Biophysical Chemistry 195 (2014) 32–42

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

The complex folding behavior of HIV-1-protease monomer revealed by optical-tweezer single-molecule experiments and molecular dynamics simulations

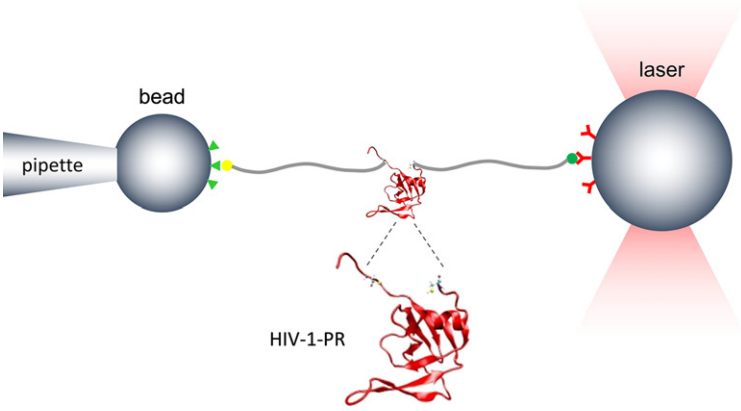
M. Caldarini [a,1], P. Sonar [b,c,1], I. Valpapuram [b,c], D. Tavella [a], C. Volonté [a], V. Pandini [d], M.A. Vanoni [d], A. Aliverti [d], R.A. Broglia [a,e], G. Tiana [a,*], C. Cecconi [b,c,*]

[a] Dipartimento di Fisica, Università degli Studi di Milano and INFN, Milano, Italy
[b] Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, Via Giuseppe Campi 213/A, 41125 Modena, Italy
[c] CNR Institute of Nanoscience S3, Via Giuseppe Campi 213/A, 41125 Modena, Italy
[d] Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy
[e] The Niels Bohr Institute, University of Copenhagen, Denmark

HIGHLIGHTS

- HIV-1-PR plays a key role in the life cycle of the human immunodeficiency virus.
- Complex (un)folding behavior was revealed by single-molecule and MD studies.
- HIV-1-PR unfolds either in a two-state manner or by populating an intermediate state.
- Folding of denaturated HIV-1-PR monomer is a multi-pathway process.
- Folding information can guide rational design of HIV-1-PR folding inhibitors.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form 4 August 2014
Accepted 4 August 2014
Available online 14 August 2014

Keywords:
HIV-1-protease
Single-molecule studies
Protein folding
Intermediate states
Molecular dynamics simulations

ABSTRACT

We have used optical tweezers and molecular dynamics simulations to investigate the unfolding and refolding process of a stable monomeric form of HIV-1-protease (PR). We have characterized the behavior under tension of the native state (*N*), and that of the ensemble of partially folded (*PF*) conformations the protein visits *en route* to *N*, which collectively act as a long-lived state controlling the slow kinetic phase of the folding process. Our results reveal a rich network of unfolding events, where the native state unfolds either in a two-state manner or by populating an intermediate state *I*, while the *PF* state unravels through a multitude of pathways, underscoring its structural heterogeneity. Refolding of mechanically denatured HIV-1-PR monomers is also a multiple-pathway process. Molecular dynamics simulations allowed us to gain insight into possible conformations the protein adopts along the unfolding pathways, and provide information regarding possible structural features of the PF state.

* Corresponding authors.
E-mail addresses: tiana@mi.infn.it (G. Tiana), ciro.cecconi@gmail.com (C. Cecconi).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bpc.2014.08.001
0301-4622/

1. Introduction

Human immunodeficiency virus type 1 protease (HIV-1-PR) is a dimeric aspartyl protease, made of two identical monomers of 99 residues each, which plays an essential role in the life cycle of HIV [1]. The virus expresses its protease embedded in a polyprotein, where it is largely monomeric [2] and the monomer adopts a fold very similar to that observed in the functional and mature dimer [3–5]. The structures of the wild-type and variant HIV-1-PR forms have been widely studied because of the importance of this enzyme as a target for antiviral therapy [6,7]. Likewise, the maturation process leading to the active dimer [8], comprising the folding of the monomer, its cleavage from the polyprotein and the assembly of the monomers into the dimer, has been intensively investigated as each one of these steps could be potentially blocked to inhibit HIV replication [9–11].

In spite of its small size, the folding process of the HIV-1-PR monomer seems quite complex as compared to that of proteins of similar length. Under biological conditions, it takes about 50 s to reach its native conformation, not involving any proline isomerization, and events occurring at least on two different time scales were identified in stopped-flow fluorescence experiments [12]. The native conformation itself appears more complicated than that of proteins of similar size. In fact, if we quantify the complexity of the native topology with the Plaxco's contact order *CO* [13], we find that compared to proteins of similar size, such as ACBP ($CO = 11$), protein G ($CO = 9$), CI2 ($CO = 10$) and SH3 ($CO = 10$), the monomer of the HIV-1-PR has a higher *CO* of 15.

Simulations have suggested the existence of multiple pathways in the unfolding and refolding processes of HIV-1-PR monomer [14]. However, to date no experimental evidence of this complex folding scenario has been provided. Several studies have examined the folding mechanism of HIV-1-PR using a variety of ensemble techniques, such as differential scanning calorimetry [15], NMR [16], urea denaturation [17] and sedimentation equilibrium studies [18]. Although these studies have given information on the overall thermodynamics and kinetics of the process, none has provided insight into the ensemble of folding routes that the individual molecules follow to reach their native state (*N*), as these experimental methods provide only averaged information. The advent of single-molecule techniques, such as optical tweezers, has made it possible to revisit protein folding with a new approach, allowing us to go beyond the ensemble average and to dissect folding mechanisms in unprecedented detail. Through mechanical manipulation, it is possible to follow in real time the trajectories of individual molecules and to describe the inherent heterogeneity of biological processes that are stochastic in nature, such as the folding of a polypeptide chain into a functional protein [19–24]. Moreover, these studies can be complemented by state-of-the-art molecular dynamics (MD) simulations, which, through a comparative analysis with the experimental data, can provide atomic details of the process under investigation [25,26].

In this study, we used optical tweezers and MD simulations to shed light on the HIV-1-PR folding mechanism. We manipulated a stable monomeric form of HIV-1-PR to show that it unfolds mainly along two trajectories, sometimes populating a mechanically stable intermediate state. Taking advantage of the slow folding kinetics of HIV-1-PR observed by fluorescence and circular-dichroism experiments [12], we were also able to manipulate partially folded (*PF*) conformations that the molecule adopts along its journey to the native state. Our data indicate that the *PF* state is long-lived, compact and structurally inhomogeneous, unfolding along a multitude of trajectories when mechanically stretched. State-of-the-art MD simulations revealed the sequence of events characterizing the unfolding and refolding trajectories, providing insight into possible structural features of the intermediate states populated by the protein and suggesting a role for non-native contacts in the unfolding process.

2. Results

In order to study the folding process of HIV-1-protease at the single-molecule level while avoiding the complication of monomer/dimer equilibria [27], the R87K substitution was introduced [28] in the Q7K/C67A/C95A protease variant [29]. Cysteine residues were also introduced at the N- and C-terminal ends of the protein by engineering the L5C and N98C substitutions. Urea denaturation of the protease variant, as monitored in solution by tryptophan fluorescence, was reversible and well described by a two-step process (Fig. 1), as previously reported for both the well-characterized dimeric, but inactive, HIV-1 protease variant and the monomeric (inactive) protein form described by Noel et al. [12]. However, the mutant is marginally stable ($\Delta G^\circ = 9.84 \pm 1.34$ kJ/mol) and prone to aggregation. Thus, suitable preparation protocols had to be set up to generate protein–DNA chimeras for use in single-molecule experiments (see Materials and methods for details).

Single monomers of HIV-1-PR were manipulated as depicted in Fig. 2. The molecules were stretched and relaxed by moving the pipette relative to the optical trap, while the applied force and molecular extension were measured as previously described [19,30]. As the protein unfolds and refolds under tension its extension suddenly changes, generating discontinuities (rips) in the stretching and relaxation traces that can be analyzed to gain insight into the kinetic and mechanical properties of the molecule [20]. We used this experimental approach to explore the unfolding and refolding processes of native and partially folded conformations of the monomeric HIV-1-PR.

2.1. The native state unfolds along two main pathways

To study the unfolding trajectories of the native state of monomeric HIV-1-PR, experiments were performed in which the force applied on the protein was first ramped (~5 pN/s) to 55 pN to induce the unfolding of the molecule; then, it was decreased (~−5 pN/s) to 2 pN and kept constant at such low value for 90 s, before the next stretching-relaxation cycle started. In the absence of force the HIV-1-PR monomer takes ~50 s to refold into its native state from a urea-induced denatured state [12]. In the present experiments, the HIV-1-PR is given approximately twice as much time (90 s) to fold into *N* because under tension, although very small (2 pN), folding is slowed down [32,33]. The force

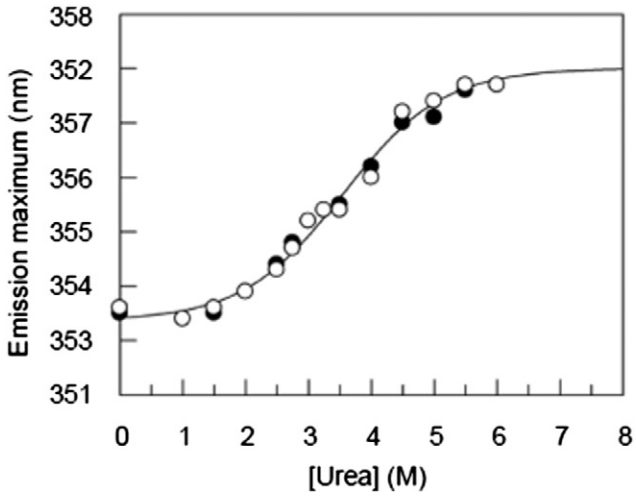


Fig. 1. Equilibrium folding properties of the L5C/R87K/N98C HIV-1-PR variant. Fluorescence emission of protease samples (2 μM in 20 mM sodium phosphate buffer, pH 6.0, 1 mM DTT) containing increasing (empty symbols) or decreasing (closed symbols) urea concentration was measured at 25 °C at equilibrium (excitation light, 290 nm). The wavelength of the maximum of the emission spectrum was plotted as a function of urea concentration. The curve shows the common fit to a two-state model (Eq. (1)) with a ΔG° for folding of −9.84 ± 1.34 kJ/mol and an *m* value of 2.85 ± 0.42 kJ/mol.

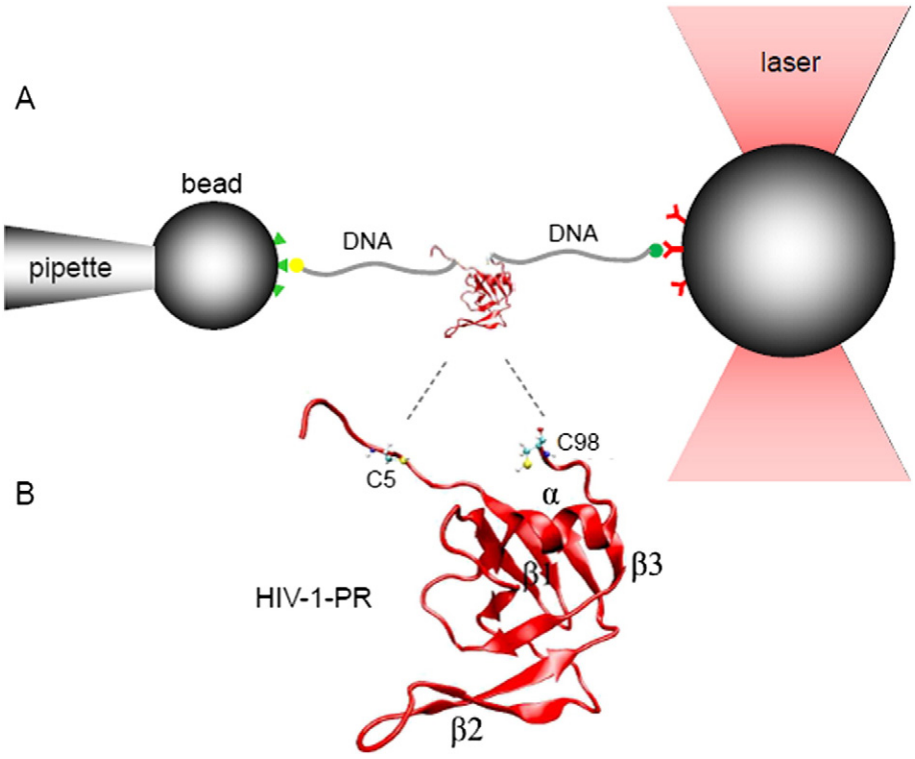


Fig. 2. Experimental setup and structure of HIV-1-PR. A) Individual molecules were manipulated by means of DNA handles (~500 bp DNA molecules) covalently attached to cysteine residues engineered at positions 5 and 98. Each DNA handle is tethered to a functionalized bead, as previously described [18,47]. B) NMR structure of the monomeric form of HIV-1-PR (Protein Data Bank code 1BVG).

applied on the molecule is never lowered to zero, in order to avoid direct contact between the polystyrene beads, which would generate unwanted interactions between the tethering surfaces. The unfolding events observed during the first stretching of the molecule, which corresponds to the unfolding of the native state, are indistinguishable from those observed during subsequent stretching-relaxation cycles, suggesting proper folding of HIV-1-PR during the 90 s wait period. It is worth mentioning that these experiments are technically very challenging for two main reasons. First, the interaction between digoxigenin and antibodies usually cannot stand forces larger than 25–30 pN. However, in spite of its poor thermodynamic stability, HIV-1-PR monomer is a mechanically resistant protein and force must be raised up to 55 pN to ensure complete unfolding. As a consequence, in each experiment many tethered molecules had to be tested to find one that could withstand such high tension for at least one cycle. Second, at 2 pN of force the two tethering surfaces are quite close to each other and in a 90 s period of time there is always a good chance that another molecule gets tethered between them, causing the end of the measurement.

Under these experimental conditions, the HIV-1-PR monomer was observed to unfold through two main mechanisms, as shown in Fig. 3. In 64% of the cases, it unfolded in an all-or-none fashion, transiting from *N* to the unfolded state (*U*) along a two-state unfolding pathway. Other times (31%) it unfolded in a three-state manner, populating an intermediate state (*I*). Only rarely (5% of the cases) it unraveled through different mechanisms (Figure S1). During multiple stretching/relaxation cycles, the same molecule is observed to take different unfolding pathways (Fig. 3).

When the HIV-1-PR unfolds in a two-state fashion, it generates a single discontinuity in the stretching trace at ~25 pN, which corresponds to the sudden increase in the extension of the molecule as it passes from its compact native state to its extended unfolded state (average rip width $d = 22 \pm 4$ nm), left trace in Fig. 3B. The increment in contour length (ΔL_c) upon unfolding of the molecule was estimated by fitting the force-extension curves to the worm-like chain model [36,37]. This analysis yielded a ΔL_c of 30 ± 3 nm, which compares well with the value of 31.5 nm expected for the complete unfolding of the molecule (see Materials and methods). Analysis of the force distribution of the two-state unfolding events (Fig. 3C) provided information on the unfolding rate constant (k^0_u) and distance ($x^{\ddagger}_u$) between the native state and the transition state along the reaction coordinate [19,35]. However, it is important to notice that rate constants estimated through optical tweezers experiments contain contributions from experimental parameters, such as DNA handle length and bead size; thus, they cannot be compared directly to rate constants of the free protein in solution [35]. On the contrary, the estimated $x^{\ddagger}_u$ value reflects an intrinsic property of the molecule. The distance $x^{\ddagger}_u$ is a measure of the mechanical compliance of the native state, i.e.: how much it can be deformed along the pulling axis without crossing the transition state. The larger is $x^{\ddagger}_u$, the more pliant a protein structure is. Analysis of the force distribution of Fig. 3C yielded a $x^{\ddagger}_u$ of 1 ± 0.1 nm, indicating that the compliance of HIV-1-PR's native state is relatively large compared to that of other all-β strand proteins [38–40]. This unusual pliability of the three-dimensional structure, if maintained in the dimeric species, might play a role during the catalytic activity of the enzyme, allowing the protease to adopt the proper conformations to interact with its substrates.

When the HIV-1-PR unfolds in a three-state manner (Fig. 3B, right trace), it first populates an intermediate state *I* at ~25 pN (*N*–*I* transition; average rip width $d = 17 \pm 4$ nm), and then it transits into its unfolded state at higher forces, (*I*–*U* transition; average rip width $d = 7 \pm 3$ nm). The partial unfolding of the protein into *I* results in a change in contour length ΔL_c of 22 ± 4 nm, while the subsequent unfolding of the intermediate state results in a ΔL_c of 9 ± 3 nm. The sum of the two ΔL_c values (31 ± 5 nm) is consistent with the full unfolding of the protein. Assuming a value of 0.36 nm/residue, a ΔL_c of 9 ± 3 nm for the *I*–*U* transition indicates that ~25 residues are structured in the intermediate state [41]. Analysis of the unfolding force distributions provided kinetic information on the unfolding transitions (Fig. 3).

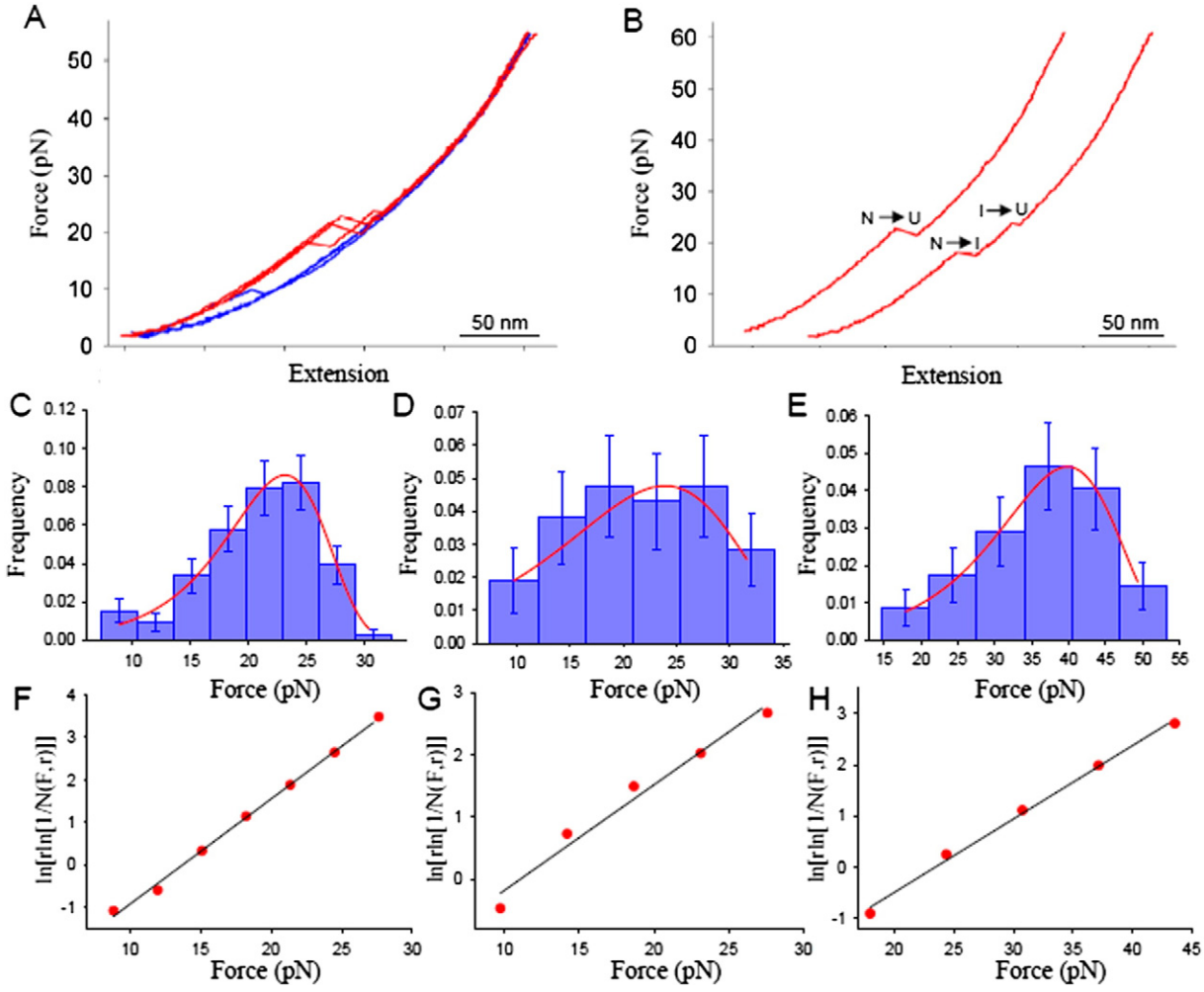


Fig. 3. Forced unfolding of the monomeric HIV-1-PR native state. A) Three superimposed stretching (red) and relaxation (blue) cycles acquired by manipulating one HIV-1-PR molecule at 100 nm/s (~5 pN/s). The molecule unfolded twice in a two-state manner, and once it populated an intermediate state *I*. B) Stretching traces from panel A) shifted laterally for clarity. The native state (*N*) transits into the unfolded state (*U*) through an all-or-none transition on the left trace, while it populates an intermediate state (*I*) in the other case. C) Force distribution of the two-state unfolding events ($n = 105$) fit to a probability distribution function [48], (red line, Eq. (2)), to estimate the position of the transition state along the reaction coordinate and the zero force unfolding rate. The best fit (least squares) values gave $x^{\ddagger}_u = 1 \pm 0.1$ nm and $k^0_u = 7\ (\pm 1) \times 10^{-3}\ s^{-1}$. D) and E) show the force distributions of the first and second unfolding events along the three-state unfolding pathway ($n = 46$) and the corresponding fit to the probability distribution function (red lines). The best fit values gave $x^{\ddagger}_u = 0.5 \pm 0.1$ nm and $k^0_u = 4\ (\pm 1) \times 10^{-2}\ s^{-1}$ for D), and $x^{\ddagger}_u = 0.5 \pm 0.1$ nm and $k^0_u = 6\ (\pm 1) \times 10^{-3}\ s^{-1}$ for E). Force distributions of C), D) and E) were also linearized and analyzed as previously reported (Eq. (3)) [24,34] in F), G) and H) respectively. This analysis provided kinetic values quite similar to those reported above, that is $x^{\ddagger}_u = 1 \pm 0.1$ nm and $k^0_u = 8\ (\pm 1) \times 10^{-3}\ s^{-1}$ for F), $x^{\ddagger}_u = 0.7 \pm 0.1$ nm and $k^0_u = 3\ (\pm 0.6) \times 10^{-2}\ s^{-1}$ for G), and $x^{\ddagger}_u = 0.6 \pm 0.1$ nm and $k^0_u = 5\ (\pm 0.7) \times 10^{-3}\ s^{-1}$ for H).

2.2. Partially folded conformations unfold along multiple pathways

During the ~50 s needed to fold, the molecule likely explores different conformations before finding its native state. To gain insight into this ensemble of partially folded conformations, we performed measurements in which individual HIV-1-PR molecules were pulled before they could reach their native state. In these experiments, a tethered molecule was stretched and relaxed multiple times between an upper and lower force value (55 pN and 2 pN, respectively), with no waiting time at low tension. Under these conditions, the whole stretching-relaxation cycle takes only 20 s to be completed. Thus, an unfolded HIV-1-PR has no time to fold back into its native state before it is pulled again, and the molecule is stretched while being in one of the many possible partially folded (*PF*) conformations it adopts *en route* to *N*, which collectively we call the *PF* state.

In these measurements, the HIV-1-PR monomer unfolds either through an all-or-none transition (23% of the cases) or, more often, in a three-state manner involving an intermediate state (71% of the cases), as shown in Fig. 4. Only rarely (6% of the cases), it unfolds through different mechanisms (Figure S1). The same molecule is observed to take different unfolding pathways when stretched and relaxed multiple times (Fig. 4).

The sharp transitions characterizing the stretching traces indicate that the *PF* state is a well-defined thermodynamic state, separated from other states of HIV-1-PR by free-energy barriers. The ΔL_c associated with the two-state unfolding events (Fig. 4B, left race) was estimated to be 31 ± 3 nm, a value that is quite similar to that measured for the unfolding of the native state, revealing a native-like extension of the *PF* conformations. Similarly, the changes in contour length associated with the three-state unfolding pathway were estimated to be 22 ± 3 nm for the first transition, and 7 ± 2 nm for the second one, summing to a total ΔL_c of 29 ± 4 nm.

The ensemble of conformations associated with the *PF* state is, overall, less stable than the native state, unfolding mostly at 10 pN (Fig. 4C). However, unfolding events at higher forces are also observed, resulting in a broad force distribution. The shape and broadness of the force distribution suggests a continuum of two-state unfolding pathways, characterized by different kinetic barriers. Interestingly, the force

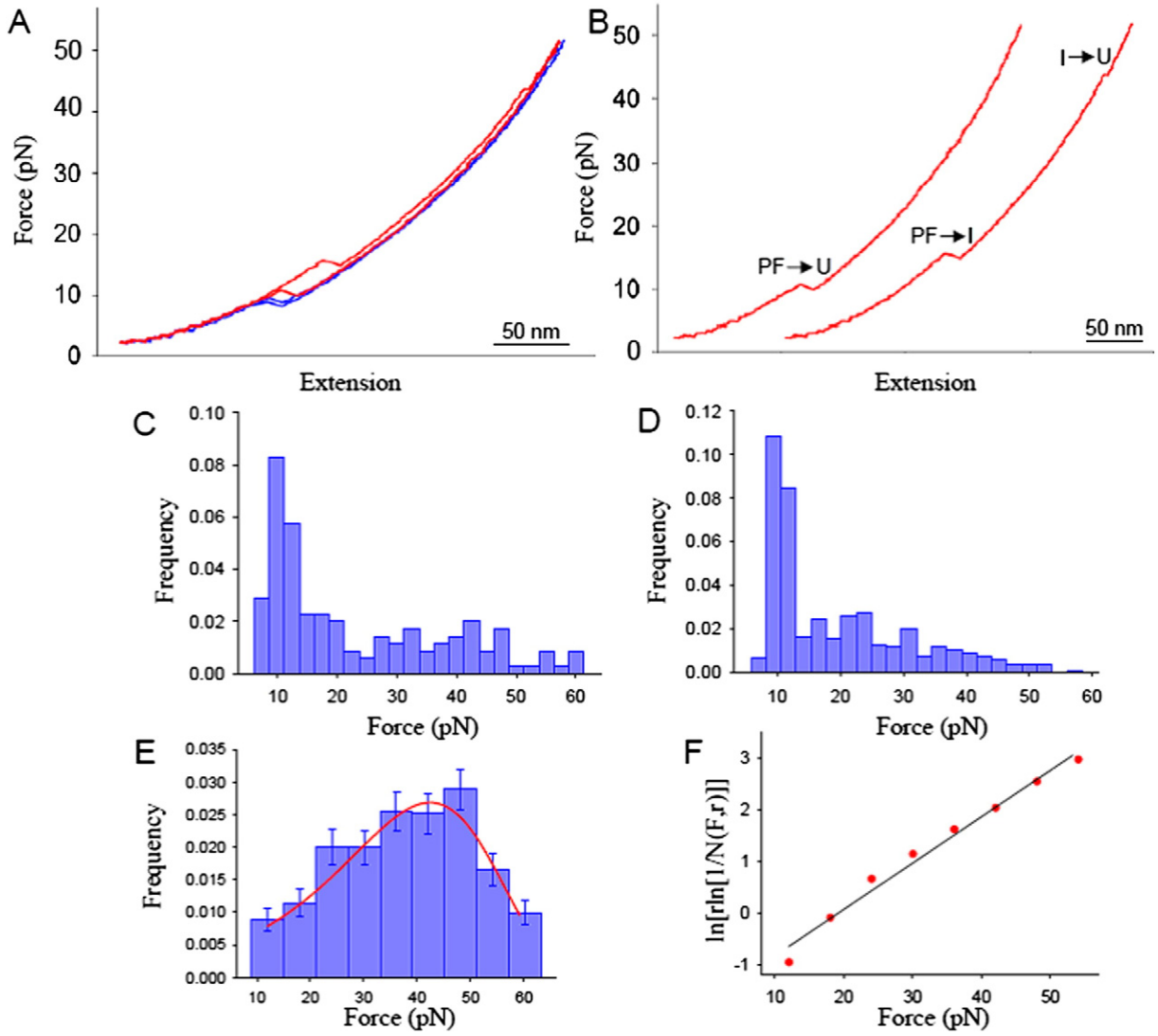


Fig. 4. Forced unfolding of partially folded (*PF*) conformations of HIV-1-PR monomer. A) Two superimposed force vs. extension cycles acquired by stretching and relaxing one HIV-1-PR molecule at 100 nm/s. B) Stretching traces of panel A), where the molecule transits into the unfolded state either through an all-or-none transition (left trace), or by populating an intermediate state *I* (right trace). C) Force distribution of the two-state unfolding events ($n = 139$). D) Force distribution of the first unfolding events along the three-state unfolding pathway ($n = 432$). E) Force distribution of the second unfolding events along the three-state unfolding pathway ($n = 432$). Fit of these data to a probability distribution function (red line, Eq. (2)) [48] yielded $x^{\ddagger}_u = 0.3 \pm 0.1$ nm and $k_0 = 3\ (\pm 0.5) \times 10^{-2}\ s^{-1}$. F) Plot of $\ln[r \ln[1/N(F,r)]]$ vs. force for the force distribution of E), and corresponding fit (black lines, Eq. (3)) [34] yielding a $x^{\ddagger}_u = 0.4 \pm 0.1$ nm and $k_0 = 2\ (\pm 0.2) \times 10^{-2}\ s^{-1}$.

distribution of the first unfolding event along the three-state unfolding pathway (Fig. 4D) has a shape similar to that of Fig. 4C. These similarities suggest that the heterogeneity of the observed unfolding pathways most likely reflects a heterogeneity of the initial states of such pathways, indicating a high structural variability of the *PF* state. After crossing the first barrier for unfolding, the molecule transits directly into the unfolded state (23% of the cases), or is kinetically trapped into an intermediate state that then unfolds at higher forces (71% of the cases). However, the unfolding force distribution of the intermediate state (Fig. 4E) is not scattered and can be well fitted by a two-state unfolding probability function, suggesting a well-defined molecular conformation for the intermediate.

2.3. HIV-1-PR monomer refolds through multiple pathways

Refolding of the monomeric HIV-1-PR is also a multi-pathway process, where the molecule regains its original extension either gradually, or through one or more transitions (Fig. 5). Notice that the refolding events displayed in this figure do not lead directly to the native state of the protein because, as mentioned above, folding into *N* is a very slow process. Mechanically denatured HIV-1-PR monomers compact into partially folded (*PF*) conformations that take at least 50 s to transit into the native state [12]. This *PF*–*N* transition is not detected in our measurements, implying that the change in extension involved in the associated folding process is too small to give rise to a noticeable signal in our recordings. These results indicate that the *PF* state is similar in compactness to the *N* state. The shape of the force distributions of Fig. 5D, E and F reveals the presence of multiple refolding pathways associated with the refolding events shown in Fig. 5B and C. No correlation has been noticed between type of refolding trajectory and type of unfolding events observed in the following stretching trace.

2.4. Molecular dynamics simulations

State-of-the-art ratcheted MD simulations of the mechanical denaturation of the HIV-1-PR monomer were carried out to gain insight into the molecular states visited by the protein during forced unfolding [42–44]. Starting from an equilibrated native-like

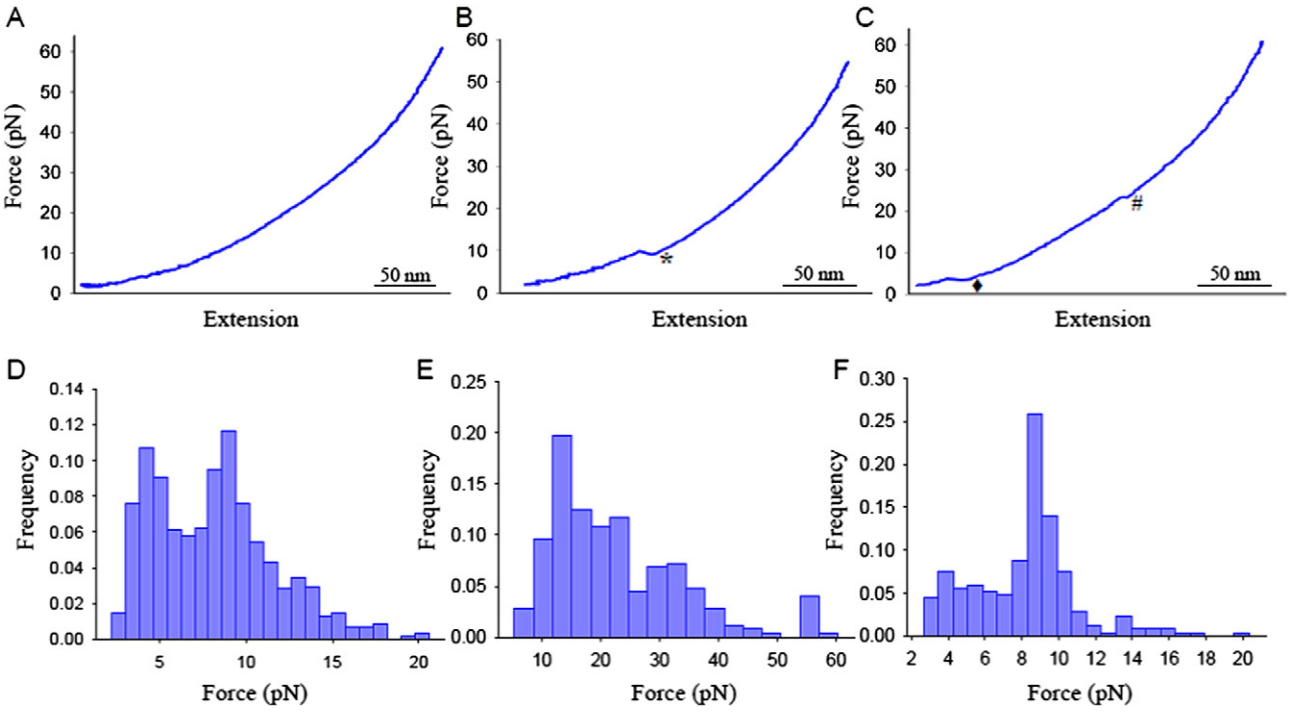


Fig. 5. HIV-1-PR monomer refolds through multiple pathways. When the tension applied on a mechanically denatured HIV-1-PR monomer is relaxed, the molecule regains its original extension through either: a gradual compaction without any sharp transition (55% of the cases, panel A); a single sharp transition associated with a ΔL_c of 31 ± 4 nm (28% of the cases, panel B), or a two-step process with a first (small, $\Delta L_c = 8 \pm 3$ nm) and second (large, $\Delta L_c = 26 \pm 5$ nm) transition (12% of the cases, panel C)). Only rarely, it regains a compact conformation through different trajectories (5% of the cases; Fig. S1). D), E) and F) show the force distributions of the refolding events indicated by *, # and ♦, respectively.

conformation with a reference distance between Cys5 and Cys98 of 2.1 nm, and using a biasing algorithm to make the simulations feasible in a reasonable amount of time (see Materials and methods for details) we calculated 7 unfolding trajectories in explicit solvent at a constant force of 20 pN. Snapshots of the molecular conformations visited by the protein during the forced denaturation are displayed in Fig. 6.

The first elements to lose structure are hairpin β1 (residues 10–23) and the α helix (residues 87–93); then, β strand 83–87 is removed from the 23–34 loop, leaving hairpins β2 (residues 43–58) and β3 (residues 59–75) as the only native secondary structure (see Fig. 2B). Only at the end of the simulation, these two β structures are disrupted, in an undefined order. The seven trajectories mainly differ from each other for number and type of non-native interactions. Fig. 7 shows the average number of non-native interactions as a function of the distance *d* between Cys5 and Cys98. Two trajectories display the formation of a limited number of non-native contacts (up to 5, points marked with #), which do not involve well-defined regions of the protein. In the other five trajectories, the number of non-native contacts can reach values between 8 and 10 (points marked with *). The regions of the chain involved in these contacts are more restrained, as shown in panel B of Fig. 7 (see also Table S1 in the Supplementary materials). If we assume that some non-native contacts can stabilize certain secondary and tertiary structures, then the observed heterogeneity along the experimental unfolding pathways (Fig. 3), in terms of presence or absence of an intermediate state, might reflect the variability in number and type of non-native contacts forming during the forced unfolding process. Moreover, by comparing the end-to-end distance *d* of the conformations displayed in Fig. 6 with the unfolding transitions (rips) observed during the experiments (Fig. 3), we can propose that hairpins β2 and β3 are structured, at least to some extent, in the intermediate state observed by optical tweezers. In fact, at $d \approx 17$ *nm*, corresponding to the rip width of the *N–I* transition along the three-state unfolding pathway (Fig. 3B), β2 and β3 are still partly formed, and surrounded by the structure-less part of the chain, with which they make a variable number of tertiary contacts.

Biased MD simulations were also carried out to gain insight into the mechanism by which denatured HIV-1-PR monomers travel toward the native state. Four refolding trajectories were simulated in explicit

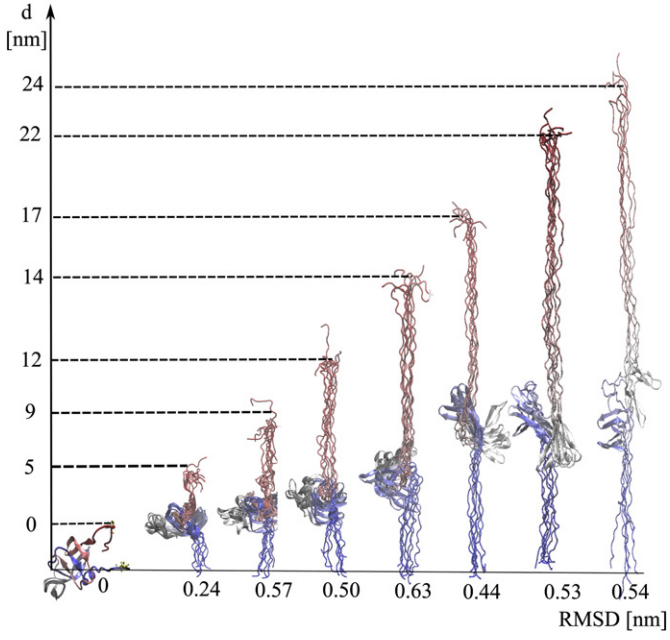


Fig. 6. Molecular dynamics simulation of the forced unfolding of the HIV-1-PR monomer. Snapshots of the molecular structures visited by the protein along the unfolding trajectories at increasing distance *d* between Cys5 and Cys98. The RMSD values indicate the structural variability within each set.

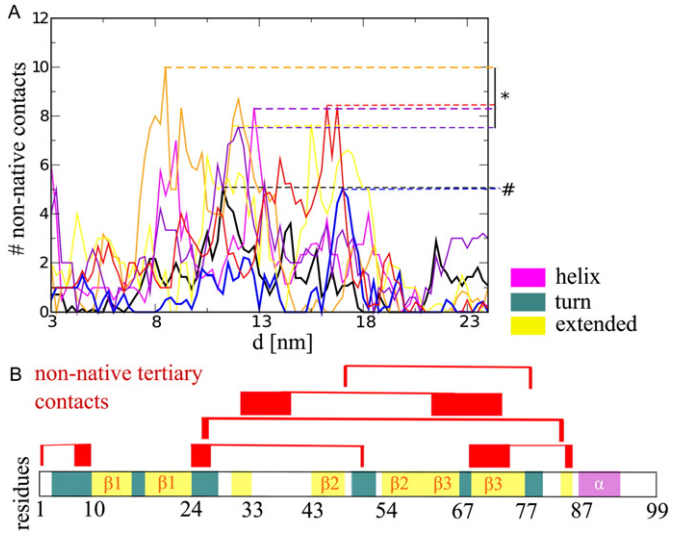


Fig. 7. Non-native contacts forming during the forced unfolding of the HIV-1-PR monomer. A) Average number of non-native contacts as a function of the distance *d* between Cys5 and Cys98. Each color marks a different unfolding trajectory. The maximum number of non-native contacts is around 5 (#) in two of them, while in the other cases it ranges between 8 and 10 (*). B) The horizontal bar highlights the secondary structures of the protein, while the red connectors indicate non-native contacts forming in at least 2% of the time in a given trajectory.

solvent at a constant force of 5 pN, starting from fully-elongated structures (right-most conformation in Fig. 6). None of the simulations reached the native state of the protein. Rather, the molecules got trapped into partially folded conformations, characterized by native secondary and tertiary contacts in hairpins β2 and β3, and non-native contacts between segments 24–34 and 83–93 (Fig. 8). The overall structures of these conformations are rather heterogeneous, displaying a mutual RMSD of 0.47 nm. However, the distance between Cys5 and Cys98 is 1.3 ± 0.4 nm, which is essentially indistinguishable from that measured in the native state (0.6 ± 0.4 nm). Therefore, the trapped conformations displayed in Fig. 8 appear to share common features with the *PF* conformations observed in single-molecule experiments, both for their native-like extension and for their structural heterogeneity

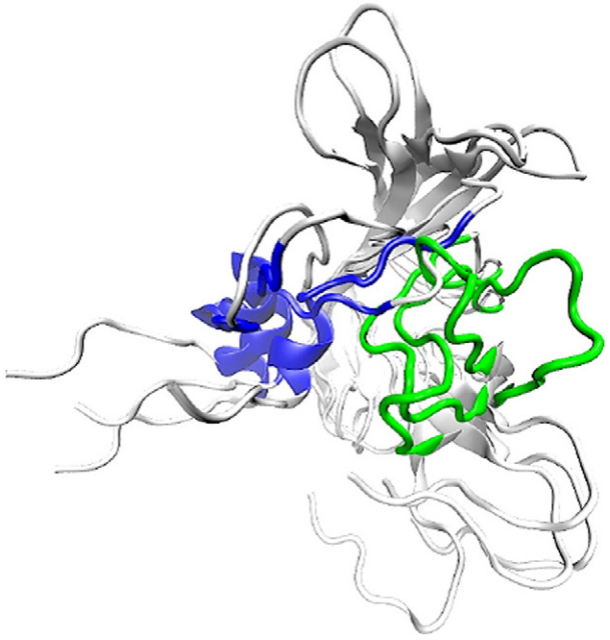


Fig. 8. Molecular dynamics simulation of the refolding process of the HIV-1-PR monomer. Four superimposed structures adopted by the protein at the end of refolding simulations carried out at 5 pN. The trajectories ended with non-native conformations in which segments 24–34 (green) and 83–93 (blue) cannot build the native β sheet.

3. Discussion

The monomer of the HIV-1-PR is a complex protein, with a non-trivial β-rich native conformation and a particularly lengthy dynamics. This complexity is not always apparent in bulk experiments, in which the overall experimental signal is the average over a very large number of molecules. In fact, the fluorescence signal recorded in the equilibrium titration experiment shown in Fig. 1 is compatible with a simple two-state model. Single-molecule force spectroscopy offers a powerful approach for deciphering the intricacies of complex folding scenarios. Through mechanical manipulation it is nowadays possible to follow the actual dynamics of single molecules as they undergo their transformations, avoiding the inherent averaging of ensemble methods. In this paper, we use optical tweezers and biased MD simulations to uncover information inaccessible to more traditional experimental techniques.

The molecular states and unfolding/refolding pathways of HIV-1-PR revealed by our studies are schematically depicted in Fig. 9. Starting from the native structure, the protein can either unfold directly, or through an intermediate state. Analysis of the force distributions of the unfolding events (Fig. 3) suggests that the first kinetic barrier of the three-state pathway is probably different from that of the two-state pathway, both because of its position along the reaction coordinate ($x^{\ddagger}_u = 0.5 \pm 0.1$ nm versus $x^{\ddagger}_u = 1 \pm 0.1$ nm, respectively) and its height ($k^0_u = 4\,(\pm 1) \times 10^{-2}\ s^{-1}$ versus $k^0_u = 7\,(\pm 1) \times 10^{-3}\ s^{-1}$, respectively). These data support the idea that the bifurcation of the energy landscape of HIV-1-PR's native state does not take place after a major kinetic barrier, common to both pathways, as reported for other proteins [45]. Rather it takes place early on during the unfolding process [20]. The structure of the intermediate state populated by the protein is not known. However, the comparison of the rip width (~17 nm) of the *N–I* transition of the three-state unfolding process (Fig. 3B) with the extension of the molecular conformations visited by the protein during the MD simulations (Fig. 6), seems to indicate that hairpins β2 and β3 are still at least partially structured in the intermediate state.

The heterogeneity of the unfolding pathways of the HIV-1-PR monomer has been previously proposed by Bonomi et al. [14], from calculations of the free-energy surface of the protein with a structure-based

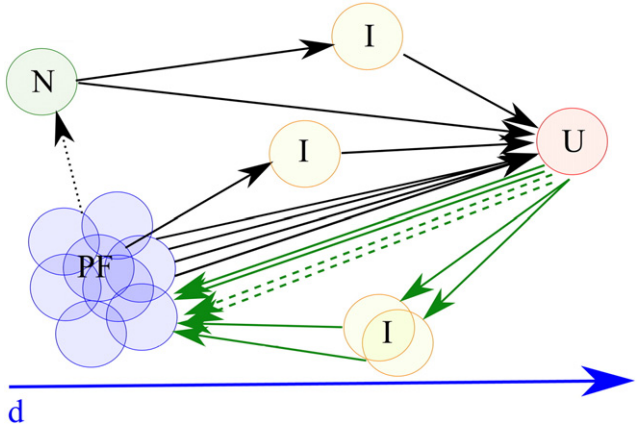


Fig. 9. Sketch of the unfolding/refolding pathways and molecular states populated by the HIV-1-PR monomer. The native state unfolds along two main pathways, one involving an intermediate state. The unfolded state follows multiple trajectories to refold first into a structurally heterogeneous *PF* state, which then transits into the native state slowly. The *PF* state can unfold either in a two-state manner or, more often, by populating an intermediate state.

model. However, no specific intermediate state was observed. These results agree well with our data if we consider that, according to our MD simulations, the intermediate state could be stabilized by non-native contacts, and thus could not be detectable in structure-based models which, by definition, do not describe non-native interactions. It should also be noticed, however, that the pathways described in this paper apply only to the folding behavior of the molecule under tension. The folding pathways followed by HIV-1-PR free in solution might differ.

Refolding of the mechanically denatured HIV-1-PR monomer into its native state is a multi-path process characterized by partially folded (*PF*) conformations that altogether act as a long-lived state controlling the slow kinetic phase of the process. In approximately half of the folding trajectories, the *PF* state is reached without encountering any major free-energy barrier (Fig. 5A). As a consequence, the overall folding kinetics is presumably solely controlled by the large free-energy barrier between the *PF* and the native state. In the other half of the cases, the HIV-1-PR monomer folds toward the *PF* state by crossing different kinetic barriers, thus generating broad refolding force distributions (Fig. 5). The *PF* state is quite compact and its transition into the native state is undetectable in our measurements. Nonetheless, despite this experimental limitation, we were able to gain insight into this long-lived state by pulling on it and generating unfolding events (Fig. 4). The resulting broad and skewed distribution of unfolding forces suggests that the *PF* state comprises a structurally heterogeneous set of conformations that unravel through a multitude of unfolding pathways. All MD simulations of the refolding process lead to trapped non-native states that might share common features with the *PF* conformations. They are characterized by a native-like arrangement of hairpins β2 and β3, but a non-native positioning of the hydrophobic core of the protein. In the native conformation, the segment 83–86 is inserted in the looped segments 21–24 and 31–34 to form a β-sheet, which is further stabilized by the β-strand 74–78. In the conformations displayed in Fig. 8 not only this β-sheet is not present, but also the relative positioning of segment 83–93 with respect to segment 24–34 is such that the molecule has to swell to reach a native-like conformation.

Despite the heterogeneity of the *PF* conformations, in ~70% of the unfolding trajectories the molecule populates an intermediate state that unravels producing a well-defined force distribution. This intermediate state appears to have similar properties to those of the intermediate state observed when pulling on the native state (Fig. 3). The height (k^0_u) and position ($x^{\ddagger}_u$) of the unfolding energy barriers are similar, and the changes in contour length associated with their unfolding are within the standard deviations. Likewise, their mechanical resistances are similar, as shown in Figs. 3E and 4E. These similarities might reflect common structural features of the two intermediate states, where hairpins β2 and β3 could be, at least partially, folded in both cases. However, our results do not allow for definitive conclusions on this matter and extracting detailed structural features of the two intermediate states will require further dedicated experiments.

Being a key enzyme in the maturation of the human immunodeficiency virus, the HIV-1-PR is the target of drugs used in anti-AIDS therapy. In essentially all these therapeutic approaches, molecules are designed to target the active site of the protease to inhibit its function. However, being associated with the error-prone replication mechanism of a retrovirus, the protease mutates under drug pressure, altering the binding site of the inhibitors and generating, in the process, pharmacological resistance. The intermediate states revealed in our study can be regarded as non-conventional targets for molecules that, by stabilizing partially folded conformations, can block the maturation of HIV-1-PR, thereby inhibiting production of mature viral particles. Much work remains to be done to implement such a strategy. In particular, a systematic NMR study of the denatured state of the HIV-1-PR by itself, and in presence of selected exogenous molecules might be very helpful in this regard. Within this context, and in keeping with the fact that folding inhibition can be viewed as an extreme form of allosteric inhibition, one can refer to Kunze et al. [46] where it is reported that structure-based computational screening of dynamic pockets of the HIV-1-PR has allowed the identification of small molecules displaying robust (IC50 ~ 10–20 μM) inhibition.

4. Materials and methods

4.1. Protein expression

Plasmid pJF19, a pXC35 derivative encoding the Q7K/C67A/C95A variant of HIV1-1-PR was a kind gift of Dr. Celia Schiffer (University of Massachusetts Medical School, Waltham, USA). Starting from this plasmid, a construct expressing the double cysteine, monomeric L5C/R87K/N98C protein variant used in single-molecule experiments was generated with the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene) in conjunction with the following oligonucleotides:

oligo1 5′-GCCGCAGATTACCTGCTGGAAACGCCCGC-3′,
oligo2 5′-CCGGTGAACATTATTGGCAAAAACCTGCTGACCCAG-3′,
oligo3 5′-GGCGCGACCCTGTGCTTTTAGGGATCC-3′,

where the codons introducing the L5C (oligo1), R87K (oligo2) and N98C (oligo3) amino acyl substitutions are underlined. The mutagenic reaction was performed according to manufacturer's instructions, while all other required DNA handling steps were carried out according to established protocols [47,48]. Both DNA strands of the insert of the resulting plasmid were sequenced to confirm the presence of the desired mutations and to rule out the introduction of unwanted ones. The engineered protein was overproduced using the *Escherichia coli* TAP106 strain as the expression host, under the growth and induction conditions reported elsewhere [12,27].

4.2. Protein purification

The recombinant protein was isolated from inclusion bodies under denaturing conditions, and stored frozen in dried form until needed. Cells obtained from a 5 l culture were resuspended in 20 mM Tris–HCl, pH 7.5, containing 1 mM EDTA, 1 mM 2-mercaptoethanol (buffer A) and 0.2 mg/ml mg/ml lysozyme, and disrupted by applying five cycles of 1 min sonication (Branson, Model 250 sonicator) on ice. After centrifugation of the homogenate at 20,000 ×*g* for 20 min at 20 °C, the recombinant protein was purified from inclusion bodies according to the protocol described in [49]. Briefly, inclusion bodies were washed once with 2 M urea, 0.2% Triton X-100 in buffer A, once with 2 M urea in buffer A, and twice with buffer A. For each washing, 5 ml of buffer per gram of the initial cell pellet was used and the resuspended sample was centrifuged at 20,000 ×*g* for 30 min at 20 °C. The pellet was dissolved in glacial acetic acid (60 mg/ml) and then diluted with 50% acetic acid to a concentration of 12 mg/ml. After centrifugation at 45,000 ×*g* for 30 min, the supernatant was passed through a 0.22 μm filter and chromatographed on a Sephadex G-75 Superfine column (2.6 cm × 85 cm, 450 ml) equilibrated and eluted with 50% acetic acid. Fractions containing homogenous protein, as judged by SDS-PAGE, were pooled, brought to dryness under vacuum and stored at −80 °C. The resulting protein was monomeric as judged by dynamic light scattering measurements and analytical gel filtration.

4.3. Protein preparation for single-molecule experiments

DNA handles were prepared essentially as described by [27,31]. The following oligonucleotides were obtained from Primm srl (Milan, Italy) and they were used to amplify a 558 bp portion of the pGEMEX-1 vector.

oligo4 5′Biotin-CAAAAAACCCCTCAAGACCC-3′.
oligo5 5′Digoxigenin-CAAAAAACCCCTCAAGACCC-3′.
oligo6 5′Thiol-GCTACCGTAATTGAGACCAC-3′.

Handles (400–500 μg) were generated by PCR using ten 1.2 ml reaction mixtures. The two types of handles were generated using oligo6 with either oligo4 or oligo5. The reaction mixture contained 0.5 μM oligo4 or oligo5, 0.5 μM oligo6, 0.04 ng/μl p-GEMEX, 0.025 U/μl *Taq*-polymerase in PCR Buffer (Qiagen) 16.2 mM DTT, 1.5 mM $MgCl_2$. Amplification was performed by applying 34 cycles consisting of denaturation at 94 °C for 45 s, annealing at 67 °C for 40 s for oligo4 or annealing at 63 °C for 40 s for oligo5, and extension at 72 °C for 90 s.

After purification with a HiSpeed Plasmid Maxi Kit (Qiagen), the DNA handles were attached to the protein by following the procedure described in [27,31]. Considering the observed tendency of the protease to aggregate at concentrations above 1 mg/ml, activation of cysteine residues was carried out on the denatured protein, which was then refolded before the attachment of the DNA handles.

The lyophilized purified protease (1 mg) was resuspended in 20 mM Na-phosphate, pH 6.5, containing 1 mM DTT and 7 M guanidine-HCl (200 μl). After 30 min incubation at 25 °C the sample was diluted to 3 mg/ml protein concentration (270 μM) with 20 mM Na-phosphate, pH 6.5, 1 mM DTT to lower the guanidine-HCl concentration to 4 M. DTT was removed by centrifugal chromatography on Bio Spin P6 columns (Bio-Rad) equilibrated and eluted with 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl. The resulting protein (typically 80–100 μM) was reacted with 2- to 5-fold molar excess of dithiodipyridine (DTDP) added from a 10 mM stock solution in 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl, 15% (v/v) acetonitrile. The reaction mix was incubated at room temperature for 30 min, which was sufficient to bring about the complete derivatization of the protease cysteine residues, as judged spectrophotometrically by monitoring absorbance changes at 343 nm [27,31].

The sample was then diluted to a protein concentration of 0.33 mg/ml with 20 mM Na-phosphate, pH 6.5, 4 M guanidine-HCl. Excess DTDP was removed by dialysis at 4 °C against 0.1 M formic acid (two 1 l changes for 4 h each) and 30 mM formic acid, pH 2.8 (1 l, overnight). After centrifugation at 19,000 ×*g* for 30 min at 4 °C to remove precipitated material, protein refolding was started by sample dilution in 5 volumes of 10 mM Na-acetate, pH 6.0, to shift the pH to 4.0, and dialysis for 4 h at 4 °C against 4 l of 20 mM Na-phosphate buffer, pH 6.0, which was made partially anaerobic by bubbling nitrogen gas. After concentration in centrifugal filter units (Millipore) followed by centrifugation at 16,000 ×*g* for 30 min at 4 °C, the protein (0.6–0.8 mg/ml) was stored in aliquots at −80 °C.

Attachment of the DNA handles to the protein was performed as described in [27] using an optimized DNA/protein ratio of 1:1 (mol/mol) instead of the 5:1 ratio suggested in Cecconi et al. [27]. Attachment of the DNA handles to the protein was verified by SDS-PAGE on 6% gels prepared in TBE buffer. The coupling of protein–DNA chimera to polystyrene beads was performed as described in Cecconi et al. [27].

4.4. Denaturation/renaturation studies

Protein was refolded as described above by skipping the cysteine activation step. Protein folding and unfolding were studied by monitoring fluorescence emission changes of protease samples (2 μM) in 20 mM Na-phosphate buffer, pH 6.0, 1 mM DTT and decreasing or increasing urea concentration at 25 °C with a Cary Eclipse fluorimeter (Varian; excitation light, 290 nm). At least 5 spectra were recorded after reaching equilibrium. The maximum wavelength of the emission spectrum was plotted as a function of urea concentration and the data were fitted to the two-state process equation:

$$\lambda_{\max\,obs} = \lambda_{\max\,F} + (\lambda_{\max\,U} - \lambda_{\max\,F}) * (\exp((-m * x + \Delta G^{\circ})/(RT)) / (1 + (\exp((-m * x + \Delta G^{\circ})/(RT))))) \quad (1)$$

where ΔG° is the unfolding free energy, x is the denaturant concentration, R is the gas constant (8.314 J/(K * mol)), T is the absolute temperature (283 K), m is a measure of the dependence of ΔG on the denaturant concentration, $\lambda_{\max\,F}$ is the maximum wavelength of the native emission spectrum and $\lambda_{\max\,U}$ is the maximum wavelength of the unfolded emission spectrum. This analysis yielded a $\Delta G^{\circ} = 9.84 \pm 1.34$ kJ/mol.

4.5. Optical tweezers experiments

Single-molecule manipulation experiments were performed using a custom-built optical tweezers instrument with a dual-beam laser trap of 840 nm wavelength [25]. The experiments were conducted at room temperatures in 10 mM Tris–HCl buffer, pH 7.0, 250 mM NaCl, 10 mM $CaCl_2$ or in 20 mM Na-phosphate buffer, pH 6.0, with similar results. DNA–protein constructs were manipulated between a 3.10 μm antidigoxigenin-coated bead (Spherotec) held in the optical trap, and a 2.2 μm streptavidin-coated bead (Spherotec) held at the end of a micropipette by suction. The force applied to the molecule was varied by moving the micropipette relative to the optical trap by means of a piezoelectric flexure stage (MAX311/M, Thorlabs, Newton, NJ). The applied force was determined by measuring the change in light momentum of the laser beams leaving the trap [30], while changes in the extension of the molecule were determined by video microscopy [30]. The molecules were stretched and relaxed at a constant speed of 100 nm s^{-1}, while force and molecular extension were recorded at a rate of 40 Hz. Only those molecules that displayed the characteristic DNA overstretching transition at 67 pN were used in the analysis [19].

4.6. Changes in contour length of HIV-1 PR

Theoretical changes in contour lengths (ΔL_c) upon unfolding/refolding of HIV-1 PR were calculated as described [19]. The distance between positions 5 and 98 in the native structure was measured using the NMR structure of HIV-1-PR (Protein Data Bank code 1BVG).

4.7. Analysis of unfolding force distributions

Unfolding force distributions of two-state unfolding events were fit to the probability distribution function [34]:

$$P(F) = \left(k^0{}_u / r\right) \exp\left(F x^{\ddagger}{}_u / k_B T\right) \exp\left[-\left(\left(k^0{}_u k_B T\right) / \left(r x^{\ddagger}{}_u\right)\right)\left(\exp\left(F x^{\ddagger}{}_u / k_B T\right) - 1\right)\right] \quad (2)$$

where $k^0{}_u$ is the unfolding rate at zero force, r is the loading rate (pN s^{-1}), $x^{\ddagger}{}_u$ is the distance from the native state to the transition state along the reaction coordinate, k_B is the Boltzmann's constant and T is the absolute temperature.

The same force distributions were also fit to the equation [25,35]:

$$\ln\left[r \ln 1/N(F,r)\right]] = \ln\left(k^0{}_u k_B T / x^{\ddagger}{}_u\right) + \left(x^{\ddagger}{}_u / k_B T\right) F \quad (3)$$

where $N(F,r)$ is the fraction of folded molecules at force F and loading rate r (pN s^{-1}). $N(F,r)$ can be calculated by integrating the unfolding force distributions over the corresponding range of forces. Eq. (3) can be used to fit $\ln[r\ln[1/N(F,r)]]$ vs. F graphs.

4.8. Simulations

Unfolding simulations were carried out with the Amber03 force field [47], using a modified version of the GROMACS package [50,51] . The system is solvated with 35,000 water molecules in a fixed box of sides $6 \times 6 \times 30$ nm. The temperature is maintained at 300 K by Nosè–Hoover thermostat [52].

First, we carried out a 100 ns MD simulation at constant forces of 8 pN, 16 pN, 22 pN and 35 pN. This time is much shorter than that

needed for the system to undergo any consistent conformational change. Thus, the simulations provided reference distances d between the Cys5 and Cys98 positions. According to these simulations (Fig. S2), the application of a force on the two cysteine residues up to 22 pN sets the protein in a stationary initial conformation characterized by varying average values of d, ranging from 0.2 nm (calculated with respect to the native conformation) at 8 pN to 2.1 nm at 22 pN. Application of a force of 35 pN does not increase the average value of d beyond 2.1 nm. Such increase of d is due to the fact that the termini of the chain, which eventually, in the presence of another monomer, build the dimeric interface of HIV-1-PR, are weakly stabilized in the monomer, and consequently can be unraveled even by the application of small forces (Fig. S3). Since the transitions observed in the optical tweezer experiments occur typically at forces of ≈20 pN or higher, we used 2.1 nm as the reference distance between the cysteine residues in the native state.

The biasing scheme is that described in Ref. [25]. However, the ratchet is here applied both on the distance d between Cys5 and Cys98 and on the RMSD to a fully-unfolded, linear conformation of the protein. The biasing potential has thus the form

$$U_{rat}(t) = \frac{1}{2}k_d\left(\rho_d(t) - min_{t'<t}\,\rho_d\left(t'\right)\right)^2 + \frac{1}{2}k_r\left(\rho_r(t) - min_{t'<t}\,\rho_r\left(t'\right)\right)^2, \quad (4)$$

where

$$\rho_d(t) = (d(t) - 30\text{nm})^2$$

and

$$\rho_r(t) = [\text{RMSD}(t)]^2,$$

the harmonic constants being $k_d = k_r = 1$ kJ/mol/nm^2.

A contact between residues is defined if any pair of their atoms is closer than 0.35 nm. Non-native contacts are calculated as contacts between residues which are separated by at least 3 other residues along the chain [53], and which are at least 2 residues away from pairs of residues which form native contacts. They are considered relevant if they appear in at least 2% of the time in a given trajectory.

Acknowledgments

C.C. gratefully acknowledges the financial support from Fondazione Cassa di Risparmio di Modena, the EU through a Marie Curie International Re-Integration Grant (n. 44952), the Italian MIUR (Grant n. 17DPXLNBEK), and partial support from Italian MIUR FIRB RBPR05JH2P "ITALNANONET". We thank Celia Schiffer (University of Massachusetts Medical School, Worcester, MA, USA) for the plasmid. Financial support from FoldLESs S.r.l (Monza, Italy) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bpc.2014.08.001.

References

[1] J.R. Huff, HIV protease: a novel chemotherapeutic target for AIDS, J. Med. Chem. 34 (1991) 2305–2314.
[2] J.M. Louis, A. Aniana, I.T. Weber, J.M. Sayer, Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors, Proc. Natl. Acad. Sci. 108 (2011) 9072–9077.
[3] J.M. Louis, E.M. Wondrak, A.R. Kimmel, P.T. Wingfield, N.T. Nashed, Proteolytic processing of HIV-1 protease precursor, kinetics and mechanism, J. Biol. Chem 274 (1999) 23437–23442.
[4] R. Ishima, D.A. Torchia, J.M. Louis, Mutational and structural studies aimed at characterizing the monomer of HIV-1 protease and its precursor, J. Biol. Chem. 282 (2007) 17190–17199.
[5] Y. Levy, A. Caflisch, J.N. Onuchic, P.G. Wolynes, The folding and dimerization of HIV-1 protease: evidence for a stable monomer from simulations, J. Mol. Biol. 340 (2004) 67–79.
[6] T. Yamazaki, A.P. Hinck, Y.-X. Wang, L.K. Nicholson, D.A. Torchia, P. Wingfield, S.J. Stahl, J.D. Kaufman, C.-H. Chang, P.J. Domaille, P.Y.S. Lam, 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. 5 (1996) 495–506.
[7] A.G. Tomasselli, R.L. Heinrikson, Targeting the HIV-protease in AIDS therapy: a current clinical perspective, Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 1477 (2000) 189–214.
[8] S. Kimura, M. Caldarini, R.A. Broglia, N.V. Dokholyan, G. Tiana, The maturation of HIV-1 protease precursor studied by discrete molecular dynamics, Proteins 82 (2014) 633–639.
[9] R.A. Broglia, Y. Levy, G. Tiana, HIV-1 protease folding and the design of drugs which do not create resistance, Curr. Opin. Struct. Biol. 18 (2008) 60–66.
[10] R.A. Broglia, G. Tiana, L. Sutto, D. Provasi, F. Simona, Design of HIV-1-PR inhibitors that do not create resistance: blocking the folding of single monomers, Protein Sci. 14 (2005) 2668–2681.
[11] R. Zutshi, J. Franciskovich, M. Shultz, B. Schweitzer, P. Bishop, M. Wilson, J. Chmielewski, Targeting the dimerization interface of HIV-1 protease: inhibition with cross-linked interfacial peptides, J. Am. Chem. Soc. 119 (1997) 4841–4845.
[12] A.F. Noel, O. Bilsel, A. Kundu, Y. Wu, J.A. Zitzewitz, C.R. Matthews, The folding free-energy surface of HIV-1 protease: insights into the thermodynamic basis for resistance to inhibitors, J. Mol. Biol. 387 (2009) 1002–1016.
[13] K.W. Plaxco, K.T. Simons, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins, J. Mol. Biol. 277 (1998) 985–994.
[14] M. Bonomi, A. Barducci, F.L. Gervasio, M. Parrinello, Multiple routes and milestones in the folding of HIV-1 protease monomer, PLoS ONE 5 (2010) e13208.
[15] M.J. Todd, N. Semo, E. Freire, The structural stability of the HIV-1 protease, J. Mol. Biol. 283 (1998) 475–488.
[16] S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Real time NMR monitoring of local unfolding of HIV-1 protease tethered dimer driven by autolysis, FEBS Lett. 497 (2001) 59–64.
[17] S.K. Grant, I.C. Deckman, J.S. Culp, M.D. Minnich, I.S. Brooks, P. Hensley, C. Debouck, T.D. Meek, Use of protein unfolding studies to determine the conformational and dimeric stabilities of HIV-1 and SIV proteases, Biochemistry 31 (1992) 9491–9501.
[18] D. Xie, S. Gulnik, L. Collins, E. Gustchina, L. Suvorov, J.W. Erickson, Dissection of the pH dependence of inhibitor binding energetics for an aspartic protease: direct measurement of the protonation states of the catalytic aspartic acid residues, Biochemistry 36 (1997) 16166–16172.
[19] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule, Science 309 (2005) 2057–2060.
[20] E.A. Shank, C. Cecconi, J.W. Dill, S. Marqusee, C. Bustamante, The folding cooperativity of a protein is controlled by its chain topology, Nature 465 (2010) 637–640.
[21] M. Shayegan, N. Rezaei, N.H. Lam, T. Altindal, A. Wieczorek, N.R. Forde, Probing multiscale mechanics of collagen with optical tweezers, Proc. SPIE 8810 (2013) 88101P.
[22] N. Varongchayakul, S. Johnson, T. Quabili, J. Cappello, H. Ghandehari, S.D.J. Solares, W. Hwang, J. Seog, Direct observation of amyloid nucleation under nanomechanical stretching, ACS Nano 7 (2013) 7734–7743.
[23] A. Borgia, P.M. Williams, J. Clarke, Single-molecule studies of protein folding, Annu. Rev. Biochem. 77 (2008) 101–125.
[24] D. Aioanei, S. Lv, I. Tessari, A. Rampioni, L. Bubacco, H. Li, B. Samorì, M. Brucale, Single-molecule-level evidence for the osmophobic effect, Angew. Chem. Int. Ed. 50 (2011) 4394–4397.
[25] P.O. Heidarsson, I. Valpapuram, C. Camilloni, A. Imparato, G. Tiana, F.M. Poulsen, B.B. Kragelund, C. Cecconi, A highly compliant protein native state with a spontaneous-like mechanical unfolding pathway, J. Am. Chem. Soc. 134 (2012) 17068–17075.
[26] G. Arad-Haase, S.G. Chuartzman, S. Dagan, R. Nevo, M. Kouza, B.K. Mai, H.T. Nguyen, M.S. Li, Z. Reich, Mechanical unfolding of acylphosphatase studied by single-molecule force spectroscopy and MD simulations, Biophys. J. 99 (2010) 238–247.
[27] C. Cecconi, E.A. Shank, S. Marqusee, C. Bustamante, DNA molecular handles for single-molecule protein-folding studies by optical tweezers, Methods Mol Biol, 749, The Humana Press Inc, New York, U.S.A., 2011, pp. 255–271.
[28] R. Ishima, R. Ghirlando, J. Tozser, A.M. Gronenborn, D.A. Torchia, J.M. Louis, Folded monomer of HIV-1 protease, J. Biol. Chem. 276 (2001) 49110–49116.
[29] J.M. Louis, R. Ishima, I. Nesheiwat, L.K. Pannell, S.M. Lynch, D.A. Torchia, A.M. Gronenborn, Revisiting monomeric HIV-1 Protease: characterization and redesign for improved properties, J. Biol. Chem. 278 (2003) 6085–6092.
[30] S.B. Smith, Y. Cui, C. Bustamante, [7] Optical-trap force transducer that operates by direct measurement of light momentum, in: M. Gerard, P. Ian (Eds.), Methods in Enzymology, 361, Academic Press, 2003, pp. 134–162.
[31] C. Cecconi, E. Shank, F. Dahlquist, S. Marqusee, C. Bustamante, Protein-DNA chimeras for single molecule mechanical folding studies with the optical tweezers, Eur. Biophys. J. 37 (2008) 729–738.
[32] M. Carrion-Vazquez, A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, J.M. Fernandez, Mechanical and chemical unfolding of a single protein: a comparison, Proc. Natl. Acad. Sci. 96 (1999) 3694–3699.
[33] J. Stigler, F. Ziegler, A. Gieseke, J.C.M. Gebhardt, M. Rief, The complex folding network of single calmodulin molecules, Science 334 (2011) 512–516.
[34] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, Biophys. J. 72 (1997) 1541–1555.
[35] J. Liphardt, B. Onoa, S.B. Smith, I. Tinoco, C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force, Science 292 (2001) 733–737.

[36] C. Bustamante, J. Marko, E. Siggia, S. Smith, Entropic elasticity of lambda-phage DNA, Science 265 (1994) 1599–1600.
[37] J.F. Marko, E.D. Siggia, Stretching DNA, Macromolecules 28 (1995) 8759–8770.
[38] M. Schlierf, H. Li, J.M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 7299–7304.
[39] H. Dietz, M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 16192–16197.
[40] H.C. Kotamarthi, R. Sharma, Sri R. Koti Ainavarapu, Single-molecule studies on polySUMO proteins reveal their mechanical flexibility, Biophys. J. 104 (2013) 2273–2281.
[41] G. Yang, C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F. W. Dahlquist, C. Bustamante, Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme, Proc. Natl. Acad. Sci. 97 (2000) 139–144.
[42] M. Marchi, P. Ballone, Adiabatic bias molecular dynamics: a method to navigate the conformational space of complex molecular systems, J. Chem. Phys. 110 (1999) 3697–3702.
[43] E. Paci, M. Karplus, Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations, J. Mol. Biol. 288 (1999) 441–459.
[44] C. Camilloni, R.A. Broglia, G. Tiana, Hierarchy of folding and unfolding events of protein G, CI2, and ACBP from explicit-solvent simulations, J. Chem. Phys. 134 (2011).
[45] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules, Proc. Natl. Acad. Sci. 103 (2006) 12724–12728.
[46] J. Kunze, N. Todoroff, P. Schneider, T. Rodrigues, T. Geppert, F. Reisen, H. Schreuder, J. Saas, G. Hessler, K.-H. Baringhaus, G. Schneider, Targeting dynamic pockets of HIV-1 protease by structure-based computational screening for allosteric inhibitors, J. Chem. Inf. Model. 54 (2014) 987–991.
[47] F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Sthrul, Current Protocols in Molecular Biology, John Wiley & Sons. Inc., Hoboken NJ, 2005.
[48] T. Maniatis, E.F. Frisch, J. Sambrook, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA, 1988.
[49] J. Hui, A. Tomasselli, I. Reardon, J. Lull, D. Brunner, C.-S. Tomich, R. Heinrikson, Large scale purification and refolding of HIV-1 protease from *Escherichia coli* inclusion bodies, J. Protein Chem. 12 (1993) 323–327.
[50] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[51] C. Camilloni, A. Guerini Rocco, I. Eberini, E. Gianazza, R.A. Broglia, G. Tiana, Urea and guanidinium chloride denature protein L in different ways in molecular dynamics simulations, Biophys. J. 94 (2008) 4654–4661.
[52] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. 81 (1984) 511–519.
[53] L. Bellucci, S. Corni, R. Di Felice, E. Paci, The Structure of Neuronal Calcium Sensor-1 in Solution Revealed by Molecular Dynamics Simulations, PLoS ONE 8 (2013) e74383.