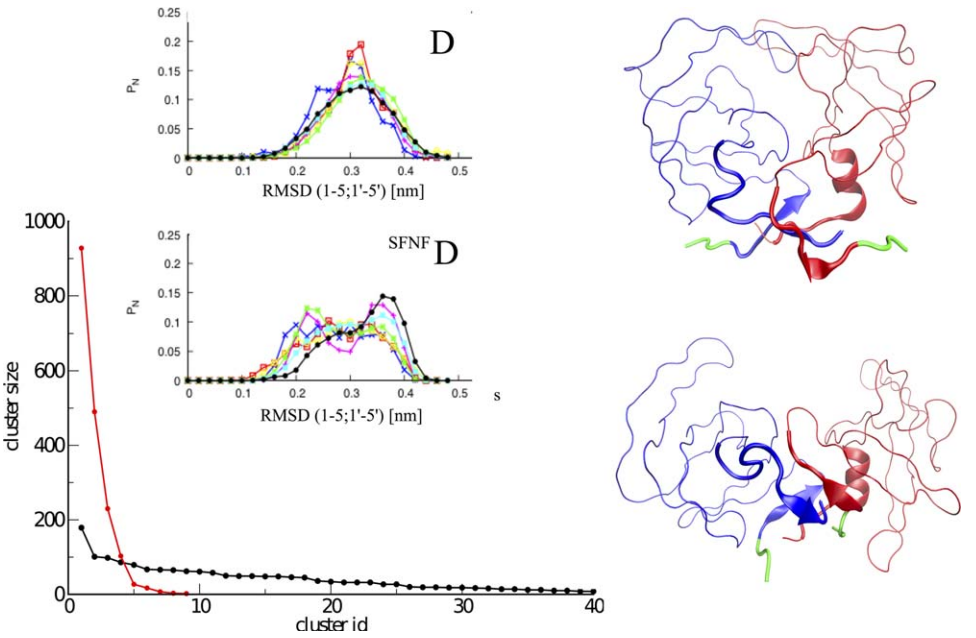

Figure 7

In the lower, left part of the figure, the distribution of cluster sizes obtained for the precursor (in black) and for the mature dimer (in red) is displayed. Top: The distribution of RMSD of the N-terminal segments 1–5 of the dimer and of the precursor at increasing temperature from $0.8T_{room}$ (blue curve) to $1.4T_{room}$ (in red) is shown. On the right, the average structures of the first two clusters found for the precursor are shown. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

When studying the folding of small monomeric proteins or the binding of small molecules, MD simulations carried out with standard explicit-solvent force fields can be of help, complementing the experimental data and giving an atomistic description of the system. However, in the case of such a complicated phenomenon as the dimerization of HIV-1-PR precursor, the possibility of obtaining an equilibrium sampling of two separate chains of 103 amino acids each is computationally not feasible. In fact, the largest systems whose equilibrium properties have been calculated correspond to biomolecules involving monomeric chains of less than one-fourth the size of the present system[11] or the binding of molecules of few tens of atoms to a protein.[28]

A sensible trade-off between computational economy and a realistic description of the system comes from implicit-solvent protein models. Besides involving a much smaller number of atoms than the corresponding explicit-solvent representation of the system, these models offer a further advantage, allowing the use of computational algorithms[15] that are more efficient than plain MD.

In order to be reliable, models must be validated. In the case of explicit-solvent models, the validation process took more than two decades. In the case of the Medusa force field, several studies[12,13] show its ability to identify correctly the native state of proteins up to 100 residues and to account for their thermodynamic properties. In this work, we increase our expectations and challenge the model to reproduce a more complicated phenomenon, involving the folding and dimerization of a large protein, perturbed by the interaction with the precursor tail.

The model can reproduce to a significant degree of approximation the NMR conformation of the mature dimer, the two-state thermodynamics of the dimer and the broadening of the calorimetric peak in the monomer, and the effect of the precursor tail on the dimerization constant. To a more atomistic level, the model can correctly predict the structural differences between monomer and dimer and between monomer and precursor, as reported by secondary chemical shifts in NMR experiments. Furthermore, the PRE data describing the interaction of the monomers in the precursor are correctly reproduced.

The cluster analysis of the conformations visited by the precursor at room temperature indicates that multiple quaternary structures are possible, according to the different ways the precursor tail interacts with the termini of the protein, whereas the tertiary structure of the monomers remain overall native-like. This finding is qualitatively in agreement with the analysis carried out in Ref. 17, where the authors back-calculate PRE data from rigid-body sampling of the relative positions of the native monomers. They concluded that a single conformation is not able to reproduce the experimental PRE, an average of four conformations being the minimum to obtain sensible results, whereas a larger number of

conformations would lead to an over-fitting of the data. Our simulations indicate that indeed the quaternary structure of the precursor involves multiple states, but that the number of these states (although depending on the precise definition) is larger than 4. As many of these states correspond to an arrangement of the monomers quite different from the native one, we expect them to be enzymatically inactive, and consequently, we do not expect a simple relationship between the dimerization constant of the protein and its Michaelis constant.

CONCLUSIONS

Computational modeling of biological systems can be a very useful tool to complement and interpret experimental data, provided that a suitable model is chosen. Not all models are suitable to describe all phenomena. In the case of studying the equilibrium properties associated with the maturation of the HIV-1-PR precursor, explicit-solvent models, although very reliable for many purposes, are computationally too demanding to obtain equilibrium properties. We have shown that a simpler model can reproduce a large amount of experimental data and help to understand the equilibrium properties of the protease precursor at atomistic level.

ACKNOWLEDGMENTS

N.V.D. acknowledges the financial support of the NIH (Grant R01AI102732).

REFERENCES

1. Ishima R, Torchia DA, Louis JM. Mutational and structural studies aimed at characterizing the monomer of HIV-1 protease and its precursor. J Biol Chem 2007;282:17190–17199.
2. Louis JM, Nashed NT, Parris KD, Kimmel AR, Jerina DM. Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. Proc Natl Acad Sci USA 1994;91:7970–7974.
3. Louis JM, Wondrak EM, Kimmel AR, Wingfield PT, Nashed NT. Proteolytic processing of HIV-1 protease precursor, kinetics and mechanism. J Biol Chem 1999;274:23437–23442.
4. Louis JM, Aniana A, Weber IT, Sayer JM. Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors. Proc Natl Acad Sci USA 2011; 108:9072–9077.
5. Ishima R, Torchia DA, Lynch SM, Gronenborn AM, Louis JM. Solution structure of the mature HIV-1 protease monomer: insight into the tertiary fold and stability of a precursor. J Biol Chem 2003;278: 43311–43319.
6. Cecconi F, Micheletti C, Carloni P, Maritan A. Molecular dynamics studies on HIV-1 protease drug resistance and folding pathways. Proteins 2001;43:365–372.
7. Levy Y, Caflisch A, Onuchic JN, Wolynes PG. The folding and dimerization of HIV-1 protease: evidence for a stable monomer from simulations. J Mol Biol 2004;340:67–79.
8. Bonomi M, Barducci A, Gervasio FL, Parrinello M. Multiple routes and milestones in the folding of HIV-1 protease monomer. PLoS One 2010;5:e13208.
9. Tozzini V, Trylska J, Chang C-E, McCammon JA. Flap opening dynamics in HIV-1 protease explored with a coarse-grained model. J Struct Biol 2007;157:606–615.
10. Hornak V, Okur A, Rizzo RC, Simmerling C. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proc Natl Acad Sci USA 2006;103:915–920.
11. Piana S, Lindorff-Larsen K, Shaw DE. Protein folding kinetics and thermodynamics from atomistic simulation. Proc Natl Acad Sci USA 2012;109:17845–17850.
12. Ding F, Tsao D, Nie H, Dokholyan NV. Ab initio folding of proteins with all-atom discrete molecular dynamics. Structure 2008;16: 1010–1018.
13. Shirvanyants D, Ding F, Tsao D, Ramachandran S, Dokholyan NV. Discrete molecular dynamics: an efficient and versatile simulation method for fine protein characterization. J Phys Chem B 2012;116: 8372–8382.
14. Yin S, Ding F, Dokholyan NV. Eris: an automated estimator of protein stability. Nat Methods 2007;4:466–467.
15. Dokholyan NV, Buldryev SV, Stanley HE, Shakhnovich EI. Discrete molecular dynamics studies of the folding of a protein-like model. Fold Des 1998;3:577–587.
16. Proctor EA, Ding F, Dokholyan NV. Discrete molecular dynamics. WIREs Comput Mol Sci 2011;1:80–92.
17. Tang C, Louis JM, Aniana A, Suh J-Y, Clore GM. Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. Nature 2008;455:693–696.
18. Ferrenberg A, Swendsen R. Optimized Monte Carlo data analysis. Phys Rev Lett 1989;63:1195–1198.
19. Noel JK, Whitford PC, Sanbonmatsu KY, Onuchic JN. SMOG@ctbp: simplified deployment of structure-based models in GROMACS. Nucleic Acids Res 2010;38:W657–W661.
20. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: when simulation meets experiment. Angew Chem Int Ed Engl 1999;38:236–240.
21. Yamazaki T, Hinck AP, Wang Y-X, Nicholson LK, Torchia DA, Wingfield P, Stahl SJ, Kaufman JD, Chang CH, Domaille PJ, Lam PY. Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy. Protein Sci 1996;5:495–506.
22. Zhang Y. Progress and challenges in protein structure prediction. Curr Opin Struct Biol 2008;18:342–348.
23. Sayer JM, Liu F, Ishima R, Weber IT, Louis JM. Effect of the active site D25N mutation on the structure, stability, and ligand binding of the mature HIV-1 protease. J Biol Chem 2008;283:13459–13470.
24. Noel AF, Bilsel O, Kundu A, Wu Y, Zitzewitz JA, Matthews CR. The folding free-energy surface of HIV-1 protease: insights into the thermodynamic basis for resistance to inhibitors. J Mol Biol 2009;387: 1002–1016.
25. Liwo A, Khalili M, Scheraga HA. Ab initio simulations of protein-folding pathways by molecular dynamics with the united-residue model of polypeptide chains. Proc Natl Acad Sci USA 2005;102: 2362–2367.
26. Louis JM, Clore GM, Gronenborn AM. Autoprocessing of HIV-1 protease is tightly coupled to protein folding. Nat Struct Biol 1999; 6:868–875.
27. Spinelli S, Liu QZ, Alzari PM, Hirel PH, Poljak RJ. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. Biochimie 1991;73:1391–1396.
28. Limongelli V, Bonomi M, Parrinello M. Funnel metadynamics as accurate binding free-energy method. Proc Natl Acad Sci USA 2013; 110:6358–6363.