Journal of Molecular Graphics and Modelling 31 (2011) 65–75

Contents lists available at SciVerse ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Identification of amino acid residues of a designed ankyrin repeat protein potentially involved in intermolecular interactions with CD4: Analysis by molecular dynamics simulations

Piyarat Nimmanpipug[a,b,*], Chalermpon Khampa[a,b], Vannajan Sanghiran Lee[b,c], Sawitree Nangola[d], Chatchai Tayapiwatana[d,e,**]

[a] Department of Chemistry, Center of Excellence for Innovation in Chemistry and Materials Science Research Center, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
[b] Thailand Center of Excellence in Physics, Commission on Higher Education, Bangkok 10400, Thailand
[c] Department of Chemistry, Faculty of Science, University of Malaya, 50603, Kuala Lumpur, Malaysia
[d] Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand
[e] Biomedical Technology Research Unit, National Center for Genetic Engineering and Biotechnology, National Science and Technology Development Agency at the Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand

ARTICLE INFO

Article history:
Received 1 June 2011
Received in revised form 30 August 2011
Accepted 1 September 2011
Available online 10 September 2011

Keywords:
Ankyrin
CD4
HIV
Protein docking
Molecular dynamics simulations

ABSTRACT

We applied molecular dynamics simulations to investigate the binding properties of a designed ankyrin repeat protein, the DARPin–CD4 complex. DARPin 23.2 has been reported to disturb the human immunodeficiency virus (HIV) viral entry process by Schweizer et al. The protein docking simulation was analysed by comparing the specific ankyrin binder (DARPin 23.2) to an irrelevant control (2JAB) in forming a composite with CD4. To determine the binding free energy of both ankyrins, the MM/PBSA and MM/GBSA protocols were used. The free energy decomposition of both complexes were analysed to explore the role of certain amino acid residues in complex configuration. Interestingly, the molecular docking analysis of DARPin 23.2 revealed a similar CD4 interaction regarding the gp120 theoretical anchoring motif. In contrast, the binding of control ankyrin to CD4 occurred at a different location. This observation suggests that there is an advantage to the molecular modification of DARPin 23.2, an enhanced affinity for CD4.
© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The evolution and mutation of the human immunodeficiency virus (HIV) are matters of concern in contemporary HIV therapy. There are currently a large number of HIV variants that are able to avoid human immune recognition and evade host cytotoxic lymphocytes (CTL). The HIV-1 virus incessantly changes its surface by mutation of the key epitopes (minimum structural units that can be recognised by a B or T cell receptor) [1]. HIV drugs are one of the success stories in drug design, and several have been developed to suppress HIV-1 replication, including HIV-1 reverse transcriptase (HIV-1RT) inhibitors [2–4] such as Abacavir, Lamivudine, Zidovudine, Didanosine, Emtricitabine, Tenofovir, Lamivudine, and Stavudine. Other drugs act as HIV-1 protease (HIV-1PR) inhibitors, for example, Amprenavir, Indinavir, Lopinavir, Nelfinavir, Ritonavir, and Saquinavir [5–11]. More recently, integrase inhibitors have been developed, including Raltegravir [12–14]. The current treatments have encountered drug resistance problems from virus mutations and, in addition, can cause serious side effects. Analysis of the HIV-CD4 binding site has shown that it is relatively stable against mutations and is the main entry receptor of HIV [15]; therefore, this site is a promising drug target, regardless of protein alteration problems.

Recently, a novel approach for inhibiting HIV infection using Designed Ankyrin Repeat Protein (DARPin) technology has been developed [16,17]. Ankyrin repeat proteins are a family of proteins that are found across multiple species and mediate protein–protein interactions in various cell compartments [18]. These proteins are composed of stacked repeats containing 33 amino acids. Each repeat is formed by two antiparallel α-helices and a β-turn connecting to the next repeat [19]. These repeats are flanked by

* Corresponding author at: Department of Chemistry, Center of Excellence for Innovation in Chemistry and Materials Science Research Center, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand. Tel.: +66 53943341; fax: +66 53892277.
** Corresponding author at: Division of Clinical Immunology, Department of Medical Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand Tel.: +66 818845141; fax: +66 53946042.
E-mail addresses: piyaratn@gmail.com (P. Nimmanpipug), asimi002@chiangmai.ac.th (C. Tayapiwatana).

1093-3263/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2011.09.003

(a)

```
SDLGKKLLEAARAGQDDEVRILMANGADVNAK        = n-cap
DEYGLTPLYLATAHGHLEIVEVLLKNGADVNAV       = 1st repeat
DAIGFTPLHLAAFIGHLEIAEVLLKHGADVNAQ       = 2nd repeat
DKFGKTAFDISIGNGNEDLAEILQKL              = c-cap
```

(b)

```
DLGKKLLEAARAGQDDEVRILMANGADVNAT         = n-cap
DTLGRTPLHMAAAWGHLEIVDVLLKHGADVNAI       = 1st repeat
EEVGMTPLHLAAFLGHLEIVEVLLKSGADVNAQ       = 2nd repeat
DKFGKTAFDISIDYGNEDLAEILQ                = c-cap
```

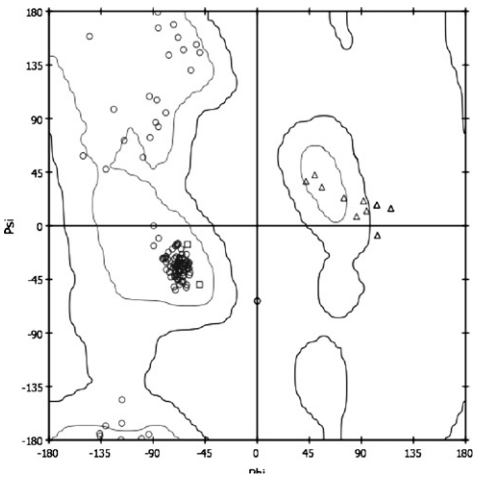
(c)

```
                        10        20        30        40        50
                        |         |         |         |         |
CD4-Ank control     KKVVLGKKGDTVELTCTASQKKSIQFHWKNSNQIKILGNQGSFLTKGPSK
CD4-DARPin 23.2     KKVVLGKKGDTVELTCTASQKKSIQFHWKNSNQIKILGNQGSFLTKGPSK
                        60        70        80        90        100
                        |         |         |         |         |
CD4-Ank control     LNDRADSRRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLLVFGL
CD4-DARPin 23.2     LNDRADSRRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQLLVFGL
                        110       120       130       140       150
                        |         |         |         |         |
CD4-Ank control     TANSDTHLLQGQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLE
CD4-DARPin 23.2     TANSDTHLLQGQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLE
                        160       170       180       190       200
                        |         |         |         |         |
CD4-Ank control     LQDSGTWTCTVLQNQKKVEFKIDIVVLA/SDLGKKLLEAARAGQDDEVRI
CD4-DARPin 23.2     LQDSGTWTCTVLQNQKKVEFKIDIVVLA/-DLGKKLLEAARAGQDDEVRI
                        210       220       230       240       250
                        |         |         |         |         |
CD4-Ank control     LMANGADVNAKDEYGLTPLYLATAHGHLEIVEVLLKNGADVNAVDAIGFT
CD4-DARPin 23.2     LMANGADVNATDTLGRTPLHMAAAWGHLEIVDVLLKHGADVNAIEEVGMT
                        260       270       280       290       300
                        |         |         |         |         |
CD4-Ank control     PLHLAAFIGHLEIAEVLLKHGADVNAQDKFGKTAFDISIGNGNEDLAEIL
CD4-DARPin 23.2     PLHLAAFLGHLEIVEVLLKSGADVNAQDKFGKTAFDISIDYGNEDLAEIL
                        310
                        |
CD4-Ank control     QKL-------
CD4-DARPin 23.2     Q---------
```

Fig. 1. Amino sequence and CD4 complex alignment of (a) irrelevant ankyrin control, pdb id = 2JAB (binding site: Glu45, Tyr46, Leu48, Ala56, His57, Val76, Ala78, Ile79, Phe81, Phe89, Ile90, Gln109), (b) DARPin 23.2 (binding site: Thr33, Lue34, Arg36, Ala44, Trp45, Ile64, Glu66, Val67, Met69, Phe77, Leu78, Gln97), and (c) full sequence of the ankyrin control–CD4 and the DARPin 23.2–CD4 model with the amino acid differences highlighted in bold (D1 and D2 domains of CD4 are residue numbers 1–89 and 90–178, respectively).

constant capping regions, forming one contiguous polypeptide chain. Target protein binding involves the tips of the β-hairpins and the surface of the helical bundle facing the concave ankyrin groove [20]. Advantages of using ankyrin repeat proteins are that they displays tight, specific binding at the nanomolar range to their target proteins, and each repeat can contribute to target binding [21–23]. This provides direct evidence that DARPins can bind rapidly and selectively to their intended target cells *in vivo* and are one promising avenue of research in which problems caused by protein mutations and alterations may be less severe.

DARPin proteins specific for human CD4 were selected from a DARPin library using a ribosome display [16]. CD4-specific DARPins are highly specific and efficiently prevent viral infection *in vitro* in a wide range of HIV isolates without disturbing basic cellular functions. In the initial step leading to HIV infection, gp120 interacts with CD4 molecules on the surface of T lymphocytes, and this crucial process can be efficiently blocked by these selected DARPins



Number of residue in core region	111 (91.7%)
Number of residue in allowed region	7 (5.8%)
Number of residue in disallowed region	3 (2.5%)

Fig. 2. Homology structures of DARPin 23.2 using the 2JAB template in the DS studio 2.0 program package. The amino acid residues located in the binding site i.e. Thr33, Lue34, Arg36, Ala44, Trp45, Ile64, Glu66, Val67, Met69, Phe77, Leu78, and Gln97 were designated in black. The best model, with 81.5% sequence identity and 89.5% sequence similarity, had the highest verify score of 56.01.

[24]. Studies have shown that the CD4-specific DARPin 57.2 efficiently binds to CD4+ T cells, DCs, and monocytes *in vivo* but has no effect on SHIV plasma levels after a single intravenous injection into SHIV-infected rhesus macaques. Moreover, the clearance of DARPin 57.2 from the plasma is rapid, with the rate proportional to the dosage [25]. Therefore, understanding the interaction between DARPins and CD4 molecules may provide a solution for improving the binding activity of DARPins.

Our study aims to clarify the roles of crucial residues in the binding interface and their interaction contributions in a potential ankyrin repeat protein–CD4 complex, compared to an irrelevant ankyrin binder. For large, multidomain proteins like this system, domain docking is the most logical method of analysis, mimicking the hierarchical nature of protein folding, as the complexity of the problem is very large. The field of protein–protein docking is highly computationally oriented, and shares approaches with small-molecule docking. A variety of methods have been developed and reviewed [26,27]. A popular technique, using a fast Fourier transform (FFT) correlation-based method to check for surface complementarity, was initially proposed by Katchalski-Katzir et al. [28]. The programs DOT [29], GRAMM [30], FTDOCK [31], and ZDOCK [32,33] include modifications of earlier concepts to successfully predict protein complex structures. For the initial stage of docking to pre protein–protein interactions, we selected a robust rigid-body docking algorithm with a well-parameterised scoring function, ZDOCK. This method has been shown to be highly effective in predicting protein–protein complexes. The method takes into account the major determinants of protein–protein association, that is, shape, surface, and electrostatic complementarities, as well as desolvation, using the fast Fourier transform algorithm. In the next stage, the refinement stage, slight improvements to the initial docking stage were performed. The molecular recognition was further investigated using molecular dynamics simulations to follow molecular motions, changes and key interactions governing the binding potential of a designed ankyrin repeat protein with CD4. The free energy of binding and its decomposition per residue were clarified to identify crucial residues involved in the improvement in CD4 binding of the designed ankyrin.

2. Methodology

2.1. Ankyrin model design

The amino acid sequence of a synthetic ankyrin repeat protein that binds specifically to the CD4 molecule clone 23.2 was selected from the previous work of Schweizer et al. [16] and homology modelling was used to generate a model based on the structure of irrelevant ankyrin binders that have high affinity for DARPin binding to human epidermal growth factor receptor 2 (Her2) (pdb code: 2JAB) [21] (Fig. 1). From experimental evident [16,25], selected DARPin, 23.2 interacted specifically with CD4 and competed with gp120 for binding to CD4 with nM scale K_D and the CD4-specific DARPin can rapidly and selectively bind its target cells *in vivo* while controlled or irrelevant ankyrin shows non-specific binding. Therefore, we are interested in performing the theoretical study for comparison. Each domain of CD4 is specified in the figure. The template structure was modified at several positions to obtain the designed ankyrin, including residues located at the binding site: Thr33, Lue34, Arg36, Ala44, Trp45, Ile64, Glu66, Val67, Met69, Phe77, Leu78, Gln97. The model was built by MODELLER in Discovery Studio using the Discrete Optimised Protein Energy (DOPE) scoring function, which is suitable for single chain modelling [34].

Table 1
Residues involved in the binding of ankyrin/designed ankyrin–CD4 binding complexes evaluated by protein docking.

CD4–designed ankyrin		CD4–ankyrin template	
Amino acids interface of CD4	Amino acids interface of designed ankyrin	Amino acids interface of CD4	Amino acids interface of ankyrin template
GLN20, LYS22, SER23, ILE24, GLN25^{HB1,HB2,HB3}, PHE26, HIS27, LYS35, GLY38, ASN39, GLN40 ^{HB4,HB5}, GLY41^{HB6},SER42, PHE43, THR45, ARG59, GLU87	THR33, LEU34, GLY35, ARG36, GLU65, GLU66, VAL67^{HB1}, GLY68, MET69, HIS73, LEU74, PHE77, ASP98^{HB4,HB6}, LYS99^{HB2,HB3}, PHE100, GLY101, LYS102^{HB5}, ASP106, ILE107, ASP110, TYR111	LYS7^{pi}, LYS8, GLY9, ASP10, THR11, GLU13, ARG58, LYS72, ASN73^{HB1}, ILE76, PHE98, GLY99, LEU100, THR101, ALA102, THR106, GLU119, THR156, TRP157, THR160, LYS167, VAL168, GLU169^{HB2}, PHE170, LYS171^{HB3,HB4,HB5}, ILE172, ASP173^{HB6}	SER12, LYS16, ARG23^{HB1}, ALA24, LYS43^{HB3,HB6}, ASP44, GLU45, TYR46^{HB4}, GLY47, LEU48, TYR52^{pi}, LEU53, ALA56, HIS57, VAL76, ASP77, ALA78^{HB5}, ILE79, GLY80, PHE81, PHE89, ILE90, LYS111^{HB2}, PHE112

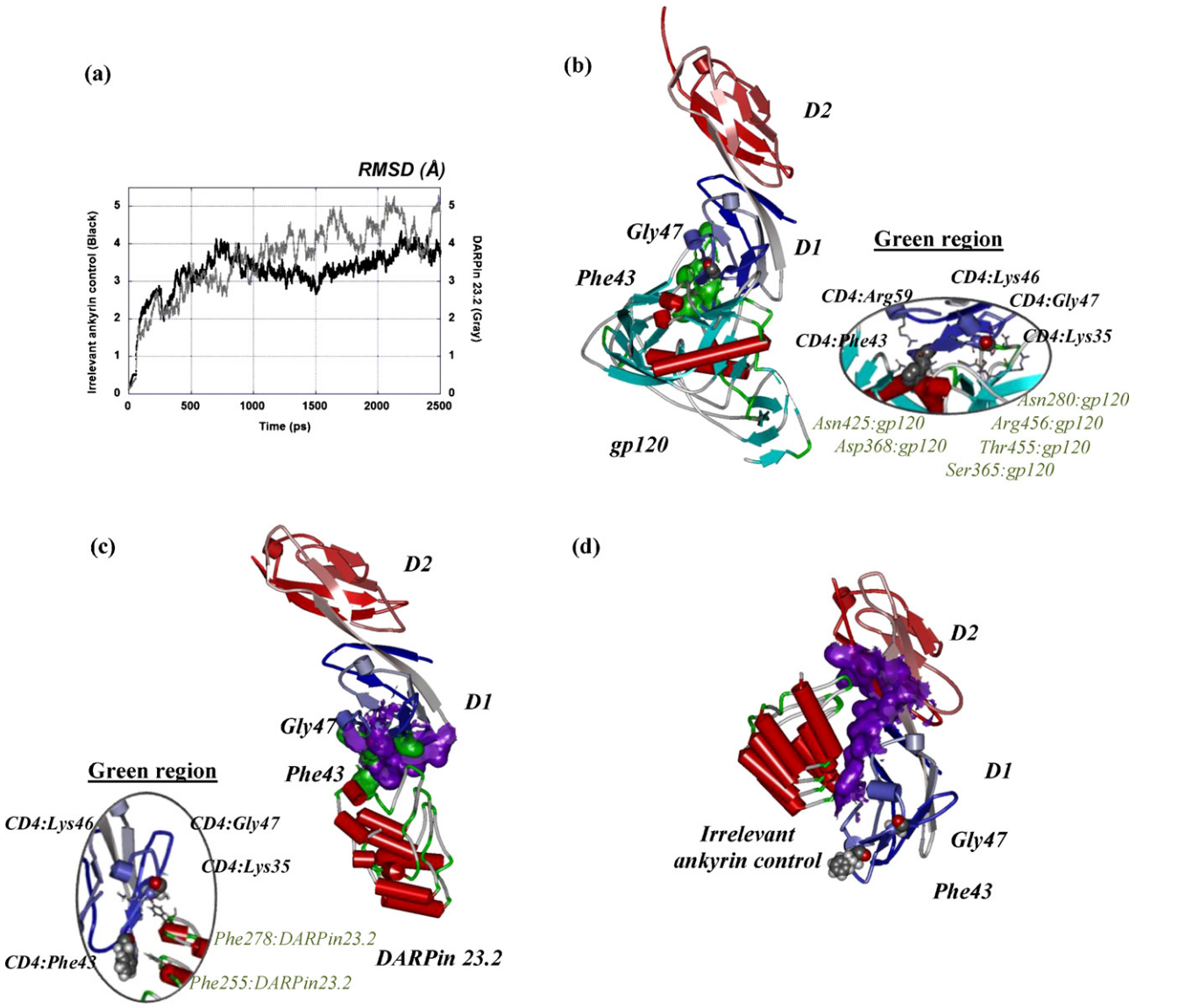


Fig. 3. Ankyrin–CD4 binding complexes evaluated from MD simulations using average structures after reaching equilibrium. (a) RMSD plot of ankyrin–CD4 binding complexes evaluated from MD simulations. The binding interface of (b) gp120–CD4 (pdb id = 2NXY) with the CD4's amino acid interface of Lys35, Phe43, Lys46, Gly47 and Arg59 (green) coinciding with critical residues in the gp120 recognition site [44,42,15] as in stick model (zoom view), of (c) the DARPin 23.2–CD4 with the amino acid interface of Lys35, Phe43, Lys46, Gly47 (green). The amino acids of DARPin 23.2 in binding with CD4 are shown in stick model (zoom view). The binding interface of (d) the ankyrin control–CD4 with different amino acids interface is shown. The amino acids that shift from the binding interface in comparison to gp120–CD4 are indicated in purple.

The optimum model was chosen based on the "verify" score [35] available in the DS program and the root mean square distance (RMSD). The RMSD of a model from its templates should be very small and indicates how well a model fits the templates on which it was built.

2.2. Protein docking

Designed ankyrin–CD4 docking simulations were carried out using ZDOCK [32,33] in the DS 2.5 program package. Atomic charges were assigned in a CHARMm polarH force field. Rigid-body sampling around CD4 (pdb code: 3CD4) [36] was used to generate 54,000 structures with a small cluster radius of 6.0 Å for the RMSD cutoff and a smaller interface cutoff of 9.0 Å around each hypothesised binding region. The obtained models were clustered to 2000 complexes with the highest scores for refinement concerning all short/long length interactions (*van der Waals* and electrostatic interactions) and desolvation energies, based on the atomic contact energy implemented in the program. Residues involving the binding site together with intermolecular interactions were investigated.

2.3. Molecular dynamics simulations

Designed ankyrin–CD4 complexes were solvated, prior to dynamic simulations, at the molecular mechanic level using the parm99 force field in the AMBER9 program suite [37]. All molecules were solvated in a cubic box of TIP3P water that extended to within 10 Å from the solute in each direction, and the cutoff distance was kept at 12 Å to compute the non-bonded interactions. All simulations were performed under periodic boundary conditions, and long-range electrostatics were treated using the particle-mesh-Ewald method [38].

Molecular dynamics began by heating each of the minimised systems to 310 K while constraining the protein atoms with a force of 5 kcal/mol/Å^2, followed by a 200 ps pressure-constant period to raise the density while keeping the protein atoms constrained. For the next step, the run time was 200 ps for each volume-constant

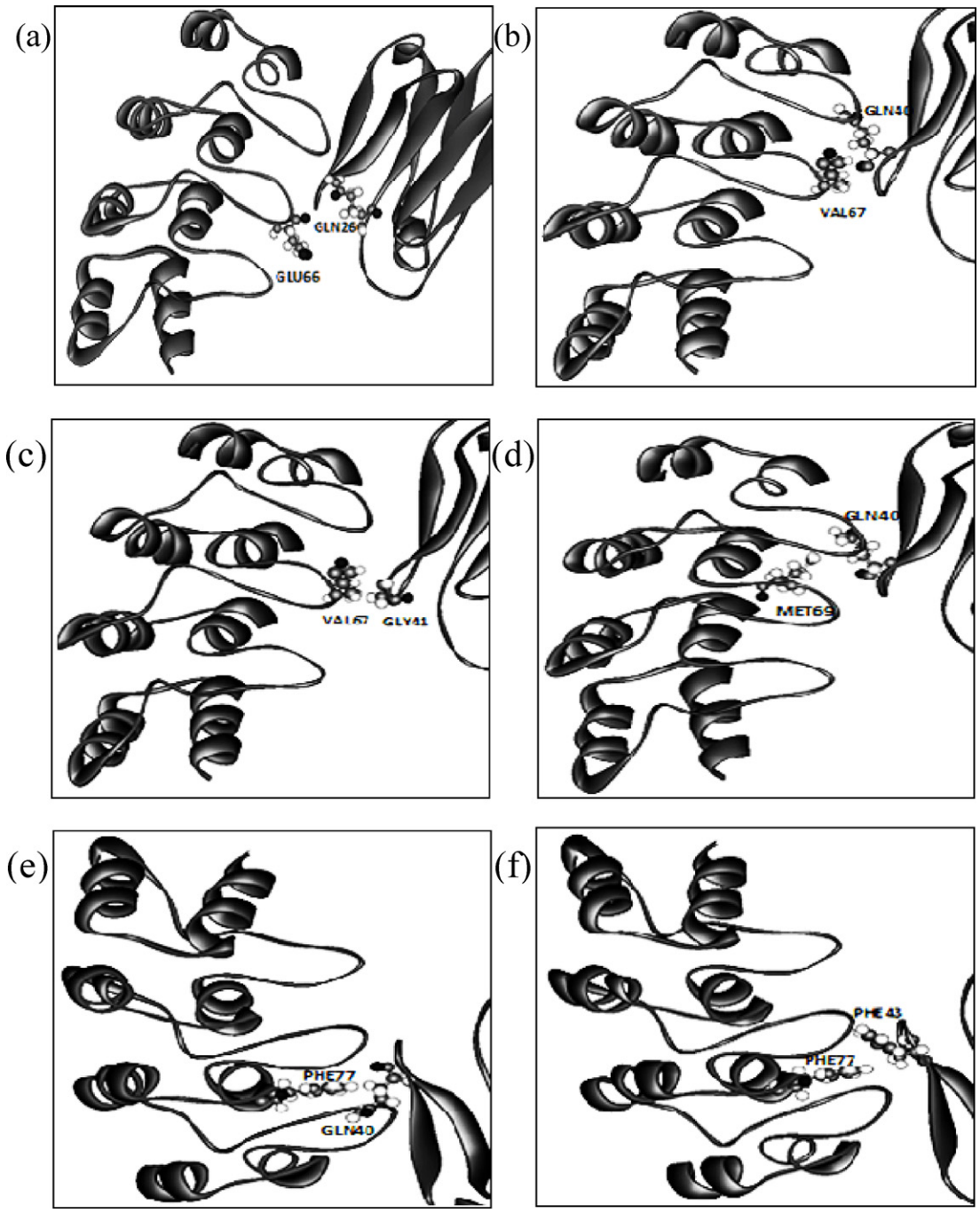


Fig. 4. Relevant mutated residues of the designed ankyrin governing intermolecular interactions with CD4 in 3 Å: (a) Glu66: DARPin 23.2 interacting with Gln26:CD4, (b) Val67:DARPin 23.2–Gln40:CD4, (c) Val67: DARPin 23.2–Gly41:CD4, (d) Met69: DARPin 23.2–Gln40:CD4, (e) Phe77: DARPin 23.2–Gln40:CD4, and (f) Phe77: DARPin 23.2–Phe43:CD4.

period with a force of 2.5 and 1.25 kcal/mol/Å^2, followed by a 400 ps simulation period without the force constrained. Finally, 1-ns dynamics calculations were performed for each free system in the NVT ensemble at a constant temperature of 310 K. For all of the simulations, the total run time was 2 ns with a time step of 2 fs throughout the simulation. The trajectory was recorded every 1 ps.

2.4. Binding free energy and free energy decomposition calculations: MM-PBSA/MM-GBSA

Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) [39] computes the binding free energy using a thermodynamic path that includes the solvation contribution. A total of 500 snapshots were taken from the last 500 ps of the MD trajectory. The binding free energies ($\Delta G_{\text{binding}}$) were evaluated and determined from the free energies of the complex, proteins and peptides, according to the equation:

$$\Delta G_{\text{binding}} = \Delta G_{\text{water}}(\text{ankyrin–CD4 complex}) - [\Delta G_{\text{water}}(\text{ankyrin}) + \Delta G_{\text{water}}(\text{CD4})]$$

The binding free energies for each species, in turn, were estimated from the absolute molecular mechanical energies (E_{MM}), the solvation free energies ($G_{\text{PB}} + G_{\text{nonpolar}}$) and the vibration, rotation

Table 2
Hydrogen bonds distances, angles, and percentages of occupation for binding. Intra-residue interactions are not included. The distance cutoff is 3.0 Å.

Donor	Acceptor H	Acceptor	%Occupied	Distance (Å)	Angle (°)
Residue@atom	Residue@atom	Residue@atom			
CD4--designed ankyrin					
98@OD1:ankyrin	40@HE22:CD4	40@NE2:CD4	88.00	2.82 (−0.09)	19.13 (−10.06)
87@OE2:CD4	99@HZ3:ankyrin	99@NZ:ankyrin	40.14	2.81 (−0.09)	23.49 (−11.47)
25@OE1:CD4	99@HZ2:ankyrin	99@NZ:ankyrin	18.07	2.84 (−0.09)	34.17 (−12.51)
87@OE1:CD4	99@HZ3:ankyrin	99@NZ:ankyrin	15.03	2.83 (−0.10)	37.00 (−15.08)
87@OE2:CD4	99@HZ2:ankyrin	99@NZ:ankyrin	13.67	2.81 (−0.09)	23.09 (−13.14)
87@OE1:CD4	99@HZ1:ankyrin	99@NZ:ankyrin	11.84	2.82 (−0.10)	22.09 (−11.20)
100@O:ankyrin	35@HZ3:CD4	35@NZ:CD4	10.40	2.86 (−0.09)	26.09 (−13.15)
CD4–ankyrin template					
9@O:CD4	23@HH11:ankyrin	23@NH1:ankyrin	40.43	2.87 (0.08)	26.72 (−10.40)
10@OD2:CD4	52@HH:ankyrin	52@OH:ankyrin	37.83	2.80 (0.12)	23.98 (−13.41)
100@O:CD4	46@HH:ankyrin	46@OH:ankyrin	34.93	2.81 (0.11)	24.05 (−12.49)
78@O:ankyrin	171@HZ1:CD4	171@NZ:CD4	26.39	2.84 (0.09)	28.44 (−13.75)
10@OD1:CD4	52@HH:ankyrin	52@OH:ankyrin	23.19	2.82 (0.11)	26.21 (−13.63)
169@OE1:CD4	111@HZ2:ankyrin	111@NZ:ankyrin	13.94	2.80 (0.10)	32.32 (−14.13)
78@O:ankyrin	171@HZ2:CD4	171@NZ:CD4	13.54	2.84 (0.09)	29.75 (−13.70)
123@OD1:ankyrin	58@HH12:CD4	58@NH1:CD4	13.09	2.85 (0.08)	28.76 (−13.82)
169@OE2:CD4	111@HZ2:ankyrin	111@NZ:ankyrin	10.74	2.84 (0.10)	31.72 (−14.17)

and translation entropies. Each of these terms was calculated using the following equations:

$$\Delta G_{\text{water}} = E_{\text{MM}} + \Delta G_{\text{solvation}} - T \quad S$$

$$G_{\text{solvation}} = G_{\text{solvation}-\text{electrostatic}} + G_{\text{nonpolar}}$$

$$E_{\text{MM}} = E_{\text{internal}} + E_{\text{electrostatic}} + E_{\text{vdW}}$$

$$E_{\text{internal}} = E_{\text{bond}} + E_{\text{angle}} - E_{\text{torsion}}$$

where the internal energy E_{internal} has three contributions: E_{bond}, E_{angle}, and E_{torsion}, which represent the strain energy in bonds, angles and torsion angles caused by their deviation from the equilibrium values; $E_{\text{electrostatic}}$ and E_{vdW} are the electrostatic and *van der Waals* interaction energies, respectively; T is the temperature, S is the entropy. E_{MM} was determined with the *anal* program from AMBER modelling software. An infinite cutoff for all interactions and parm99 force field parameters were used. The contribution of the change to conformational entropy during complex formation was estimated with the *nmode* module from AMBER. For the molecular mechanics-generalised Born solvent area (MM-GBSA) method, G_{PB} was replaced by G_{GB}. The Hawkins, Cramer and Truhlar pairwise generalised Born model [1,40] was used with parameters described by Tsui and Case [41]. The GB model makes this variant attractive because it is much faster than the PB model and allows for the decomposition of the electrostatic solvation free energy into atomic contributions in a straightforward manner. This model allows simple and rapid free energy decomposition.

3. Results and discussion

3.1. Probing ankyrin–CD4 interactions compared with the CD4–gp120 complex: residues involved in binding

Designed ankyrin structures modelled from homology modelling were all well built with a low RMSD value below 1 Å. The best candidate with the highest verify score had 111 residues (91.7%) in the core region and 7 residues (5.8%) in allowed regions, as

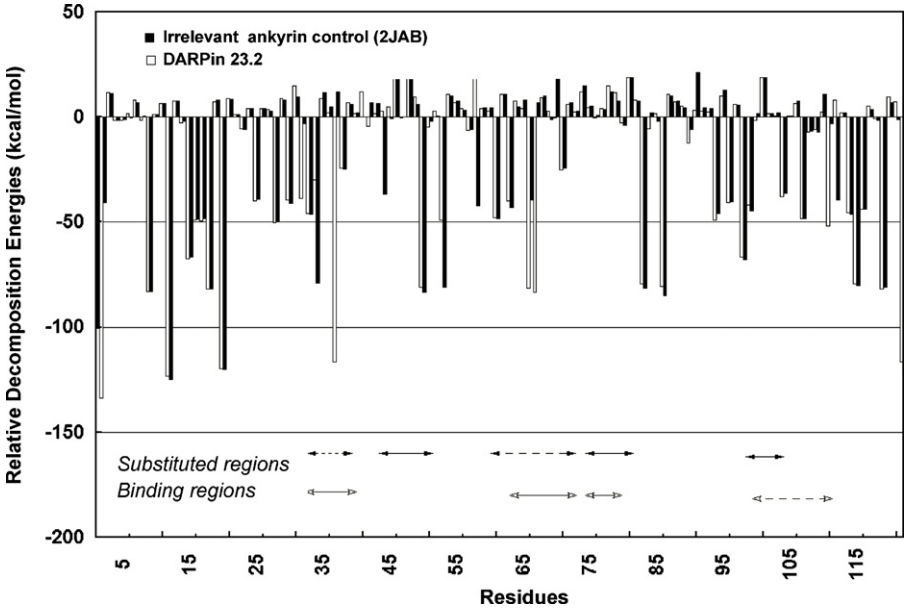


Fig. 5. Relative decomposition energies (kcal/mol) of ankyrin–CD4 (irrelevant ankyrin control/DARPin 23.2) complexes. Substituted region which indicated the modification of the control at several positions to obtain the designed ankyrin and binding region were shown in the plot.

shown in Fig. 2. Although there were 3 residues (2.8%) in the disallowed region, these amino acids (Gly58, Gly68, and Gly101) are not part of the binding site of ankyrin. Targeting of CD4 by high-affinity DARPins can occur without a loss of CD4 T-cell function and unwanted side effects [16]. Thus, we attempted to clarify the specific binding properties of the designed ankyrin with CD4 (23.2) compared with an irrelevant ankyrin control (2JAB). Binding of ankyrin–CD4 complexes was investigated via molecular docking. Intermolecular interactions and changes to the binding interface between complexes were evaluated. According to several structural and kinetic studies of the gp120 receptor, the amino acids located at the CD4–gp120 binding site are Lys35, Phe43, Lys46, Gly47 and Arg59; these are critical residues of the gp120 recognition site [15,42,43].

Key residues involved in the binding of CD4–ankyrin binding complexes identified from protein docking (Table 1) were defined as any residue of the designed ankyrin that had any atom within 5 Å to CD4 and vice versa; the binding residues of the designed ankyrin with Phe43 of the CD4 cavity, which contains highly conserved binding in the D1 domain, were Gln20, Lys22, Ser23, Ile24, Gln25, Phe26, His27, Lys35, Gly38, Asn39, Gln40, Gly41, Ser42, Phe43, Thr45, Arg59, and Glu87. Such binding is similar to that determined from the X-ray structure of the CD4–gp120 binding scaffold. However, the ankyrin control exhibited a different binding pattern from gp120 when bound to CD4. The binding interface was covered by different amino acids in both domains D1 and D2; these amino acids were Lys7, Lys8, Gly9, Asp10, Thr11, Glu13, Arg58, Lys72, Asn73, Ile76, Phe98, Gly99, Leu100, Thr101, Ala102, Thr106, Glu119, Thr156, Trp157, Thr160, Lys167, Val168, Glu169, Phe170, Lys171, Ile172, and Asp173. After MD simulations, the binding scaffolds were retained in essentially the same positions as the dock structures (Fig. 3). The green surface in the figure indicates the amino acid interface coinciding with critical residues of the gp120 recognition site [15,42,43]. The purple surface represents the binding interface between ankyrin and CD4 at 5 Å, which did not correlate with the CD4–gp120 recognition site. Our analysis strongly suggests that the designed DARPins have unique properties and that the specific residues of CD4 involved in binding have been identified.

3.2. Hydrogen bond analysis and intermolecular interactions of CD4–ankyrin recognition dynamics

Atomic analysis of the CD4 binding-induced conformational changes in HIV-1 gp120 from molecular dynamics simulations based on the gp120/CD4 crystal structure were previously explored by Hsu and Bonvin 2004 [44]. Three different interaction modes of electrostatics, *van der Waals* forces, and hydrogen bonding between CD4 and gp120 were found. This suggested that the increase in rigidity observed in the complex originates primarily from the formation of intermolecular hydrogen bonds. Many stable intermolecular hydrogen bonds bridge the tips of the mobile loops of gp120 interacting with CD4 and a hemisphere of CD4-D1. This hydrogen bond network nicely encloses the docking cavity of Phe43, which is the main element for the binding mechanism. The percentages of occurrence of the hydrogen bonds of F43:CD4-G473:gp120; L44:CD4-D368:gp120; K46:CD4-S365:gp120; and K46:CD4-G366:gp120 are 11.5, 67.8, 14.2, and 15.7%, respectively. A large number of hydrogen bonds of side chain-backbone (K35, S42, L44, D53, R59, Q64, N66) or side chain–side chain (K22, Q25, K29, R59, Q64, E87, D88) interactions of CD4 cover a large area of the complex interface in a nonspecific manner.

In comparison, we observed the dynamics of torsions and intermolecular hydrogen bonding distances of key residues involved in the binding of ankyrin/designed ankyrin–CD4 binding complexes evaluated by protein docking (Table 1). We initially observed a

Table 3
Energy components contributing to the binding of the CD4–ankyrin and CD4–designed ankyrin complexes in terms of MM/PBSA and MM/GBSA.

Method	Contribution	CD4–Ankyrin						ΔE (kcal/mol)		CD4–Designed Ank						ΔE (kcal/mol)	
		Complex		CD4		Ankyrin				Complex		CD4		Ankyrin			
		MEAN	STD	MEAN	STD	MEAN	STD			MEAN	STD	MEAN	STD	MEAN	STD		
MM																	
	ELE	−6308.24	71.37	−4101.14	55.16	−1531.05	52.94	−676.05	±28.55	−5679.95	97.23	−4066.11	70.99	−1033.72	52.26	−580.12	±25.14
	VDW	−1199.04	26.39	−663.15	19.52	−475.96	17.3	−59.93	± 3.81	−1221.57	25.13	−681.29	20.41	−499.29	15.89	−40.99	± 3.51
	INT	6481.97	47.3	3864	37.53	2617.97	28.9	0		6382.78	45.06	3842.19	36.59	2540.59	28.95	0	
	GAS	−1025.31	85.11	−900.29	69.15	610.96	54.95	−735.97	±29.20	−518.74	99.45	−905.21	72.79	1007.58	55.93	−621.11	±25.23
PBSA																	
	PB_{SUR}	115.42	1.6	75.53	0.85	49.92	0.61	−10.03	± 0.78	114.2	1.09	73.04	0.94	47.98	0.71	−6.82	± 0.38
	PB_{CAL}	−4197.32	64.52	−2442.25	52.94	−2467.41	46.91	712.33	±31.38	−4783.59	87.86	−2424.97	65.39	−2960.56	49.62	601.94	±24.45
	PB_{SOL}	−4081.91	63.71	−2366.72	52.46	−2417.48	46.81	702.3	±30.95	−4669.39	87.37	−2351.93	64.77	−2912.58	49.35	595.12	±24.47
	PB_{ELE}	−10505.6	29.93	−6543.39	22.8	−3998.45	16.61	36.28	± 9.55	−10463.5	27.9	−6491.08	22.36	−3994.27	16.69	21.82	± 5.89
	PB_{TOT}	−5107.22	49.62	−3267.01	38.47	−1806.53	31.08	−33.68	± 7.81	−5188.13	47.17	−3257.15	38.06	−1905	29.57	−25.98	± 5.23
GBSA																	
	GB_{SUR}	115.42	1.6	75.53	0.85	49.92	0.61	−10.03	± 0.78	114.2	1.09	73.04	0.94	47.98	0.71	−6.82	± 0.38
	GB_{CAL}	−4215.51	64.5	−2438.41	53.76	−2508.14	47.43	731.04	±29.03	−4846.67	87.3	−2446.72	64.46	−3005.17	49.68	605.22	±24.87
	GB_{SOL}	−4100.09	63.77	−2362.88	53.28	−2458.22	47.33	721.01	±28.68	−4732.47	86.86	−2373.68	63.86	−2957.19	49.43	598.4	±24.91
	GB_{ELE}	−10523.8	24.9	−6539.55	19.63	−4039.19	14.75	54.99	± 5.17	−10526.6	25.16	−6512.83	19.08	−4038.88	15.17	25.1	± 3.99
	GB_{TOT}	−5125.4	48.95	−3263.17	37.98	−1847.26	30.4	−14.97	± 4.05	−5251.21	46.26	−3278.9	36.69	−1949.61	29.54	−22.7	± 3.23

Table 4
Decomposition energies (kcal/mol) of ankyrin residues in CD4–ankyrin and CD4–designed ankyrin complexes.

Residue	Ankyrin								Designed ankyrin							
	INT	VDW	ELE	GAS	GB	GBSUR	GBSOL	GBTOT	INT	VDW	ELE	GAS	GB	GBSUR	GBSOL	GBTOT
1	10.4	−0.58	−102.7	−92.85	−8.5	0.98	−7.53	−100.37	12.25	−2.03	−79.74	−69.52	−65.12	1.04	−64.08	−133.6
2	17.22	−2.34	4.56	19.44	−60.72	0.65	−60.07	−40.63	22.9	−3.27	−5.8	13.83	−2.43	0.47	−1.96	11.87
3	23.49	−3.86	−3.88	15.75	−4.67	0.38	−4.29	11.46	13.88	−2.7	−10.64	0.54	−2.12	0.14	−1.98	−1.44
4	13.68	−3.33	−6.19	4.16	−5.58	0.02	−5.56	−1.4	26.23	−2.65	−48.84	−25.25	22.88	0.96	23.84	−1.41
5	26.49	−2.98	−41.22	−17.72	15.54	1.02	16.56	−1.16	26.07	−4.41	−44.19	−22.53	23.51	0.89	24.4	1.87
6	26.23	−3.98	−55.54	−33.3	32.41	0.79	33.2	−0.1	25.09	−7.64	−6.44	11.01	−2.71	0	−2.71	8.3
7	23.62	−7.58	−6.21	9.84	−2.9	0	−2.9	6.94	23.05	−8.04	−13.21	1.8	−3.23	0.02	−3.2	−1.41
8	24.38	−7.4	−13.64	3.34	−2.84	0.02	−2.82	0.52	20.36	−3.59	7.81	24.58	−108.3	0.68	−107.6	−83
9	20.22	−3.58	9.9	26.53	−109.9	0.6	−109.34	−82.81	14.83	−5.13	−5.57	4.14	−2.92	0.04	−2.88	1.26
10	15.09	−5.06	−7.57	2.46	−1.03	0.02	−1	1.45	15.03	−5.88	0.8	9.95	−3.5	0	−3.5	6.45
11	15.17	−5.62	−0.57	8.98	−2.59	0	−2.59	6.39	28.62	−6.42	−213.6	−191.4	67.35	0.53	67.89	−123.5
12	29.38	−4.42	−228.8	−203.9	78.18	0.66	78.84	−125.02	14.79	−2.6	0.52	12.71	−5.43	0.44	−4.99	7.72
13	14.9	−1.77	−2.64	10.49	−3.46	0.54	−2.92	7.57	13.25	−3.12	−8.97	1.16	−4.06	0.1	−3.96	−2.79
14	13.75	−2.2	−11.18	0.37	−2.31	0.15	−2.16	−1.79	30.34	−3.54	−84.02	−57.23	−10.92	0.66	−10.26	−67.49
15	29.93	−3.25	−82.32	−55.64	−11.7	0.62	−11.08	−66.72	18.45	−5.33	83.23	96.35	−145.5	0.23	−145.2	−48.9
16	19.33	−4.19	64.2	79.34	−128.3	0.41	−127.93	−48.59	16.99	−3.26	70.88	84.6	−134.5	0.7	−133.8	−49.21
17	17.45	−2.88	66.29	80.87	−129.7	0.77	−128.97	−48.1	20.52	−3.34	21.99	39.16	−121.6	0.52	−121	−81.86
18	20.74	−4.08	29.8	46.46	−128.9	0.5	−128.36	−81.9	20.91	−6.03	−10.6	4.29	3.19	0	3.19	7.48
19	20.45	−5.54	−9.36	5.55	2.55	0	2.55	8.1	28.51	−5.48	−232.3	−209.3	88.71	0.88	89.59	−119.7
20	28.77	−5.86	−236.9	−214	93.3	0.63	93.93	−120.07	24.33	−1.07	−16.18	7.08	1.13	0.83	1.96	9.04
21	23.99	−1.12	−16.26	6.61	1.23	0.76	1.99	8.6	22.79	−6.07	−16.28	0.45	0.73	0.1	0.83	1.27
22	22.91	−5.71	−17.35	−0.15	1.22	0.11	1.33	1.18	21.1	−8.06	−14.23	−1.19	−4.53	0.19	−4.35	−5.54
23	20.81	−7.95	−15	−2.14	−3.91	0.22	−3.69	−5.83	15.01	−1.35	−6.03	7.62	−4.29	0.69	−3.6	4.02
24	14.97	−1.27	−6.13	7.57	−4.29	0.7	−3.59	3.99	25.46	−1.8	−55.63	−31.97	−8.74	0.84	−7.9	−39.87
25	26.05	−1.81	−54.43	−30.19	−9.61	0.8	−8.81	−39	13.7	−1.14	−6.13	6.43	−2.91	0.43	−2.48	3.95
26	14.2	−1.38	−6.62	6.2	−2.33	0.38	−1.95	4.25	14.36	−2.93	−0.97	10.46	−6.68	0.09	−6.59	3.87
27	14.31	−3.45	−2.03	8.83	−6.01	0.04	−5.97	2.86	16.79	−2.27	54.4	68.92	−119.6	0.51	−119	−50.12
28	17.26	−3.74	43.84	57.35	−107.2	0.46	−106.71	−49.36	19.54	−2.91	−8.76	7.87	0.83	0.11	0.94	8.81
29	20.59	−4.42	−10.29	5.88	2.13	0.04	2.17	8.06	24.97	−0.78	−53.1	−28.91	−11.34	0.98	−10.37	−39.28
30	24.56	−1.57	−60.89	−37.91	−3.96	0.7	−3.26	−41.17	14.14	−2.96	9.56	20.73	−5.86	0.14	−5.72	15.01
31	14.14	−3.05	3.74	14.83	−5.41	0.11	−5.3	9.53	19.39	−2.41	−50.93	−33.95	−5.07	0.57	−4.5	−38.45
32	24.99	−3.23	−68.21	−46.45	42.33	1.07	43.4	−3.06	16.74	−4.38	40.65	53	−98.92	0.22	−98.7	−45.7
33	17.27	−3.62	36.88	50.54	−96.94	0.18	−96.76	−46.22	19.91	0.53	−36.81	−16.38	−14.35	0.98	−13.37	−29.75
34	20.34	−0.43	23.17	43.08	−123.2	1.21	−122.01	−78.93	22.53	−1.09	−10.09	11.35	−3.02	0.72	−2.3	9.05
35	24.45	2.11	−6.8	19.77	−9.25	1.39	−7.87	11.9	12.78	−2.8	−4.91	5.08	−3.04	0.15	−2.89	2.19
36	13.07	−2.03	−4.88	6.16	−1.6	0.2	−1.41	4.75	27.84	−9.14	−218.2	−199.5	82.62	0.36	82.98	−116.6
37	23.58	−4.91	−8.41	10.26	1.61	0.19	1.79	12.05	19.62	−3.33	−37.59	−21.3	−3.09	0.12	−2.97	−24.27
38	19.3	−4.39	−38.74	−23.83	−1.09	0.16	−0.93	−24.76	26.29	−7.41	−10.36	8.53	−1.54	0	−1.54	6.99
39	26.13	−7.95	−10.48	7.7	−1.52	0	−1.52	6.18	22.43	−7.05	−11.49	3.89	−2.27	0	−2.27	1.62
40	22.24	−5.94	−13.13	3.16	−1.07	0.05	−1.02	2.14	39.55	−12.89	−14.49	12.18	0.04	0	0.04	12.22
41	23.74	−7.13	−9.31	7.29	−7.52	0.29	−7.23	0.06	20.41	−10.01	−12.52	−2.12	−2.19	0.02	−2.17	−4.29
42	23.77	−7.04	−7.27	9.47	−2.62	0.09	−2.53	6.94	15.4	−5.6	−6.47	3.33	−1.56	0	−1.56	1.77
43	15.04	−6.25	−0.29	8.5	−2.16	0	−2.16	6.33	15.36	−5.88	−2.53	6.96	−4.04	0	−4.04	2.92
44	19.82	−5.71	−45.58	−31.46	−5.21	0	−5.21	−36.7	15.01	−3.69	−1.26	10.06	−5.15	0.18	−4.98	5.09
45	15.28	−2.9	−8.38	3.99	−5.03	0.35	−4.67	−0.68	29.18	−4.89	7.53	31.82	−7.23	1.03	−6.2	25.61
46	39.51	−5.89	−4.03	29.59	−8.28	0.65	−7.63	21.96	13.48	−2.62	−9.85	1.02	−1.38	0.08	−1.29	−0.28
47	13	−2.36	−9.43	1.21	−1.18	0.15	−1.04	0.17	39.76	−8.12	−2.53	29.11	−6.98	0.33	−6.65	22.46
48	41.02	−6.59	−8.03	26.39	−5.94	0.49	−5.45	20.94	24.8	−2.81	−10.21	11.78	−2.49	0.57	−1.92	9.86
49	24.2	−4.3	−13.92	5.99	−0.24	0.38	0.14	6.13	20.18	−2.26	76.41	94.32	−176.4	0.94	−175.5	−81.13
50	21.09	−2.82	51.64	69.91	−154.1	0.76	−153.36	−83.45	23.82	−5.38	−22.01	−3.57	−0.92	0	−0.91	−4.49
51	23.92	−4.99	−20.75	−1.82	−0.19	0.02	−0.17	−1.99	20.38	−4.77	−14.16	1.45	1.55	0.01	1.56	3.01
52	20.01	−6.09	−14.89	−0.97	1.34	0	1.34	0.37	17.15	−3.61	66.48	80.02	−129.7	0.51	−129.2	−49.19
53	20.02	−4.54	35.66	51.14	−132.7	0.45	−132.23	−81.09	20.1	−4.3	−8.1	7.7	3.04	0.13	3.16	10.87
54	21	−5.01	−9.13	6.86	3.26	0.16	3.42	10.28	22.43	−7.49	−8.61	6.33	0.46	0	0.46	6.79
55	22.65	−7.83	−7.5	7.31	0.27	0	0.27	7.59	22.83	−4.92	−12.04	5.87	−1.99	0.23	−1.76	4.11
56	23.51	−6.32	−13.6	3.59	−0.5	0.04	−0.46	3.13	25.28	−1.16	−99.17	−75.05	67.53	1.28	68.81	−6.24
57	25.49	−2.14	−85.94	−62.59	55.59	1.22	56.82	−5.77	39.51	−5.48	−4.54	29.49	−6.6	0.64	−5.96	23.53
58	25.42	−3.92	−56.82	−35.32	−7.1	0.38	−6.71	−42.03	14.14	−1.23	−6.25	6.67	−3.05	0.42	−2.63	4.04
59	14.23	−1.42	−5.3	7.5	−3.49	0.36	−3.13	4.37	14.38	−3.1	−3.26	8.02	−5.14	0.08	−5.06	2.96
60	14.92	−2.32	−4.69	7.91	−3.69	0.1	−3.59	4.32	17.67	−3.05	79.47	94.09	−142.5	0.63	−141.9	−47.81
61	16.75	−2.82	70.02	83.94	−132.6	0.61	−131.98	−48.05	20.77	−4.57	−4.2	12	−1.37	0.12	−1.25	10.75

Table 4 (*Continued*)

Residue	Ankyrin								Designed ankyrin							
	INT	VDW	ELE	GAS	GB	GBSUR	GBSOL	GBTOT	INT	VDW	ELE	GAS	GB	GBSUR	GBSOL	GBTOT
62	19.74	−3.7	−4.92	11.13	−0.29	0.02	−0.27	10.86	26.12	−3.57	−56.95	−34.4	−5.85	0.43	−5.41	−39.82
63	24.94	−2.26	−62.66	−39.98	−3.81	0.59	−3.22	−43.21	14.14	−2.4	−1.99	9.75	−2.1	0.19	−1.9	7.84
64	14.49	−2.36	−1.85	10.28	−5.91	0.4	−5.52	4.76	23.39	−1.93	−13.67	7.8	−4.44	0.86	−3.58	4.22
65	19.88	−2.22	−6.19	11.47	−3.96	0.6	−3.36	8.11	21.1	−6.15	32.72	47.67	−129.3	0.08	−129.2	−81.53
66	17.39	−3.62	47.62	61.4	−100.8	0.09	−100.67	−39.27	21.07	−2.55	45.9	64.42	−148.4	0.79	−147.6	−83.23
67	15.82	−1.17	−5.69	8.96	−2.43	0.45	−1.98	6.98	19.34	−0.28	−11.75	7.31	1.45	0.56	2.01	9.32
68	23.4	0.09	−8.59	14.9	−5.46	0.79	−4.67	10.23	12.78	−2.85	−5.45	4.49	−1.75	0.05	−1.7	2.79
69	13.16	−2.41	−9.07	1.68	−2.9	0.13	−2.77	−1.09	20.04	−9.09	−11.94	−0.99	0.55	0.07	0.62	−0.37
70	25.33	−5.33	−0.13	19.86	−2.09	0.19	−1.91	17.96	18.88	−4.48	−35.05	−20.65	−4.5	0.02	−4.47	−25.13
71	19.93	−5.23	−36.62	−21.92	−2.47	0.01	−2.46	−24.38	26.44	−8.34	−10.07	8.03	−1.91	0	−1.91	6.12
72	26.6	−7.09	−12.07	7.44	−0.42	0.04	−0.38	7.06	22.95	−6.83	−11.13	4.99	−2.5	0	−2.5	2.49
73	23.29	−6.82	−11.98	4.49	−1.44	0.01	−1.43	3.07	39.64	−12.38	−11.02	16.24	−4.07	0	−4.06	12.17
74	39.94	−11.89	−10.88	17.17	−2.33	0.01	−2.32	14.85	22.75	−7.64	−8.32	6.79	−2.47	0.08	−2.39	4.39
75	23.11	−6.34	−10.83	5.94	−0.58	0.1	−0.48	5.45	14.86	−5.61	−8.78	0.46	−0.71	0	−0.71	−0.24
76	15.32	−5.59	−7.41	2.32	−1.51	0	−1.51	0.81	14.9	−5.29	−2.67	6.94	−2.98	0.01	−2.97	3.97
77	14.99	−5.64	−1.94	7.4	−3.65	0	−3.65	3.75	22.92	−2.47	−2.17	18.28	−4.35	0.82	−3.54	14.74
78	23.17	−2.31	−3.21	17.64	−6.3	0.93	−5.37	12.27	22.67	−3.24	−6.26	13.17	−2.13	0.55	−1.58	11.59
79	23.93	−2.54	−8.51	12.89	−5.78	0.53	−5.25	7.64	13.37	−2.09	−13.11	−1.83	−0.78	0.14	−0.65	−2.48
80	13.5	−2.02	−13.43	−1.95	−2.25	0.18	−2.06	−4.01	41.36	−7.34	−10.25	23.77	−5.23	0.38	−4.85	18.92
81	40.28	−7.69	−10.45	22.14	−3.72	0.36	−3.36	18.78	23.85	−2.44	−11.21	10.2	−2.78	0.54	−2.24	7.96
82	23.94	−3.17	−12.03	8.74	−1.39	0.53	−0.85	7.89	20.92	−3.07	88.6	106.4	−186.7	0.75	−186	−79.53
83	21.01	−3	60.06	78.07	−160.1	0.72	−159.42	−81.34	24.83	−5.81	−22.92	−3.91	−1.44	0	−1.44	−5.35
84	24.73	−5.62	−15.59	3.52	−1.34	0	−1.34	2.18	20.75	−5.39	−14.7	0.66	1.1	0	1.1	1.76
85	15.32	−4.8	−11.98	−1.45	−0.37	0	−0.37	−1.82	20.03	−3.55	87.18	103.7	−184.9	0.68	−184.2	−80.57
86	20.02	−3.64	55.03	71.42	−157.2	0.68	−156.53	−85.11	19.6	−3.06	−9.68	6.85	3.74	0.25	3.99	10.85
87	21.18	−4.51	−9.31	7.36	2.53	0.14	2.67	10.03	22.97	−8.08	−8.52	6.37	0.93	0	0.93	7.3
88	22.8	−7.7	−7.24	7.87	−0.33	0	−0.33	7.54	23.22	−4.1	−13.3	5.82	−0.85	0.39	−0.46	5.36
89	23.02	−4.09	−12.86	6.08	−1.81	0.28	−1.53	4.55	25.32	−1.34	−110.1	−86.1	72.53	1.35	73.88	−12.22
90	25.43	−1.61	−95.96	−72.14	65.07	1.31	66.38	−5.76	16.56	−1.96	−2.15	12.44	−9.54	0.42	−9.12	3.33
91	39.5	−6.11	−8.72	24.67	−3.88	0.57	−3.3	21.37	13.86	−1.23	−7.99	4.64	−2.13	0.41	−1.71	2.92
92	13.49	−1.73	−3.99	7.78	−3.36	0.3	−3.06	4.71	14.65	−3.74	−4.19	6.71	−4.39	0.06	−4.33	2.38
93	14.06	−2.07	−3.85	8.13	−4.13	0.19	−3.94	4.19	17.43	−3.47	89.35	103.3	−153	0.49	−152.5	−49.19
94	16.26	−3.34	70.96	83.88	−130	0.33	−129.62	−45.74	20.32	−4.06	−4.86	11.4	−1.55	0.22	−1.33	10.06
95	20.32	−2.19	−4.31	13.82	−1.16	0.31	−0.85	12.97	24.47	−1.83	−58.42	−35.79	−5.54	0.71	−4.83	−40.62
96	25.38	−1.08	−57.52	−33.23	−7.91	0.93	−6.98	−40.21	14.02	−2.74	−2.55	8.74	−2.82	0.12	−2.69	6.04
97	14.46	−2.4	−4.95	7.11	−1.32	0.11	−1.21	5.9	27.69	−3.88	−75.6	−51.79	−15.76	0.84	−14.92	−66.71
98	27.49	−3.59	−77.09	−53.18	−15.34	0.8	−14.54	−67.72	16.77	−6.28	84.19	94.68	−136.7	0.09	−136.6	−41.96
99	16.88	−4.98	46.91	58.81	−103.5	0.12	−103.34	−44.53	24.55	−1.63	−87.89	−64.97	62.07	1.33	63.4	−1.56
100	25.6	−0.77	−44.23	−19.41	19.86	1.39	21.25	1.84	22.71	1.28	−3.21	20.78	−3.14	1.33	−1.81	18.97
101	22.61	1.37	−2.7	21.28	−3.75	1.26	−2.49	18.79	13.32	−1.81	−7.88	3.63	−2.11	0.29	−1.83	1.8
102	13.16	−1.47	−7.69	4.01	−2.45	0.29	−2.16	1.85	25.88	−4.91	−133.9	−112.9	113	0.4	113.44	0.52
103	24.74	−4.42	−65.84	−45.52	47.1	0.73	47.83	2.3	18.09	−4.64	−47.79	−34.34	−3.74	0.13	−3.61	−37.96
104	18.5	−3.9	−45.94	−31.34	−5.01	0.2	−4.81	−36.15	15	−5.51	−8.22	1.26	−0.76	0	−0.76	0.5
105	14.92	−4.85	−8.54	1.52	−1.1	0	−1.1	0.43	22.85	−5.94	−9.46	7.45	−1.6	0.49	−1.11	6.34
106	22.97	−4.68	−7.89	10.4	−3.28	0.48	−2.8	7.6	16.92	−5.37	94.31	105.9	−154.4	0.31	−154.1	−48.19
107	17.11	−4.81	65.06	77.37	−126.2	0.51	−125.71	−48.34	23.34	−5.7	−20.55	−2.91	−4.12	0.11	−4	−6.91
108	23.52	−5.12	−21.77	−3.37	−3.57	0.07	−3.5	−6.86	16.61	−5.78	−14.22	−3.39	−2.58	0	−2.58	−5.97
109	16.69	−5.81	−14.2	−3.32	−3.89	0	−3.89	−7.21	23.26	−2.41	−14.29	6.57	−4.69	0.66	−4.03	2.53
110	24.36	−1.57	−8.29	14.51	−4.32	0.67	−3.65	10.86	17.16	−2.8	99.72	114.1	−166.5	0.72	−165.8	−51.69
111	13.49	−1.14	−12.66	−0.31	−3.14	0.51	−2.63	−2.94	23.92	−1.45	−6.93	15.54	−8.42	1.06	−7.36	8.18
112	26.04	−2.72	−55.81	−32.49	−7.7	0.59	−7.11	−39.6	13.49	−1.21	−9.06	3.22	−1.63	0.4	−1.23	1.99
113	13.09	−0.86	−7.27	4.96	−3.24	0.4	−2.84	2.12	26.72	−4.16	−61.78	−39.22	−6.51	0.3	−6.21	−45.43
114	25.96	−4.44	−62.5	−40.98	−5.3	0.25	−5.05	−46.04	20.17	−2.16	99.74	117.7	−198	0.9	−197.1	−79.33
115	19.49	−1.71	69.24	87.02	−168.1	0.99	−167.09	−80.07	17.5	−2.64	137.1	152	−196.6	0.76	−195.9	−43.91
116	17	−2.92	105	119.1	−163.8	0.73	−163.02	−43.94	23.5	−5.52	−10.48	7.5	−2.34	0.01	−2.33	5.17
117	22.41	−6.15	−12.19	4.07	−0.44	0.01	−0.43	3.64	15.05	−4.71	−8.08	2.26	−2.81	0.07	−2.74	−0.48
118	15.07	−4.74	−12.11	−1.79	0.3	0.04	0.34	−1.45	19.77	−2.16	108.7	126.3	−209	0.9	−208.1	−81.81
119	20.61	−2.81	52.78	70.58	−152.1	0.7	−151.44	−80.86	23.85	−0.95	−8.51	14.4	−5.43	0.59	−4.84	9.56
120	23.68	−3.58	−11.03	9.08	−2.54	0.45	−2.08	7	22.97	−5.6	−7.32	10.05	−2.65	0.03	−2.62	7.43
121	21.95	−6.5	−13.46	2	−3.25	0.02	−3.24	−1.24	24.63	−2.22	69.16	91.58	−209.4	1.35	−208	−116.5

hydrogen bond with Phe43. From atomistic dynamic simulations of ankyrin–CD4 binding complexes, explicit analysis of hydrogen bond properties, such as donor–acceptor assignments and hydrogen bond occupancies, was performed for 1 ns trajectories. The criteria for an intermolecular hydrogen bond in AMBER were used with a donor–acceptor distance of <3.0 Å. However, several hydrogen bonds from docking structures were lost during the molecular dynamics simulations. Stable hydrogen bonds that were observed and a list of hydrogen bonds occupations, distances, and angles are shown in Table 2. The designed ankyrin occupied 88% of a stable NH–O hydrogen bond between Gln40:CD4 and Asp98:ankyrin, and 40% of an O–NH hydrogen bond between Glu87:CD4 and Lys99:ankyrin. In contrast, weaker hydrogen bond formations in the CD4–ankyrin template were found, with approximately 40% occupation of O–NH between Gly9:CD4 and Arg23:ankyrin, of O–OH between Asp10:CD4 and Tyr52:ankyrin, and of O–OH between Leu100:CD4 and Tyr46:ankyrin. This indicates that the CD4–designed ankyrin is more stabilised with hydrogen bonding, while the CD4–ankyrin control forms dynamic hydrogen bonds, implying weaker interactions for CD4 binding.

3.3. Binding scaffold, free energies and decomposed energy

To elucidate the binding interactions between CD4 and ankyrin, conformational data that were generated from the molecular dynamics simulations were used to conduct post-trajectory free energy analysis using the MM-PBSA and MM-GBSA approach. This method has been used in a number of recent studies to investigate protein–protein interactions, and the approach is generally well correlated with experimental results [45–47]. The free energy was expressed as the difference of the free energy of the complex with that of the receptor and ligand, averaged over a number of snapshots collected from the trajectory. The binding free energies for each of the contributing terms are shown in Table 3. An inspection of the free energy components of the complex revealed that the electrostatic component of the free energy of binding contributes unfavourably to binding ($\Delta G > 0$). However, the data showed that the hydrophobic contribution of the solvation free energy in the PB/GB calculation (PBsur/GBsur) systematically contributes favourably ($\Delta G < 0$), as expected, as the formation of complexes reduces solvent-accessible surface area. Relevant mutated residues of the designed ankyrin governing intermolecular interactions in the binding complex are shown in Fig. 4. In the case of the CD4–ankyrin complex, the binding free energy is sensitive to the calculation method, and the contribution from solvation in PBSA is almost 20 kcal/mol, for example, 702, 721 kcal/mol for PBSA, GBSA, respectively, lower than GBSA. However, both methods yielded almost the same solvation energy as in the CD4–designed ankyrin complex. According to the previous report of DARPin 23.2–CD4 binding complex [16], GBSA approach was more related to this regard, with a ΔG of −22.7 and −15.0 kcal/mol for the CD4–designed ankyrin and the CD4-control, respectively.

To clarify the roles of crucial residues in the binding interface, the decomposition energy of the potential ankyrin repeat protein–CD4 complex was calculated and compared with the ankyrin control. The decomposition energy analysis indicated that not only the mutated residues at the binding site (Fig. 5) governed the interaction, but their neighbours were also involved. For example, the binding energy of residue number 65 apparently decreases from 8.11 kcal/mol in the ankyrin template to −81.53 kcal/mol in the designed ankyrin (Table 4). The binding free energies of residues in the binding region are strongly improved in the designed ankyrin molecule (Fig. 5 and Table 4). A negative value indicates a favourable interaction. Three loop interactions between ankyrin and CD4 in the amino acid regions at the positions 31–36, 65–69, and 97–102 can be clearly seen in Fig. 5; the summations of GBTOT in the three selected regions are −219, −153, −88 kcal/mol, respectively. Further improve the affinity of DARPin 23.2 for CD4 could be at the three plausible residues; Leu34, Val67, and Phe102 on three relevant loops of ankyrin governing CD4–ankyrin interactions. Compared to the control ankyrin, the binding interaction of the designed ankyrin was significantly improved in the first and second region, which might lead to improved binding activity with CD4.

4. Conclusion

Our protein docking and molecular dynamics simulations of CD4–ankyrins formerly reported by Schweizer et al. reveal the associated residues involved in binding. Conformational changes during the dynamics simulations of docking structure display a loss of hydrogen bonds, which implies stability in the complexes. The binding free energy determined by the MM-GB method was correlated with the experimental activity of the designed ankyrin to a higher degree than that from the MM-PBSA approach. Intermolecular interactions at the CD4 interface ranging from non-specific to specific hydrogen bonds and hydrophobic contacts surrounding the Phe43 cavity are the key elements for precise calculations with the CD4–gp120 complex. The decomposition energy of the dynamics structures provided molecular information for three plausible residues of DARPin 23.2 i.e. Leu34, Val67, and Phe102 on three relevant loops, which were significant in DARPin 23.2–CD4 interactions. They will be the candidates for further improvement of the DARPin 23.2 binding affinity with CD4.

According to the *in vivo* study, CD4-specific DARPin was not totally pronounced the negative effect on SHIV plasma levels after a single intravenous injection into SHIV-infected rhesus macaques [25]. Several techniques for enhancing the binding activity either by ribosome display or phage display, however, are complicated. The information obtained from dynamics simulations will be useful for future protein engineering by rational design. The improvement of its binding affinity by performing amino acid substitution at these specific residues will generate the candidate molecules with higher binding force against CD4 for preventing the viral attachment *in vivo*.

Acknowledgements

The work was supported by the Thailand Research Fund (TRF), the National Centre for Genetic Engineering and Biotechnology (BIOTEC) of the National Science and Technology Development Agency, the Thailand Centre of Excellence in Physics (ThEP), the Centre for Innovation in Chemistry(PERCH-CIC), and the National Research University Project under Thailand's Office of the Higher Education Commission, Thailand.

References

[1] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, Pairwise solute descreening of solute charges from a dielectric medium, Chem. Phys. Lett. 246 (1995) 122–129.
[2] R.A. Kramer, M.D. Schaber, A.M. Skalka, K. Ganguly, F. Wong-Staal, E.P. Reddy, HTLV-III gag protein is processed in yeast cells by the virus pol-protease, Science 231 (1986) 1580–1584.
[3] L. Douali, D. Villemin, D. Cherqaoui, Neural networks: accurate nonlinear QSAR model for HEPT derivatives, J. Chem. Inf. Comput. Sci. 43 (2003) 1200–1207.
[4] L. Douali, D. Villemin, D. Cheraqoui, Comparative QSAR based on neural networks for the anti-HIV activity of HEPT derivatives, Curr. Pharm. Des. 9 (2003) 1817–1826.
[5] M. Moore, G. Dreyer, Substrate-based inhibitors of HIV-1 protease, Perspect. Drug Discov. Des. 1 (1993) 85–108.
[6] J.W. Erickson, Design and structure of symmetry-based inhibitors of HIV-1 protease, Perspect. Drug Discov. Des. 1 (1993) 109–128.
[7] A.G. Tomasselli, R.L. Heinrikson, Targeting the HIV-protease in AIDS therapy: a current clinical perspective, Biochim. Biophys. Acta 1477 (2000) 189–214.

[8] G.V. De Lucca, S. Erickson-Viitanen, P.Y.S. Lam, Cyclic HIV protease inhibitors capable of displacing the active site structural water molecule, Drug Discov. Today 2 (1997) 6–18.
[9] D.J. Kempf, H.L. Sham, HIV protease inhibitors, Curr. Pharm. Des. 2 (1996) 225–246.
[10] N.A. Roberts, J.A. Martin, D. Kinchington, A.V. Broadhurst, J.C. Craig, I.B. Duncan, et al., Rational design of peptide-based HIV proteinase inhibitors, Science 248 (1990) 358–361.
[11] O. Arusksakunwong, S. Promsri, K. Wittayanarakul, P. Nimmanpipug, V.S. Lee, A. Wijitkosoom, et al., Current development on HIV-1 protease inhibitors, Curr. Comput. Aided Drug Des. 3 (2007) 201–213.
[12] J. Cocohoba, B.J. Dong, Raltegravir: the first HIV integrase inhibitor, Clin. Ther. 30 (2008) 1747–1765.
[13] M. Markowitz, J.O. Morales-Ramirez, B.Y. Nguyen, C.M. Kovacs, R.T. Steigbigel, D.A. Cooper, et al., Antiretroviral activity, pharmacokinetics, and tolerability of MK-0518, a novel inhibitor of HIV-1 integrase, dosed as monotherapy for 10 days in treatment-naive HIV-1-infected individuals, J. Acquir. Immune Defic. Syndr. 43 (2006) 509–515.
[14] J. Cohen, Retrovirus meeting. Novel attacks on HIV move closer to reality, Science 311 (2006) 943.
[15] H. Wu, D.G. Myszka, S.W. Tendian, C.G. Brouillette, R.W. Sweet, I.M. Chaiken, et al., Kinetic and structural analysis of mutant CD4 receptors that are defective in HIV gp120 binding, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 15030–15035.
[16] A. Schweizer, P. Rusert, L. Berlinger, C.R. Ruprecht, A. Mann, S. Corthésy, et al., CD4-specific designed ankyrin repeat proteins are novel potent HIV entry inhibitors with unique characteristics, PLoS Pathogens 4 (2008) e1000109.
[17] H.K. Binz, M.T. Stumpp, P. Forrer, P. Amstutz, A. Plückthun, Designing repeat proteins: well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins, J. Mol. Biol. 332 (2003) 489–503.
[18] H.K. Binz, P. Amstutz, A. Plückthun, Engineering novel binding proteins from nonimmunoglobulin domains, Nat. Biotechnol. 23 (2005) 1257–1268.
[19] S.G. Sedgwick, S.J. Smerdon, The ankyrin repeat: a diversity of interactions on a common structural framework, Trends Biochem. Sci. 24 (1999) 311–316.
[20] S. Krzywda, A.M. Brzozowski, H. Higashitsuji, J. Fujita, R. Welchman, S. Dawson, et al., The crystal structure of gankyrin, an oncoprotein found in complexes with cyclin-dependent kinase 4, a 19 S proteasomal ATPase regulator, and the tumor suppressors Rb and p53, J. Biol. Chem. 279 (2004) 1541–1545.
[21] C. Zahnd, E. Wyler, J.M. Schwenk, D. Steiner, M.C. Lawrence, N.M. McKern, et al., A designed ankyrin repeat protein evolved to picomolar affinity to Her2, J. Mol. Biol. 369 (2007) 1015–1028.
[22] M.T. Stumpp, H.K. Binz, P. Amstutz, DARPins: a new generation of protein therapeutics, Drug Discov. Today 13 (2008) 695–701.
[23] M. Gebauer, A. Skerra, Engineered protein scaffolds as next-generation antibody therapeutics, Curr. Opin. Chem. Biol. 13 (2009) 245–255.
[24] L. Chen, Y. Do Kwon, T. Zhou, X. Wu, S. O'Dell, L. Cavacini, et al., Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120, Science 326 (2009) 1123–1127.
[25] P. Pugach, A. Krarup, A. Gettie, M. Kuroda, J. Blanchard, P. Michael Jr., et al., In vivo binding and retention of CD4-specific DARPin 57.2 in macaques, PLoS One 5 (2010) e12455.
[26] T. Lengauer, M. Rarey, Computational methods for biomolecular docking, Curr. Opin. Struct. Biol. 6 (1996) 402–406.
[27] I. Halperin, B. Ma, H. Wolfson, R. Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, Proteins 47 (2002) 409–443.
[28] E. Katchalski-Katzir, I. Shariv, M. Eisenstein, A.A. Friesem, C. Aflalo, I.A. Vakser, Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, Proc. Natl. Acad. Sci. U.S.A. 89 (1992) 2195–2199.
[29] J.G. Mandell, V.A. Roberts, M.E. Pique, V. Kotlovyi, J.C. Mitchell, E. Nelson, et al., Protein docking using continuum electrostatics and geometric fit, Protein Eng. 14 (2001) 105–113.
[30] I.A. Vakser, Protein docking for low-resolution structures, Protein Eng. 8 (1995) 371–377.
[31] H.A. Gabb, R.M. Jackson, M.J.E. Sternberg, Modelling protein docking using shape complementarity, electrostatics and biochemical information, J. Mol. Biol. 272 (1997) 106–120.
[32] R. Chen, L. Li, Z. Weng, ZDOCK: an initial-stage protein-docking algorithm, Proteins 52 (2003) 80–87.
[33] R. Chen, Z. Weng, A novel shape complementarity scoring function for protein–protein docking, Proteins 51 (2003) 397–408.
[34] M.Y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures, Protein Sci. 15 (2006) 2507–2524.
[35] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with three-dimensional profiles, Nature 356 (1992) 83–85.
[36] T.P.J. Garrett, J. Wang, Y. Yan, J. Liu, S.C. Harrison, Refinement and analysis of the structure of the first two domains of human CD4, J. Mol. Biol. 234 (1993) 763–778.
[37] D.A. Case, T.E. Cheatham Iii, T. Darden, H. Gohlke, R. Luo, K.M. Merz Jr., et al., The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[38] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[39] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, et al., Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, Acc. Chem. Res. 33 (2000) 889–897.
[40] G.D. Hawkins, C.J. Cramer, D.G. Truhlar, Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium, J. Phys. Chem. 100 (1996) 19824–19839.
[41] V. Tsui, D.A. Case, Theory and applications of the generalized Born solvation model in macromolecular simulations, Biopolymers 56 (2000) 275–291.
[42] U. Moebius, L.K. Clayton, S. Abraham, S.C. Harrison, E.L. Reinherz, The human immunodeficiency virus gp120 binding site on CD4: delineation by quantitative equilibrium and kinetic binding studies of mutants in conjunction with a high-resolution CD4 atomic structure, J. Exp. Med. 176 (1992) 507–517.
[43] S.C. Harrison, CD4: structure and interactions of an immunoglobulin superfamily adhesion molecule, Acc. Chem. Res. 26 (1993) 449–453.
[44] S.T.D. Hsu, A.M.J.J. Bonvin, Atomic insight into the CD4 binding-induced conformational changes in HIV-1 gp120, Proteins 55 (2004) 582–593.
[45] H. Gohlke, D.A. Case, Converging free energy estimates: MM-PB(GB)SA studies on the protein–protein complex Ras–Raf, J. Comput. Chem. 25 (2004) 238–250.
[46] N. Basdevant, H. Weinstein, M. Ceruso, Thermodynamic basis for promiscuity and selectivity in protein–protein interactions: PDZ domains, a case study, J. Am. Chem. Soc. 128 (2006) 12766–12777.
[47] R. Grünberg, M. Nilges, J. Leckner, Flexibility and conformational entropy in protein–protein binding, Structure 14 (2006) 683–693.