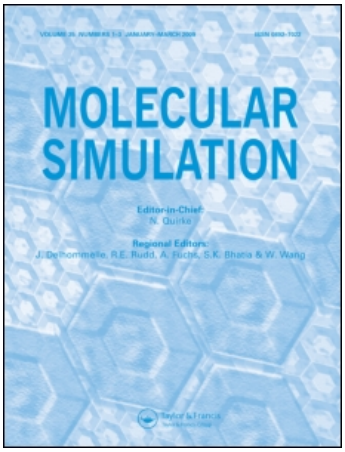

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Design of a folding inhibitor of the HIV-1 protease

G. Tiana[a]; R. A. Broglia[b]; L. Sutto[a]; D. Provasi[a]
[a] Dipartimento di Fisica, Universitá di Milano, and INFN, Milano, Italy [b] Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

To cite this Article Tiana, G. , Broglia, R. A. , Sutto, L. and Provasi, D.(2005) 'Design of a folding inhibitor of the HIV-1 protease', Molecular Simulation, 31: 11, 765 — 771
To link to this Article: DOI: 10.1080/089270500262721
URL: http://dx.doi.org/10.1080/089270500262721

Molecular Simulation, Vol. 31, No. 11, 15 September 2005, 765–771

Taylor & Francis
Taylor & Francis Group

Design of a folding inhibitor of the HIV-1 protease

G. TIANA†*, R. A. BROGLIA†‡, L. SUTTO† and D. PROVASI†

†Dipartimento di Fisica, Universitá di Milano, and INFN, via Celoria 16, 20133 Milano, Italy
‡Niels Bohr Institute, University of Copenhagen, Bledgamsvej 17, 2100 Copenhagen, Denmark

(Received June 2005; in final form July 2005)

A novel way to inhibit HIV-1 protease by destabilizing its native state is discussed. A simplified protein model is used together with Monte Carlo simulations, to assess the destabilizing effect of peptides displaying the same sequence as specific fragments of the protein which are essential for its stability. Model calculations also show that it is unlikely that the protein can escape the inhibitory peptide by point mutations.

Keywords: HIV-1 protease; Monte Carlo simulations; Local elementary structures; HIV-PR monomer

1. Introduction

Monte Carlo and molecular dynamics techniques are able, directly or by virtue of the ergodic theorem, to calculate the free energy of a system given its potential function, without the need of employing *ad hoc* effective free energy parametrizations. Their use in the design of protein inhibitors is however limited by the enormous computational cost of equilibrating even a complex protein–ligand composed by few atoms. In fact, the typical goal is to screen a large number of potential ligands and evaluate their binding free energies to a specific site of the protein (e.g. the active site of an enzyme), in order to find the best inhibitor. Consequently, it has been employed only in association with faster searching algorithms [1] or to refine results obtained with calculation based on empirical scoring functions [2].

A different strategy for the inhibition of an enzyme is to interfere with its folding, destabilizing the native state of the protein [5,6]. This is done exploiting the fact that the stabilization energy of a protein is not distributed evenly among the amino acids, but is concentrated in a core of mutually interacting residues [3,4,7]. The fragments of the protein which carry these "hot" residues (called local elementary structures (LES) in [8]) play a key role also for the folding kinetics [9,10]. The basic idea is to destabilize a protein in solution by adding short peptides with the same sequence as a LES of the protein [5,6]. We shall call these peptides p-LES. A p-LES compete with the formation of the stabilization core of the protein, shifting the equilibrium towards the unfolded state, thus inhibiting its biological activity.

In order to study the destabilization effect of the p-LES, one needs a way to calculate the equilibrium properties of the protein alone and of a solution of protein and p-LES. This is a problem computationally harder even than that of folding a protein. Consequently, one has no other way than resorting to simplified models.

The use of lattice models has proven useful to develop the idea of using p-LES as inhibitors [5], but is of no help in locating the LES and the stabilization core of a specific protein and in studying the destabilizing effect of p-LES. For these tasks one needs a model which balances computational efficiency with the ability to account for the main ingredients which make the p-LES effective. Unified-atom Gō models [11] describe each amino acid as a sphere centered at the position of its C_α and interacting with the other amino acids through a well-potential if the pair is in contact in the crystallographic native state. These models are computationally easy to handle, describe decently the conformational space and the native conformation of the protein, and have by definition, the native state as ground state of the protein. On the other hand, all the native pairs of amino acid interact with the same energy, and thus the stabilization core is unrealistically defined only by the topology of the native state, and not by the sequence of amino acids, each displaying different chemical properties.

Consequently, we have modified the Gō model, assigning to each pair of amino acid a different energy, taken from a $N \times N$ matrix B_{ij}, N being the number of residues of the chain. The matrix is calculated by means of all-atom simulations around the native conformation with semiempirical potentials (in the present case, the GROMOS96 force field [12]). In other words, the potential function which

*Corresponding author. E-mail: Guido.Tiana@mi.infn.it

Molecular Simulation
ISSN 0892-7022 print/ISSN 1029-0435 online © 2005 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/089270500262721

controls the protein is

$$U(\{r_i\}) = \sum_{i+2<j} \Big[B_{ij}\theta(R - |r_i - r_j|)\theta(R - |r_i^N - r_j^N|) + \epsilon_{HC}\theta(0.99 \cdot |r_i^N - r_j^N| - |r_i - r_j|) \Big] \times \theta(R - |r_i^N - r_j^N|) + \epsilon_{HC}\theta(R - |r_i - r_j|) \times [1 - \theta(R - |r_i^N - r_j^N|)] \quad (1)$$

where r_i is the coordinate of the C_α atom of the i-th amino acid, r_i^N is the coordinate in the crystallographic native conformation, $\theta(x)$ is a Heaviside step function, B_{ij} is the interaction energy between i-th and j-th amino acid, R the interaction range (which in the following calculations is set equal to 7.5 Å) and ϵ_{HC} is the hard core repulsion, set to 100 kT. Accordingly, the first term of the potential function describes the attraction between native pairs, the second term the hard core between native pairs, its range being equal to the 99% of the native distance, and the last term describes the repulsion between non-native pairs. Moreover, we assume that residue i interacts with residue $i + 2$ only through a hard core repulsion of range $R/2$. The matrix B_{ij} is the average interaction between all atoms of residues i and j, calculated with Gromacs in a 1 ns simulation at 300 K, during which the overall RMSD does not exceed 2.5 Å. The average of the interaction matrix is -4.4 kJ/mol, while its standard deviation is 3.2 kJ/mol. This model has proven useful in reproducing a number of site-specific properties of small globular proteins [13].

In evaluating the interactions within a p-LES and between a p-LES and the protein, use is made of the same matrix elements which act between the corresponding amino acids within the protein, in keeping with the fact that a p-LES is chemically identical to the corresponding LES. For example, if residues 17 and 26 of the protein are in contact in the native conformation, the matrix element $B_{17,26}$ is different from zero, and the interaction energy between the residues 17′ and 26′ belonging to the p-LES or between the residue 17′ of the p-LES and 26 of the protein is again $B_{17,26}$.

In the following, we will deal with the inhibition of HIV-1 protease, a protein involved in the replication of the HIV-1 virus, and thus one of the main targets of anti-viral drugs. There exists five approved drugs which bind to the active site of the protease, making it inactive. The problem is that the viral protein is able to mutate its sequence at a very high rate, causing an accelerated evolution under the selective pressure of the drug. In 6–8 months such an evolution is able to select those proteins which display low affinity for the drug, and the virus becomes resistant [14].

The HIV-1 protease is a dimer, composed of two identical monomers of 99 amino acids each (see figure 1). At neutral pH and biological temperature, the biologically-active dimer is at equilibrium with the folded monomers [15]. Consequently, destabilizing the monomer is sufficient to shift the equilibrium of the system towards biologically-inactive states.

We will first determine the "hot" sites of the protein by energy considerations, and then test a number of fragments containing those "hot" sites as inhibitor, by means of Monte Carlo samplings. A number of fragments, and in particular, that corresponding to residues 83–93 of the HIV-Pr, are shown to destabilize consistently the native state of the protein. Moreover, we show that mutations in the sequence of HIV-Pr do not affect the destabilizing efficiency of the inhibitor.

2. Identifying "hot" sites

In principle, the localization of the stabilization energy among the different residues can be investigated with the help of the total interaction energy $E_i \equiv \sum_j B_{ij}$ of each residue (see figure 2, dashed curve). But a residue belonging to the core can build other weak contacts, while a residue outside the core can display a low E_i by summing a number of weak contributions. Consequently, from

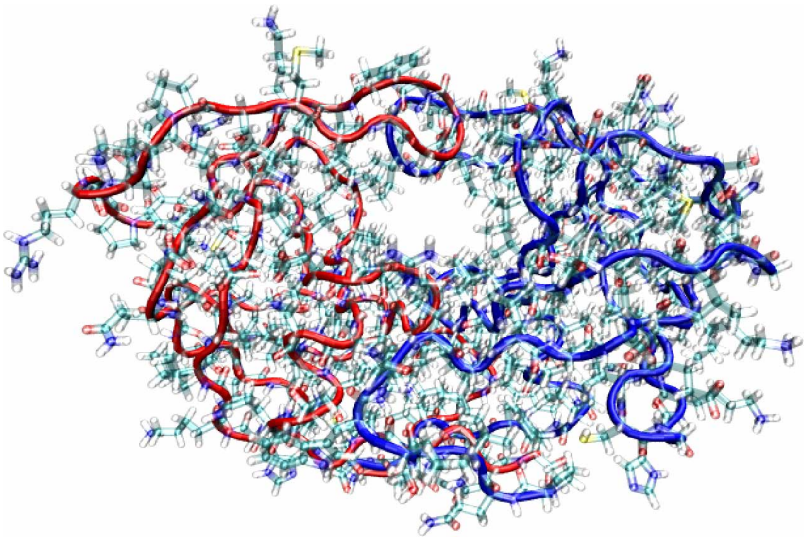

Figure 1. The crystallographic native conformation of HIV-1 protease.

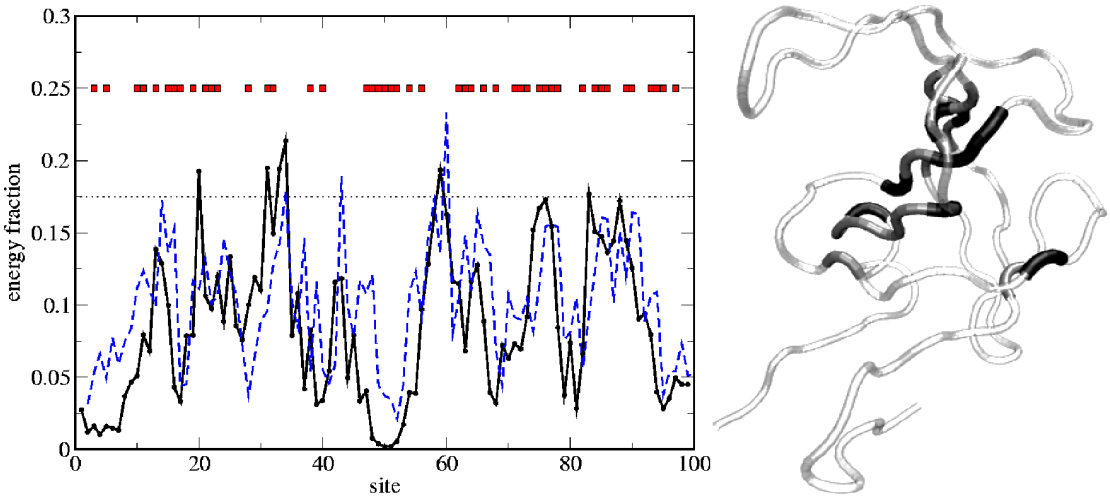

Figure 2. The principal eigenvalue μ_i^0 of the interaction matrix B_{ij} (solid curve) and the total interaction per site E_i (dashed curve). The squares on the top of the plot indicate the most hydrophobic residues (A, G, L, I, V), while the dotted curve identifies the atypical values (see text). To the right is the cartoon of the HIV-Pr monomer, where the sites are shadowed with different levels of gray according to the value of the associated eigenvector.

the distribution of E_i (see figure 3(b), solid curve), it is difficult to isolate the core residues, the tail of the distribution not departing consistently from the Gaussian distribution that one would expect from the central limit theorem if the interaction energies were uncorrelated (figure 3(b), dashed curve).

A more effective strategy is to decompose the interaction energy B_{ij} in eigenvalues, that is, $B_{ij} = \sum_k \lambda^k \mu_i^k \mu_j^k$. If the lowest eigenvalue λ^0 is much lower than the others, then the stabilization energy is concentrated in a core of few mutually interacting residues [13]. In fact, one can approximate

$$B_{ij} = \lambda^0 \mu_i^0 \mu_j^0, \tag{2}$$

where the value of μ_i^0 indicates the extent to which the i-th residue contribute to the stabilization core. The values of μ_i^0 for the 99 sites of the HIV protease are displayed in figure 2 (solid curve) and their distribution is displayed in figure 3(a) (solid curve). If the interaction matrix were random, the distribution of the components μ_i^0 should be a Gaussian with both mean and standard deviation equal to $1/99 = 0.01$ [16]. Such a distribution is shown in figure 3(a) with a dashed curve, and the high-energy tail of the distribution of μ_i^0 which differs consistently from the random distribution ($\mu_i^0 > 0.175$) is marked with filled circles. Thus, we can define preliminarily as "hot" those sites displaying $\mu_i^0 > 0.175$, that is sites 20, 31, 33, 34, 59, 76, 83 and 88. However, note that a number of approximations affect such a definition: (1) a number of sites lie close to the borderline, making it difficult to assign precisely the label "hot" (for this purpose, in [7] are introduced "warm" sites, although it is not really necessary); (2) eigenvectors other than the principal can weaken the residues identified by the above method, or strengthen those which lie closely; (3) the matrix B_{ij} does not contain solvation free energies, which are not two-body interactions, and consequently, this information has to be added separately (cf. filled squares in figure 2).

Another tool to identify the sites which are essential for the stability of the protein is to look for those, which are most conserved in evolution. There are two kinds of evolution which can be analyzed, namely the long-term evolution of the native conformation of figure 1, common to several aspartyl-proteases, and the fast evolution of the HIV-Pr in infected cells. Given a family of proteins structurally similar, but not homologous, to HIV-Pr, one can define the entropy per site $S(i) = -\sum_\sigma p_i(\sigma) \log p_i(\sigma)$, where $p_i(\sigma)$ is the probability of finding the amino acid of kind σ in the i-th site of the protein. The entropy per site assumes the value 0 if that site is perfectly conserved and $\log 20 \approx 3$ in the opposite case. Taking, for example, the value 1.5 as threshold to distinguish between conserved and non-conserved sites, one can identify two regions of the protein displaying conserved

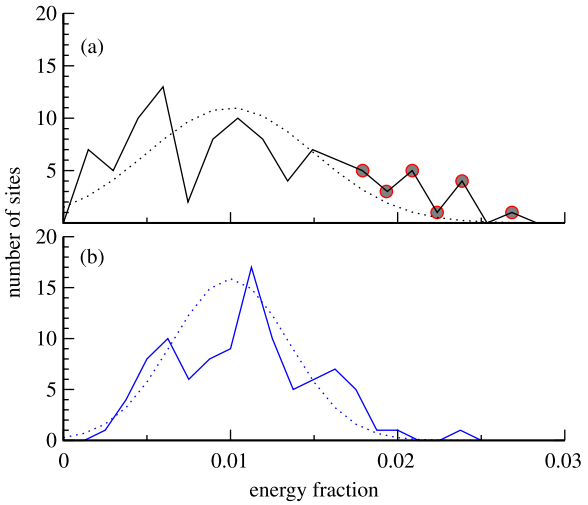

Figure 3. (a) The distribution of components of the principal eigenvector μ_i^0 (solid curve) and that associated with the eigenvectors of a random symmetric matrix (dashed curve). The shadowed circles indicate those components of μ_i^0 which are consistently more frequent than the random approximation. (b) The distribution of interaction energies per site E_i (solid curve) and its Gaussian fit (dashed curve).

residues, spanning residues 21–35 and 79–91 [6]. Note that the high conservation of site 25, 26 and 27 is expected, since they constitute the active site of the protease. The effects of the fast evolution can be analyzed by searching for those sites which have been observed to undergo non-conservative mutations (i.e. substitutions with amino acids displaying different chemical properties) in infected cells under the selective pressure of inhibitory drugs [17]. Mutable sites are 10, 12, 16, 17, 19, 20, 34, 35, 37, 39, 41, 43, 45, 54, 55, 63, 67, 69–72, 74, 81, 82, 95. The two widest regions which are not observed to mutate are essentially the same which are conserved in the long-term evolution.

The combination of these two sets of data suggests that the two fragments of the protein which, being highly energetic and consistently conserved, are crucial for the stability of the native state of HIV-Pr are 24–34 and 83–93 (with some degree of uncertainty on the exact starting and ending points). The interactions within and between these fragments have been shown by molecular dynamics simulations to be important in driving the folding of HIV-Pr, thus qualifying as LES of the protein [6]. Also the study of other models [18–20] which describe the dynamics of HIV-Pr support the importance of these fragments both for the stability and the kinetics of the protease.

3. Thermodynamical simulations of the HIV-Pr monomer

Using the simplified model described above, we have first characterized the thermodynamics of the HIV-Pr monomer, obtaining in particular, the fractional population p_N of the native state at $T = 2.5$ kJ/mol, and then investigated how the presence of peptides displaying the same sequence as fragments of the HIV-Pr destabilize the protein, decreasing p_N.

Ten Monte Carlo simulations of the monomer, lasting for 10^{10} steps each, have been performed with the Metropolis algorithm at temperatures ranging 2–5 kJ/mol. A move consists in flipping a randomly-chosen monomer of a random angle around the axis defined by the previous and the next monomer. The energy distribution obtained by the simulations have been analyzed in terms of a multihistogram algorithm [21] in order to obtain the density of states, and from this, all the thermodynamical quantities.

The specific heat of the monomer is displayed in figure 4(a) and shows a (broaden) peak at 3.8 kJ/mol and a shoulder extending towards lower temperatures. The average fraction of native energy $q_E \equiv E/E_N$ (solid curve in figure 4(b)) displays a very weak transition centered at the temperature 3.8 kJ/mol of the peak in the specific heat. The quantities which undergo a sharper variation are the fraction of energy between the fragments 24–34 and 83–93 of the HIV-Pr (dotted curve in figure 4(b)) and the dRMSD (figure 4(c)).

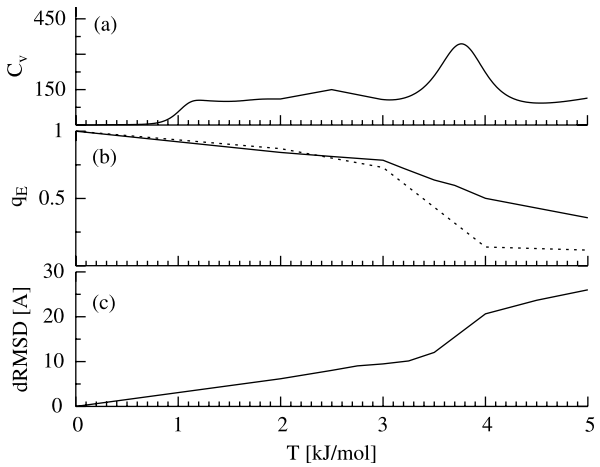

Figure 4. (a) The specific heat of the HIV-1-Pr monomer calculated with Monte Carlo simulations; (b) the order parameter q_E associated with the contacts within the whole monomers (solid curve) and within the nucleus (dotted curve; contacts between fragment 22–32 and 83–93); (c) the dRMSD associated with the monomer.

The fact that the peak in the specific heat corresponds to the temperature where q_E and the dRMSD are most rapid suggests that the peak corresponds to the folding transition (i.e. $T_f = 3.8$ kJ/mol).[†] However, such folding transition is weakly cooperative. In fact, the two-state parameter [22] κ_2 results only in 0.18. This parameter ranges from 1 for fully cooperative transitions, to 0 for non-cooperative transitions. Although it is well known that simplified models underestimate the cooperativity of the folding transition [23], the HIV-1-Pr monomer displays a value of κ_2 which is much lower than that of other proteins simulated with the same model (e.g. src-SH3 displays $\kappa_2 = 0.38$, even if it is shorter than the HIV-1-Pr).

Although the folding transition (figure 4(b), solid curve) takes place as the temperature at which the stabilization core is formed (figure 4(b), dotted curve), as for most globular proteins, the formation of the core seems not to constrain the rest of the protein to the same extent of typical globular proteins. The shoulder in C_v at temperature lower than T_f indicates that, even when the core is formed, the rest of the protein can fluctuate consistently. These fluctuations are those responsible for making the folding transition weakly cooperative and the peak in the specific heat blurred.

In figure 5, the values of q_E and the dRMSD obtained in a simulation at temperature $T = 2.5$ kJ/mol are displayed. The fluctuations in both quantities define the two states (i.e. the native and the unfolded state) already observed in the specific heat. Accordingly, we can define as equilibrium probability of the native state p_N the fraction of MC steps that the system spends in the low-dRMSD, high-q_E state of figure 5. In the following, we will use as thresholds for the definition of p_N the values dRMSD $= 10$ Å and $q_E = 0.7$, although the precise value of these thresholds is not critical. The simulation

[†]Note that the comparison of the scale of temperatures of the model with that of real life is nontrivial, due to the fact that the simplicity of the model inevitably affects the shape of the entropy function, and thus the temperature $T = ((S/(E)^{-1}$.

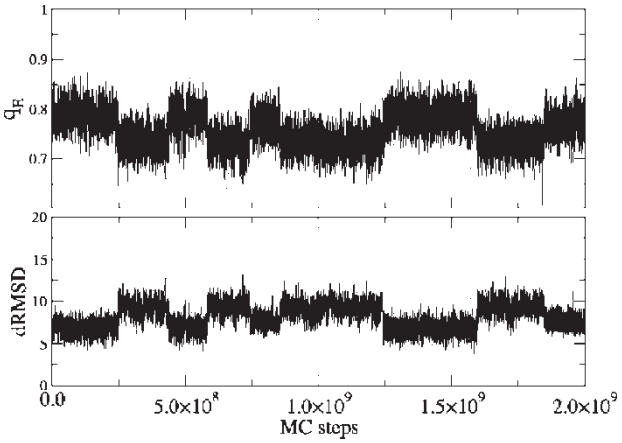

Figure 5. The values of q_E and dRMSD calculated in a Monte Carlo simulation at $T = 2.5$ kJ/mol of the HIV-1-Pr monomer.

at $T = 2.5$ kJ/mol provides a native-state probability $p_N = 0.87$.

4. Inhibitory effects of selected peptides

The effects of peptides displaying the same sequence as fragments of the HIV-Pr on the stability of the native state is evaluated by means of Monte Carlo simulations at $T = 2.5$ kJ/mol of a system composed of the HIV-Pr and n_P peptides constrained in a cubic box of side 100 Å.

First, we study the peptide corresponding to the fragment 83–93 of HIV-Pr. From the joint probability distribution $p(q_E, dRMSD)$ of the native relative energy fraction q_E and of the dRMSD, for the case of three peptides, it is possible to derive the free energy of the system, displayed in figure 6(b), to be compared with that of the monomer alone (figure 6(a)). The well corresponding to the native state (marked by dashed lines) is increased by ≈ 15 kJ/mol with respect to the unfolded state, corresponding to a shift of the equilibrium towards the latter (p_N decreases, in this case, from 0.87 to 0.28, cf. table 1). The dependence of p_N on the number of peptides n_P is shown in table 2 and indicates that a ratio between proteins and peptides of 1:3 guarantees an efficient inhibition.

The reason for the destabilizing effect is simple. Since the peptide mimics a LES of the protein, its interaction with the rest of the protein is as strong as the interaction which stabilizes the core of the protein. The entropic cost of binding depends on the concentration of peptides in solution. If the concentration of peptides is high enough, so that such entropy cost is smaller than the entropic cost associated with the folding of the protein, the binding of the peptide to the protein (which is incompatible with its folding) is favorable with respect to folding.

The behavior of the peptide corresponding to the fragment 24–34, although being a LES as well, is different. The plot of the free energy $F(q_E, dRMSD)$ of the system composed of three copies of this peptide and the protein is displayed in figure 6(c). The corresponding value of p_N is 0.6, indicating

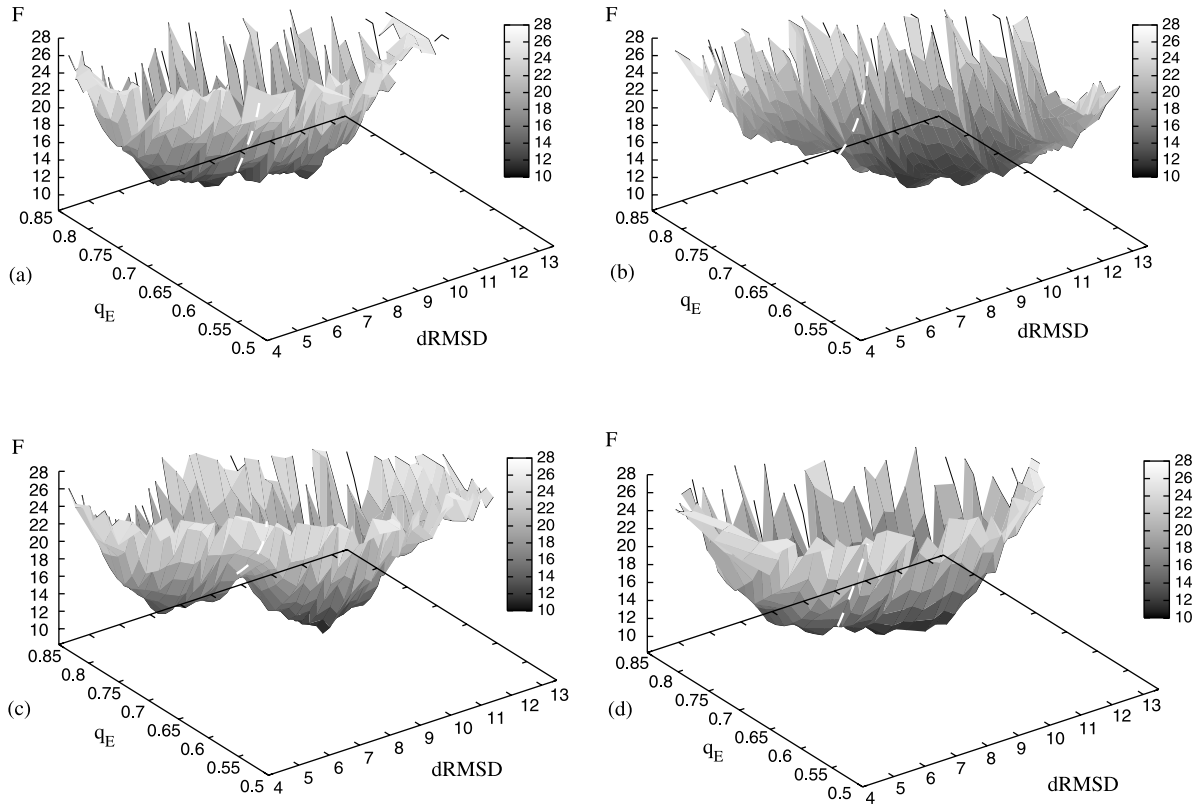

Figure 6. The free energy of the system as a function of the order parameter q_E and of the dRMSD for the HIV-Pr alone (a), the HIV-Pr together with three peptides of kind 83–93 (b), the HIV-Pr together with three peptides of kind 24–34 (c) and the HIV-Pr together with three peptides of kind 5–15 (d).

Table 1. The effect of mutations on a number of sites of the monomer on the stability p_N of the native state.

Mut	p_N *alone*	p_N *with pept*
Wt	0.87	0.28
19	0.93	0.12
31	0.85	0.14
33	0.46	0.31
37	0.95	0.16
63	0.75	0.09
67	0.90	0.14
68	0.96	0.23
72	0.91	0.06
85	0.49	0.14
95	0.90	0.09

The second column corresponds to the stability of the monomer alone ($T = 2.5$ kJ/mol), while the third column refers to the case of the monomer plus 3 peptides of kind 83–93.

an inhibitory effect of this p-LES smaller than of the peptide 83–93. The reason for this difference is that the peptide 24–34 can bind the LES 83–93 staying partially together with the native LES 24–34. In other words, the LES and the p-LES are not perfectly mutually exclusive. This is because the LES 24–34 is at the surface of the monomer, and consequently, the p-LES can bind it without clashing with the rest of the protein.

As a control, we have tested the effect of peptides which do not correspond to LES of the protein. Figure 6(d) shows the free energy of the protein together with three peptides of kind 5–15. The effect of the peptide is negligible, displaying a p_N of 0.85, essentially identical to that without peptides.

The effects of a point mutation on the wild-type sequence of HIV-1-Pr are simulated by switching off the interaction of the mutated site. This is meant to reflect a random mutation which substitutes a residues selected by evolution in order to stabilize the native conformation of the protein with a residue whose interaction with the rest of the protein are random quantities, summing up to zero.

Table 1 displays the effects of point mutations in selected sites of the protein on the equilibrium stability p_N of the native state. As expected in Section II, there are some mutations which, being "hot" sites, destabilize consistently the protein, while most of them have little effect. Thus the sites 33 and 85 are among them (note that site 85 was not identified as "hot" in figure 3, lying at the borderline).

The third column of table 1 represents the stability of the mutated protein in a simulation performed together with three peptides of kind 83–93. The plot indicates that most mutations do not affect the inhibitory effect of the peptides. The only mutations which diminish the affinity of the protein with the peptides are those in "hot" sites, but in this case, the protein is already destabilized by the mutation itself.

This result is quite important from the point of view of using the peptide 83–93 as drug, since it shows that the protein has no escape from it. If the fast mutation rate typical of HIV-1-Pr produces mutations in most of the sites, the inhibitory peptides remains effective. If, on the other hand, "hot" sites are mutated, the inhibitory peptides fail, but the protein in any case is destabilized.

Of course, the HIV-1-Pr can find more complex evolution patterns, which rely on multiple mutations to try to escape the drug. However, to do so the protein should build a new stabilization core, completely different from the wild-type one. This is a task which is likely to be lengthy enough to cause the death of the virus, or at least to give enough time to the researchers to replicate the above procedure on the mutated sequence.

Table 2. The equilibrium probability $p_N(\equiv p(q_E > 0.7,$ dRMDS < 10 Å)) of the native state of the monomer as a function of the number n_P of peptides 83–93.

n_P	p_N
0	0.87
1	0.63
2	0.43
3	0.28
4	0.25
5	0.24

5. Conclusions

We have described a novel way to design inhibitory peptides, especially effective against a retroviral protein as HIV-1 protease. The sequence of these peptides is suggested directly from the protein one wish to inhibit, corresponding to fragments containing "hot" sites which are important for its stability and cannot be mutated under the threat of unfolding. Model simulations show that a peptide corresponding to the fragment 83–93 of HIV-1-Pr is particularly effective in destabilizing the protein and cannot be escaped by point mutations.

References

[1] G.M. Verkhivker. Computational analysis of ligand binding dynamics at the intermolecular hot spots with the aid of simulated tempering and binding free energy calculations. *J. Mol. Graph. Model.*, **22**, 335 (2004).
[2] T.J. Marrone, J.M. Briggs, J.A. McCammon. Structure-based drug design. *Annu. Rev. Pharmacol. Toxicol.*, **37**, 71 (1997).
[3] L.A. Mirny, E.I. Shakhnovich. Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.*, **291**, 177 (1999).
[4] F. Ding, N.V. Dokholyan, S.V. Buldyrev, H.E. Stanley, E.I. Shakhnovich. Formation of the folding nucleus of an SH3 domain investigated by loosely coupled molecular dynamics. *Biophys. J.*, 83 (2002).
[5] R.A. Broglia, G. Tiana, R. Berera. Resistance proof, folding-inhibitor drugs. *J. Chem. Phys.*, **118**, 4754 (2003).
[6] R.A. Broglia, G. Tiana, L. Sutto, D. Provasi, F. Simona. Design of HIV-1-PR inhibitors which do not create resistance: blocking the folding of single monomers. *Protein Sci.*, (in press) (q-bio/0504011).
[7] G. Tiana, R.A. Broglia, H.E. Roman, E. Vigezzi, E. Shakhnovich. Folding and misfolding of designed protein-like folding and misfolding of designed protein-like chains with mutations. *J. Chem. Phys.*, **108**, 757 (1998).
[8] R.A. Broglia, G. Tiana, D. Provasi. Simple model of protein folding and of non-conventional drug design. *J. Phys. Cond. Matt.*, **16**, R111 (2004).
[9] R.A. Broglia, G. Tiana. Hierarchy of events in the folding of model proteins. *J. Chem. Phys.*, **114**, 7267 (2001).

[10] G. Tiana, R.A. Broglia. Statistical analysis of native contact formation in the folding of designed model proteins. *J. Chem. Phys.*, **114**, 2503 (2001).
[11] N. Go. Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioengin.*, **12**, 183 (1983).
[12] H.J.C. Berendsen, D. van der Spoel, R. van Drunen. GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.*, **91**, 43 (1995).
[13] G. Tiana, F. Simona, G.M.S. De Mori, R.A. Broglia, G. Colombo. Understanding the determinants of stability and folding of small globular proteins from their energetics. *Protein Sci.*, **13**, 113 (2004).
[14] A.G. Tomasselli, R.L. Heinrikson. Targetting the HIV-protease in AIDS therapy: a current clinical perspective. *Biochem. Biophys. Acta*, **1477**, 189 (2000).
[15] D. Xie, S. Gulnik, E. Gustchina, B. Yu, W. Shao, W. Qoronfleh, A. Nathan, J.W. Erickson. Drug resistance mutations can effect dimer stability of HIV-1 protease at neutral pH. *Protein Sci.*, **8**, 1702 (1999).
[16] M.L. Metha. *Random Matrices*. Academic Press, San Diego (1990).
[17] R.W. Shafer, P. Hsu, A.K. Patick, C. Craig, V. Brendel. Identification of biased amino acid substitution patterns in human immunodeficiency virus type 1. *J. Virol.*, **73**, 6197 (1999).
[18] I. Bahar, A.R. Atilgan, M.C. Demirel, B. Erman. Vibrational dynamics of folded proteins: significance of slow and fast motions in relation to function and stability. *Phys. Rev. Lett.*, **80**, 2733 (1998).
[19] F. Cecconi, C. Micheletti, P. Carloni, A. Maritan. Molecular dynamics studies on HIV-1 protease: drug resistance and folding pathways. *Proteins*, **43**, 365 (2001).
[20] 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**, 67 (2004).
[21] J. Borg. Optimized Monte Carlo analysis for generalized ensembles. *Eur. Phys. J. B*, **29**, 481 (2001).
[22] H. Kaya, H.S. Chan. Polymer principles of calorimetric two-state cooperativity. *Proteins*, **40**, 637 (2000).
[23] M.S. Li, D.K. Klimov, D. Thirumalai. Finite-size effects on calorimetric operativity of two-state proteins. (2004). *q-bio/0412008*