Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology

HIV-1 protease folding and the design of drugs which do not create resistance

RA Broglia[1,2], Y Levy[3] and G Tiana[1]

Human immunodeficiency virus type 1 (HIV-1) protease (PR) plays an essential role in the life cycle of the virus. Consequently, its inhibition can control acquired immunodeficiency syndrome (AIDS). Any pharmacological treatment targeting the active site of the protease is known to generate escape mutants. On the other hand, if a drug targets a site crucial for the correct folding of the protease, mutations affecting this region would denaturate the protein and thus will not be expressed. We review the progress in our understanding of the folding of the protease, which has been instrumental in the design of a (non-conventional) folding inhibitor. The transferability of these results to other proteins testify to the universality of the folding–inhibition scenario for the design of leads of drugs which are unlikely to generate resistance.

Addresses
[1] Department of Physics, University of Milano and INFN, sez. Milano, 20133 Milano, Italy
[2] The Niels Bohr Institute, University of Copenhagen, 2100 Copenhagen, Denmark
[3] Weizmann Institute of Science, Department of Structural Biology, Rehovot 76100, Israel

Corresponding author: Broglia, RA (ricardo.broglia@mi.infn.it)

Current Opinion in Structural Biology 2008, **18**:60–66

This review comes from a themed issue on
Folding and binding
Edited by Laura Itzhaki and Peter Wolynes

Available online 21st December 2007

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2007.10.004

Introduction

Many proteins carry out their biological function as multimeric assemblies. In particular, in the form of dimers, that is proteins whose native conformation is a globule built out of two disjoint chains [1–6]. If they are equal, as in the case of the HIV-1-PR, one has to do with a homodimer. The HIV-1 PR has an important role in the maturation of the virus and as such inhibiting its enzymatic activity is of high interest. Beyond the design of inhibitors that block the active site – very effective but of restricted efficacy because of the insurgence of resistance [7–11] – novel inhibition approaches have been proposed in recent years aiming at destabilizing the dimeric [12–15,16••] or the monomeric [16••,17••] forms of the HIV-1 protease (Figure 1a).

A dimerization inhibitor is based on a molecule that mimics the intertwined interface of the protease [12,13]. While this approach is particularly attractive, the large size of the designed peptide makes it only partially effective. Alternatively, it has been recently proposed – although not yet tested – the design of a small molecule that binds to regions of the protease whose dynamics are correlated with the opening motion of the flaps, so as to (allosterically) inhibit substrate entry [18] (see also [19,20]). This opinion, however, focuses on another class of inhibitors that can regulate the activity of the protease by inhibiting the folding of the protease monomers [16••,17••], in particular through a mechanism unlikely to create resistance [17••].

In neutral solution, the HIV-1 protease folds according to a three state mechanism ($2D \leftrightarrow 2N \leftrightarrow N_2$, [1–6]), populating consistently the monomeric native conformation, as found in sedimentation equilibrium experiments [21]. This result is supported by NMR studies of mutants in which the interaction across the interface is weakened [22,23], by all-atom simulations of the monomer in explicit solvent [24] and by perfectly funneled (Gō model[4] [25,26]) energy landscape simulations of the dimer [16••]. The dimer dissociation constant ($2N \leftrightarrow N_2$), is found to be $K_D = 5.8\ \mu M$ at 4 °C [21]: for instance, in a 30 μM solution, 44% of the protein is in a monomeric form.

The above-mentioned results allow one to conclude that, at neutral pH, each monomer of the protein folds[5] ($D \leftrightarrow N$) independently of each other, through the same folding mechanism of single-domain (monomeric) proteins. After the monomer has reached its native state N, it diffuses to bind another folded chain ($2N \leftrightarrow N_2$). The destabilization of the monomer will thus lead to enzymatic inhibition.

Folding of single-domain (monomeric) proteins

Upon titration with denaturants like urea or guanidinium chloride, small, single domain proteins typically display a highly concerted transition between the denatured (D)

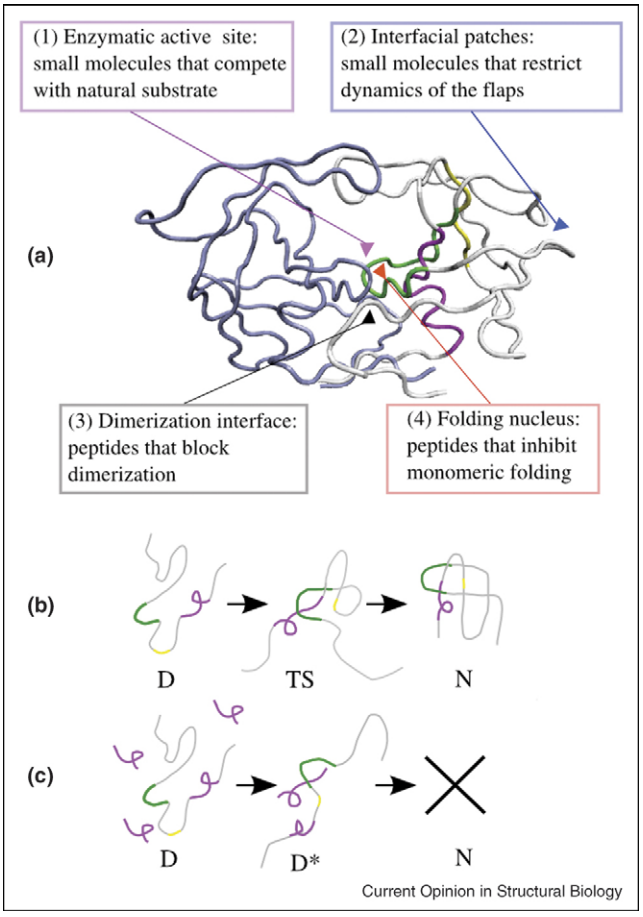
[4] In this model, that is constructed on the basis of native topology, the interaction between each pair of amino acids is described by a potential whose minimum lies at the values of the native pairs and at zero or positive energy for non-native pairs. Such a treatment that corresponds to a perfectly funneled energy landscape, insures the native conformation to be the global minimum of the system, providing at the same time a realistic description of the entropic features of the chain.

[5] We prefer to talk about the denatured state D, instead of the unfolded state, in keeping with the fact that there exists conspicuous evidence for the presence of native contacts and considerable secondary structure in the state of proteins which in cells and tissues is in equilibrium with the native state [27–29].

and the native (N) state [30,31]. This all-or-none process,[6] taking place in very short times, is consistent with a funnel-shaped energy landscape, whose bottom corresponds to the native state, and whose rational is related to 'minimal frustration' [33,34]. In other words, evolution has optimized protein sequence in such a way that the native state has not to compete with denatured conformations [35,36]. The funneled energy landscape suggests that folding is not a random process but is guided towards the folded state and may include several folding routes. Furthermore, there is experimental evidence that typical pathways in the energy landscape involve the structuring of well-defined segments of the protein [37–39]. The region around the transition state (TS) can be characterized by ϕ-value analysis,[7] and usually shows a small, well-defined set of native contacts, referred to as (critical) folding nucleus (FN, [41,42,43••]). Also the denatured state is far from a random coil and displays elements of both native and non-native structure, as shown by NMR [44–46] and fluorescence experiments [47] (see also footnote 5). These findings are consistent with the scenario emerged from simulations, which testify to the fact that folding may proceed by the formation of Local Elementary Structures (LES) that can be of various stability and their assembly is closely related to the transition state (TS) ensemble and to the folding nucleus (FN).[8]

Dynamics simulations carried out on a lattice [48,49] demonstrate that the folding of a small monomeric protein, starting from an unfolded (random coil) conformation, follows a hierarchical succession of events (cf. Figure 1b; see also [50–53]): (1) formation of local elementary structures (LES), hidden, incipient secondary structures (see also [54••]) stabilized by few, highly conserved, strongly interacting, hydrophobic ('hot') amino acids lying close along the polypeptide chain[9]; (2) docking of the LES into the (postcritical) FN (this structure is similar to that associated with the FN, but is more committed to folding. In other words, the (postcritical) FN is 'the minimal sized fragment of the native state which inevitably grows to the native state' [43••]) and (3) relaxation of the remaining amino acids in the native structure shortly after the formation of the (postcritical) FN. Note that a hierarchical succession of events is not incompatible with the cooperative, two-state folding highlighted by calorimetry and kinetic measurements. In fact, the large entropy of the denatured conformations make them behave as a single thermodynamic state even if they contain different degrees of native structure (see the model of ref. [56]). Moreover, the exponential dependence of the kinetic rates with the height of the free energy barriers allow the experiments to describe only the processes taking place on the longest timescale, hiding all faster processes.

Figure 1



(a) Four possible target sites for HIV-1 protease inhibition. The two monomers of the homodimeric protease are colored white and light blue. The LES are highlighted by green (23–33), yellow (74–78) and magenta (84–92). **(b)** Schematic representation of a unimolecular protein folding (e.g. the folding of protease monomer) highlighting the role of LES in defining the TS). **(c)** Folding inhibition through a bimolecular docking of the p-LES to the protease. Inhibition is carried out by trapping the protein in a partially denatured state, D^*, whose free energy is lower than that of the native state, N.

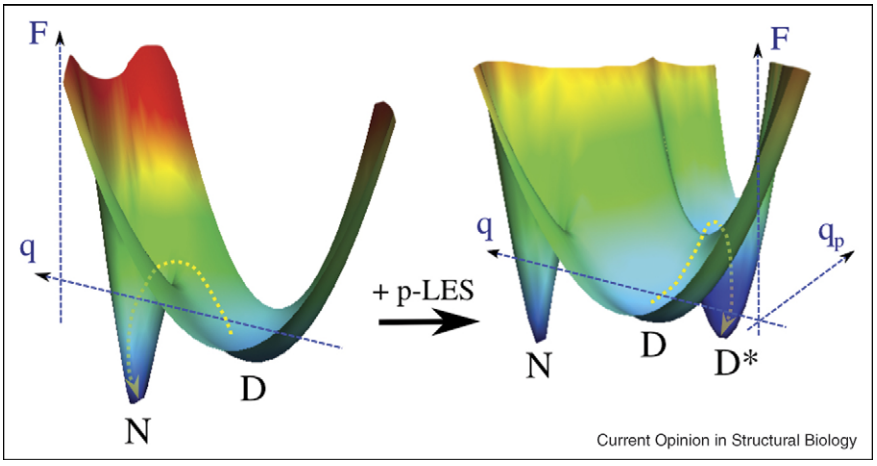
[6] Note that some proteins are proposed to display downhill folding mechanism in the absence of a significant free-energy barrier and are therefore not consistent with the all-or-none scenario [32].

[7] The quantitative measure of participation of an amino acid in a TS is a parameter called ϕ-value, which represents the ratio of the change in activation free energy upon mutation to the change in the free energy of protein stabilization. A low (high, close to 1) ϕ-value suggests that the corresponding residue does not (does) form its native contacts in the TS [40••].

[8] Transition state (TS) ensemble: the thermodynamic state associated with the top of the main free energy barrier (FEB) that separates the native from the denatured state. (Critical) folding nucleus (FN): the ensemble of conformations of the protein chain at the transition state. (Postcritical) FN: the minimal set of native contacts needed to overcome FEB and insure folding [43••]. Local elementary structures (LES): hidden, incipient secondary structures formed very early in the folding process, stabilized by few, highly conserved, strongly interacting, hydrophobic amino acids. Their docking gives rise to the (postcritical) FN.

[9] LES are likely to be related to the 'foldons' introduced in ref. [55].

Figure 2



A sketch of the effect of p-LES on the free-energy landscape of protein folding, plotted with respect to the fraction q of intra-molecular native contacts and q_p of inter-molecular contacts. Upon increase of the concentration of p-LES, the free energy of the state D^* decreases and competes with the state N.

The 'hot' sites, which stabilize the LES, are found to be very sensitive to (nonconservative) point mutations [57]. Since most of the protein stabilization energy is concentrated in these sites, mutating one or two of them has a high probability of denaturing the native state.[10] On the other hand, mutating any other site has in general little effect on the stability of the protein [59]. With the increased stability of the LES one can imagine a more hierarchical folding similar to the diffusion–collision mechanism. However, if the LES are of different stability the nucleation–condensation might be a more likely mechanism [30].

Folding inhibition

Intervening a folding reaction can, in principle, be achieved by interacting the polypeptide chain with peptides whose sequences are those of the LES that define the (postcritical) FN of the studied protein. As the concentration of the LES peptides increases, the protein may nucleate by the assembly of the protein chain with peptide LES, leading to a nonproductive folding. This can be viewed as changing the folding from a unimolecular reaction to a bimolecular reaction (cf. Figure 1c). Making use of lattice models, it has been shown that it is possible to destabilize the native conformation of a protein making use of peptides whose sequences are identical to that of the LES [60]. Such peptides (p-LES) interact with the protein (in particular with their complementary LES) with the same energy that stabilizes the (postcritical) FN. If the concentration of p-LES is large enough, the loss of entropy upon binding is smaller than that associated with folding, and thus the free energy of the partially unfolded state D^* (see Figure 1c) is stabilized (see Figure 2), thus inhibiting folding (within this context, see also [54••,61]).

There are two important advantages of these folding inhibitors with respect to conventional ones. First, their molecular structure is suggested directly by the target protein. One needs not to design or optimize anything, just find the LES of the protein to be inhibited, because the design has been performed by evolution through a myriad of generations of the organism that expresses the protein. Moreover, it is unlikely that the protein can develop resistance through mutations. In fact, a folding inhibitor binds to a LES, and a protein cannot mutate the amino acids of a LES [57,59] – in any case not those 'hot' amino acids that are essential to stabilize it as well as to bind to the other LES to form the (postcritical) FN - under risk of denaturation. This does not mean that the protein cannot decrease its affinity to p-LES. It only means that non-conventional resistance is much less likely than conventional one (see footnote 10).

Folding mechanism of the HIV-1 protease monomer

Molecular dynamics, all-atom simulations with explicit treatment of the solvent [24] have provided clear indication of the stability of the unbound folded monomer, but

[10] Because there exist quasi-degeneracies in the interaction between amino acids, in particular in the case of hot amino acids, different sequences folding to the same native conformation may do so with the help of different sets of hot sites and thus of LES and of (postcritical) FN, as shown by the existence of families of analogous sequences. Evolutionary model studies [58] have shown that proteins cannot change their FN through point mutations – as these mutations will eventually lead to non-folding sequences – but through multiple co-ordinated simultaneous mutations.

are too computationally demanding for studying folding. For this purpose, use can be made of the Gō model (see footnote 4). Results of calculations obtained making use of this model indicate that the group of amino acids 27–35 and 79–87 display relative high ϕ-values (i.e. close to 1), and are thus essential in the folding of the monomer [16••].

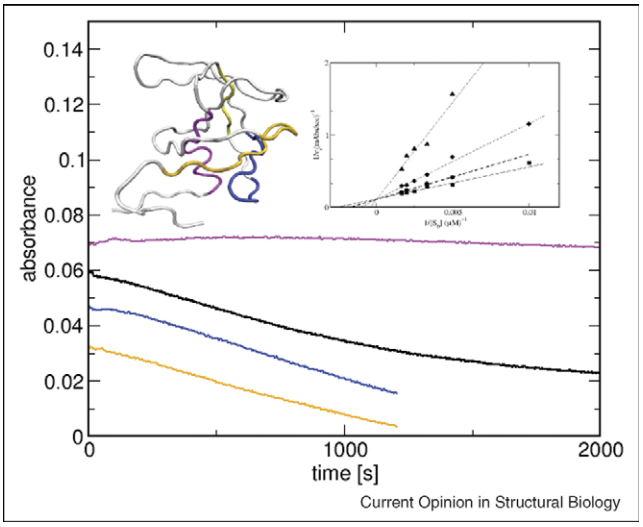
A Gō model which accounts for some of the diversities existing among the different amino acids can be developed by assigning to each native pair an interaction energy obtained by averaging the force field associated with the all-atom simulations around the native conformation [17••] (see also [62]). These simulations indicate that the LES of the monomer comprise residues 23–33, 74–78 and 83–92. The (postcritical) FN forms a spatially closed unit of a helix (83–92) with sheet (74–78) above and another beta-strand with sheet (24–34) perpendicular to these elements[11] (see Figure 1a). The LES and the associated (postcritical) FN introduced above essentially coincide with the highly protected structural units and the stabilization core defined in [63••]. Aside from these results, there exist a number of circumstantial evidences which testify to the central role played by the groups of amino acids belonging to the LES and the fact that sites 33, 75, 76, 85 and 89 can be identified as 'hot' sites, all contained in the LES (see Figure 1a and ref. [64]). It is then possible to identify the peptide displaying the same sequence as the segment 83–92 of the wildtype protease as the best candidate to play the role of folding inhibitor (p-LES (83–92)). This choice was primarily made on the basis of the fact that the other two LES are either too short (74–78) to be specific or too little structured (23–33). This last feature has been studied in detail with the help of all-atom explicit-solvent molecular dynamics simulations, which in the case of peptides containing about 10 amino acids are within the possibilities of present day computing facilities provided refined techniques like metadynamics, parallel and solute tempering, etc. are used. The same computational methods have been applied to the study of the interaction of the two peptides. Their docking results from the wrapping of a strongly fluctuating p-LES (23–33) around a well-structured p-LES (83–92), and is stabilized by three particular hydrogen bonds [65].

Folding inhibition of HIV-1 Protease: experimental supports

Gō model simulations of the folding of the HIV-1 protease monomer in the presence of three copies of p-LES (83–92), corresponding to concentration of 5 mM, indicate that the native state equilibrium population is reduced by as much as a factor of 4 as compared to the result obtained in the case of the monomer by itself [17••].

[11] In hindsight, similar results could have been obtained making use of the homogeneous Gō model [16••].

Figure 3



The enzymatic kinetics of the protease alone (black curve), of the protease inhibited with p-LES 83–92 (magenta curve) and of the protease together with control peptides 61–70 (blue curve) and 9–19 (orange curve) not related to LES. In the insets we show the segments of the monomer (with the same colors) and the Lineweaver–Burk plot associated with the protease (filled squares) and the protease complexed with the p-LES 83–93 at 3 μM (circles), 10 μM (diamonds) and 20 μM (squares).

Making use of spectrophotometric assays (Figure 3) it has been measured that the inhibition constant associated with the peptide p-LES (83–92) is $K_I = 7.72 \pm 2.34$ μM [66]. A strong decrease of the beta-sheet content between these two cases was observed by circular dichroism (from 30% to 14%), indicating that the protein is, to a large extent, in a non-folded conformation.[12] In infected cells, it was found that the peptide was able to readily cross the cell membrane and was not toxic to peripheral blood mononuclear cells [68]. Furthermore, it showed an excellent therapeutic/toxic ratio, with antiviral levels well below toxic concentration. Its IC50 (50% inhibitory concentration) was in the μM range (consistent with the results of measurements carried out on the enzyme). Even more promising, p-LES 83–92 inhibited a multidrug resistant HIV isolate, and was found to be more effective than a conventional (active-site centered) drug (Atazanavir) in a chronically infected cell line. This difference is likely to be connected with the fact that misfolded proteins within mammalian cells are usually conveyed to the proteosome after ubiquitination leading to their selective degradation, thus transforming a reversible into a no-return process.

[12] Modified Gō model dynamic simulations [17••] of the folding of the HIV-1-PR dimer indicate that p-LES (83–92) may also be able to prevent, to some extent, dimerization. A systematic study of this possibility has been carried out by means of calculations of the interaction of p-LES (83–92) with the ensemble of folded and partially folded conformations of the dimer [67].

Folding inhibition: sequence-based method

From the above discussion, it emerges that the main task to be carried out in designing a non-conventional inhibitor of a chosen target protein is that of finding the LES of the protein. This can be done, for example, as in the case of HIV-1-PR, through a careful study of its folding mechanism. In the case when the native structure of the target protein is not available and therefore the LES cannot be easily identified by simulations, one can scan all segments of the protein sequence for eventual LES [69]. In fact, whatever the size of a protein is, the associated LES must be short, of the order of 5–15 residues long, so as to be able to become structured in the early stages of the folding process. Consequently, to individuate the LES of a protein one has to test the inhibitory properties of a set of peptides of length 5–15. Each of these peptides has a sequence identical to a segment of the target protein, and displays a consistent overlap (20–50%) with the neighboring peptides. The full set of peptides covers the entire protein with a consistent amount of redundancy. The (few) peptides that mostly destabilize the protein, inhibiting its activity, are likely to be the p-LES. This protocol can be implemented equally well computationally or experimentally (by means of enzymatic or structural assays). Typically, the number of peptides to be tested is of the order of few tens. This is indeed a much smaller effort than that used in pharmaceutical high-throughput screening, where hundreds of thousand molecules are usually tested.

The general method has been applied *in silico* and *in vitro* to hen egg lysozyme [70]. Spectroscophotometric assays showed that few of the peptides tested, each composed of 10 amino acids, are able to block the enzymatic activity of the protein with μM efficiency. NMR, circular dichroism and fluorescence measurements indicate that this inhibition is because of the misfolding of the lysozyme.

Caveats

There are a number of questions that leap to mind from the above narrative. Some of the points include: (a) How specific is the p-LES effect with regards to other homologous proteins? (e.g. in the case of the HIV-1 protease, to the human aspartic proteases [71]), (b) p-LES can be marginally soluble since they are often highly hydrophobic: how easy is it to be delivered to the cell, achieving therapeutic levels?, (c) Once delivered, how prone are p-LES to hydrolization, (d) mutations will eventually occur even in the most stable sequences. How would one deal with the eventual insurgence of (non-conventional) resistance? It is likely that one can find some of the answers to these questions making use of molecules mimetic to the p-LES, each representative of one of the few different FN of the target protein.

Conclusions

Aside from the potential interest concerning the design of drugs which do not create resistance, the study of folding inhibitors will also shed light on the mechanism which is at the basis of the folding of proteins. Concerning the first point, there are still a number of questions in need of answer. Only if these answers are found, could the p-LES strategy of protein inhibition eventually display all its potential clinical interest.

Acknowledgements

This work was supported in part by a grant from the Clore Center for Biological Physics, the Kimmelman Center for Macromolecular Assemblies, and the Center for Complexity Science (Y.L.). GT and RAB acknowledge the financial support of the 2003 FIRB grant of the Italian Ministry for Scientific Research and of initiative TO61 of INFN.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Xu D, Tsai CJ, Nussinov R: **Mechanism and evolution of protein dimerization**. *Prot Sci* 1998, **7**:533-544.

2. Tiana G, Broglia RA: **Design and folding of dimeric proteins**. *Proteins* 2002, **49**:82-94.

3. Papoian GA, Wolynes PG: **The physics and bioinformatics of binding and folding-an energy landscape perspective**. *Biopolymers* 2003, **68**:333-349.

4. Caflisch A: **Folding for binding or binding for folding?** *Trends Biotechnol* 2003, **21**:423-425.

5. Levy Y, Wolynes PG, Onuchic JN: **Protein topolgy determines binding mechanism**. *Proc Natl Acad Sci U S A* 2004, **101**:511-516.

6. Levy Y, Cho SS, Onuchic JN, Wolynes PG: **A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes**. *J Mol Biol* 2005, **346**:1121-1145.

7. Fitzgerald PMD, Springer JP: **Structure and function of retroviral proteases**. *Annu Rev Biophys Chem* 1991, **20**:299-320.

8. Wlodawer A, Vondrasek J: **Inhibitors of HIV-1 protease: a major success of structure-assisted drug design**. *Annu Rev Biophys Struct* 1998, **27**:249-284.

9. Freire E: **Designing drugs against heterogeneous targets**. *Nat Biotechnol* 2002, **20**:15-16.

10. Tomasselli AG, Heinrikson RL: **Targetting the HIV protease in AIDS therapy: a current clinical perspective**. *Biochem. Biophys. Acta* 2000, **1477**:189-214.

11. Cohen J: **Confronting the limits of success**. *Science* 2002, **296**:2320-2324.

12. Zutshi R, Franciskovich J, Schultz M, Schweitzer B, Bishop P, Wilson M, Chemielewski J: **Targetting the dimerization interface of HIV-1 protease: inhibition with cross-linked interfacial peptides.**. *J Am Chem Soc* 1997, **119**:4841-4845.

13. Bowman MJ, Chemielewski J: **Novel strategies for targetting the dimerization interface of HIV protease with cross-linked interfacial peptides**. *Biopolymers* 2002, **66**:126-133.

14. Boggetto N, Reboud-Ravaux M: **Dimerization inhibition of HIV-1 protease**. *Biol Chem* 2002, **383**:1321-1324.

15. Caflisch A, Schramm HJ, Karplus M: **Design of dimerization inhibitors of HIV-1 aspartic proteinase: a computer-based combinatorial approach**. *J Comput Aided Mol Des* 2000, **14**:161-179.

16. •• 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.

Analysis of the folding of the HIV-1-PR. It presents compelling evidence that monomer folding and dimerization are decoupled. In view of the relatively high ϕ–values obtained for residues 23–35 and 79–87 in both the folding and binding transition states together with their proximity at the interface, the authors conclude that these segments are good targets for inhibitor design.

17. •• Broglia RA, Tiana G, Sutto L, Provasi D, Simona F: **Design of HIV-1-PR inhibitors that do not create resistance: blocking the folding of single monomers**. *Prot Sci* 2005, **14**:2668-2681.
On the basis of model results of ref. [60] concerning folding inhibitors and making use of a generalized Gō model, the LES of the HIV-1-PR are found. It is then shown that the corresponding p-LES peptides act as effective inhibitors of the folding of the protease while peptides of similar length and displaying a sequence identical to segments of the enzyme not associated with LES do not have any influence in the folding of the enzyme. It is predicted that the segment 83–93 should inhibit folding at μM level. Furthermore, it is suggested that the p-LES strategy can be used to inhibit other target proteins. This program is partially developed in refs. [69] and [70], while the prediction was verified in ref. [66] in the enzyme and in ref. [68] in the virus.

18. Perryman AL, Lin JH, Mccammon JA: **HIV-1 protease molecular dynamics of a wildtype and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs**. *Prot Sci* 2004, **13**:1108-1123.

19. Piana S, Carloni P, Parrinello M: **Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease**. *J Mol Biol* 2002, **319**:567-583.

20. Piana S, Carloni P, Rothlisberger U: **Drug resistance of HIV-1 protease: flexibility-assisted mechanism of compensatory mutations**. *Prot Sci* 2002, **11**:2393-2492.

21. Xie D, Gulnik S, Gustchina E, Yu B, Shao W, Qoroneh W, Nathan A, Erickson JW: **Drug resistance mutations can affect dimer stability of HIV-1 protease at neutral pH**. *Prot Sci* 1999, **8**:1702-1713.

22. Ishima R, Ghirlando R, Todzser J, Gronenborn AM, Torchia DA, Louis JM: **Folded monomer of HIV-1 protease**. *J Biol Chem* 2001, **276**:49110-49116.

23. Louis JM, Ishima R, Nesheiwat I, Pannell LK, Lynch SM, Torchia DA, Gronenborn AM: **Revisiting monomeric HIV-1 protease**. *J Biol Chem* 2003, **278**:6085-6092.

24. Levy Y, Caflisch A: **The flexibility of monomeric and dimeric HIV-1-PR**. *J Phys Chem* 2003, **107**:3068-3079.

25. Gō N: **Theory of reversible denaturation of globular proteins**. *Int J Peptide Prot Res* 1975, **7**:313-323.

26. Clementi C, Nymeyer H, Onuchic JN: **Topological and energetical factors: What determines the structural details of the transition state ensemble and 'En-route' intermediate for protein folding? An investigation of small globular proteins**. *J Mol Biol* 2000, **298**:937-947.

27. Dill KA, Shortle D: **Denatured states of proteins**. *Annu Rev Biochem* 1991, **60**:795-825.

28. Shortle D: **The denatured state (the other half of the folding equation) and its role in protein stability**. *FASEB J* 1996, **10**:27-34.

29. Bowler BE: **Thermodynamics of protein denatured states**. *Mol Biosyst* 2007, **3**:88-99.

30. Fersht AR: **Structure and Mechanism in Protein Science**. New York: W.H. Freeman and Co.; 1999.

31. Oliveberg M, Wolynes PG: **The experimental survey of protein-folding energy landscapes**. *Quart Rev Biophys* 2005, **38**:245-288.

32. Garci-Mira MM, Sadqi M, Fisher N, Sanchez-Ruiz JM, Munoz V: **Experimental identification of downhill protein folding**. *Science* 2002, **298**:2191-2195.

33. Onuchic JN, Socci ND, Luthey-Schulten Z, Wolynes PG: **Protein folding funnels: the nature of the transition state ensemble**. *Fold Des* 1996, **1**:441-450.

34. Onuchic JN, Wolynes PG: **Theory of protein folding**. *Curr Opin Struct Biol* 2004, **14**:70-75.

35. Shakhnovich EI, Gutin AM: **A new approach to the design of stable proteins**. *Prot Eng* 1993, **6**:793-800.

36. Shakhnovich EI, Gutin AM: **Engineering of stable and fast-folding sequences of model proteins**. *Proc Natl Acad Sci U S A* 1993, **90**:7195-7199.

37. Grantcharova V, Riddle DS, Santiago JV, Baker D: **Important role of hydrogen bonds in the structurally polarized transition state for folding of the src SH3 domain**. *Nat Struct Biol* 1998, **5**:714-720.

38. Martinez JC, Pisabarro MT, Serrano L: **Obligatory steps in protein folding and the conformational diversity of the transition state**. *Nat Struct Biol* 1998, **5**:721-729.

39. Khan F, Chuang JI, Gianni S, Fersht AR: **The kinetic pathway of folding barnase**. *J Mol Biol* 2003, **333**:169-186.

40. •• Fersht AR: **Characterizing transition states in protein folding: an essential step in the puzzle**. *Curr Opin Struct Biol* 1994, **5**:79-84.
The experimental approach to characterize the structural features of the critical folding nucleus through the analysis of ϕ-values is discussed in this opinion (see also [30,37–39]).

41. Shakhnovich EI, Abkhevich VI, Ptitsyn O: **Conserved residues and the mechanism of protein folding**. *Nature* 1996, **379**:96-98.

42. Itzhaki L, Otzen D, Fersht AR: **The structure of the transition state for folding of chymotrypsin inhibitor 2 analyzed by protein engineering methods**. *J Mol Biol* 1995, **254**:260-288.

43. •• Abkevich VI, Gutin AM, Shakhnovich EI: **Specific nucleus as the transition state for protein folding**. *Biochemistry* 1994, **33**:10026-10032.
The presence of specific native interactions which commit the protein to fold, and thus define the postcritical folding nucleus, are shown within a simplified lattice model.

44. Yi Q, Scalley-Kim ML, Alm EJ, Baker D: **NMR characterization of residual structure in the denatured state of protein L**. *J Mol Biol* 2000, **299**:1341-1351.

45. Mok YK, Kay CM, Kay LE, Forman-Kay J: **NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions**. *J Mol Biol* 1999, **289**:619-638.

46. Zhang O, Forman-Kay JD: **Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer**. *Biochemistry* 1995, **34**:6784-6794.

47. Navon A, Ittah V, Landsman P, Scheraga HA, Haas E: **Distributions of intramolecular distances in the reduced and denatured states of bovine pancreatic ribonuclease A. Folding initiation structures in the C-terminal portions of the reduced proteins**. *Biochemistry* 2001, **40**:105-118.

48. Broglia RA, Tiana G: **Hierarchy of events in the folding of model proteins**. *J Chem Phys* 2001, **114**:7267-7273.

49. Tiana G, Broglia RA: **Statistical analysis of native contact formation in the folding of designed model proteins**. *J Chem Phys* 2001, **114**:2503-2507.

50. Lesk AM, Rose GD: **Folding units in globular proteins**. *Proc Natl Acad Sci U S A* 1999, **78**:4304-4308.

51. Ptitsyn OB, Rashin AA: **A model of myoglobin self-organization**. *Biophys Chem* 1975, **3**:1-20.

52. Baldwin RL, Rose GD: **Is protein folding hierarchic?** *TIBS* 1999, **24**:26-83.

53. Hansen A, Jensen MH, Sneppen K, Zocchi G: **A hierarchical scheme for co-operativity and folding of proteins**. *Physica A* 1998, **250**:335-351.

54. •• Pincus MR: **Identification of structured peptide segments in protein folding**. *Bioploymers* 1992, **32**:347-351.
This paper presents clear evidence of the fact that proteins contain sequences that tend to adopt native-like conformations by themselves in solution. It also posits that peptides with that sequences will inhibit folding if the complementary segment is involved in folding on early stage of this process. It also reports inhibition of RNase in terms of peptides displaying the same sequence as the segment 105–123 (concentration 10 μM).

55. Panchenko AR, Luthey-Schulten Z, Wolynes PG: **Foldons, protein structural modules and exons**. *Proc Natl Acad Sci U S A* 1995, **93**:2008-2013.

56. Hansen A, Jensen MH, Sneppen K, Zocchi G: **A hierarchical scheme for co-operativity and folding in proteins**. *Physica A* 1998:355-361.

57. Tiana G, Broglia RA, Roman HE, Vigezzi E, Shakhnovich EI: **Folding and misfolding of designed protein-like chains with mutations**. *J Chem Phys* 1998, **108**:757-761.

58. Tiana G, Broglia RA, Shakhnovich EI: **Hiking in the energy landscape in sequence space: A bumpy road to good folders**. *Proteins* 2000, **39**:244-251.

59. Broglia RA, Tiana G, Pasquali S, Roman HE, Vigezzi E: **Folding and aggregation of designed protein chains**. *Proc Natl Acad Sci U S A* 1998, **95**:12930-12934.

60. Broglia RA, Tiana G, Berera R: **Resistance-proof, folding-inhibitor drugs**. *J Chem Phys* 2003, **118**:4754-4758.

61. Chie L, Chung D, Pincus MR: **Specificity of inhibition of ras-p21 signal transduction by peptides from GTPase activation protein (GAP) and the son-of sevenless (SOS) ras-specific guanine nucleotide exchange protein**. *Prot J* 2005, **24**:253-258.

62. Karanicolas K, Brooks CB: **The origins of asymmetry in the folding transition states of protein L and protein G**. *Prot Sci* 2002, **11**:2351-2361.

63. •• Wallqvist A, Smythers GW, Covell DG: **A cooperative folding unit in HIV-1 protease. Implications for protein stability and occurrence of drug-induced mutations**. *Prot Eng* 1998, **11**:999-1005.

The authors determine the structural units of the HIV-1-PR that exhibit a stronger protection against unfolding relative to other parts of the enzyme. This procedure identifies a stabilization core in the protease, on the basis of an ensemble of denatured states derived from native structures, comprising a close unit of residues 84–91, 74–78 and 22–32. These regions also display high sequence conservation.

64. Broglia RA, Tiana G, Sutto L, Provasi D, Simona F: **The physics of protein folding and of non-conventional drug design: Attacking AIDS with its own weapons**. *La Rivista del Nuovo Cimento* 2007, **29**:1-64.

65. Bonomi M, Gervasio FL, Tiana G, Provasi D, Broglia RA, Parrinello M: **Insight in the folding inhibition of the HIV-1 Protease by a small peptide**. *Biophys J* 2007, **93**:8-13.

66. Broglia RA, Provasi D, Vasile F, Ottolina G, Longhi R, Tiana G: **A folding inhibitor of the HIV-1 protease**. *Proteins* 2006, **62**:928-933.

67. Verkhivker GM, Tiana G, Camilloni C, Provasi D, Broglia RA: **Analysis of the binding mechanism of a non-conventional HIV-1-protease inhibitor: Monte Carlo simulations of folding inhibition with the ensembles of folded and partially unfolded protease conformations**, in press.

68. Rusconi S, Lo Cicero M, Laface AE, Ferramosca S, Cesana E, Tiana G, Provasi D, Longhi R, Sirianni F, Nielsen C, *et al.*: **Design and biological activity of a non-conventional (folding) inhibitor of HIV-1 Protease**, in press.

69. Broglia RA, Tiana G, Sutto L, Provasi D, Perelli V: **Low-throughput model design of protein folding inhibitors**. *Proteins* 2007, **67**:469-478.

70. Caldarini M, Vasile F, Provasi D, Longhi R, Tiana G, Broglia RA: **Identification and characterization of folding inhibitors of hen egg lysozyme**, in press.

71. Roberts NA, Martin JA, Kinchington D, Broadhurst AV, Craig JC, Duncan IB, Galpin SA, Handa BK, Kay J, Krohn A *et al.*: **Rational design of peptide-based HIV proteinases inhibitors**. *Science* 1990, **248**:348-361.