Bioorganic & Medicinal Chemistry Letters 19 (2009) 6087–6091

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

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

A dynamic target-based pharmacophoric model mapping the CD4 binding site on HIV-1 gp120 to identify new inhibitors of gp120–CD4 protein–protein interactions

Fabiana Caporuscio [a,†], Andrea Tafi [a], Emmanuel González [b], Fabrizio Manetti [a], José A. Esté [b], Maurizio Botta [a,*]

[a] Dipartimento Farmaco Chimico Tecnologico, Università degli Studi di Siena, Via Alcide de Gasperi, 2, 53100 Siena, Italy
[b] Laboratori de Retrovirologia irsiCaixa, Hospital Universitari Germans Trias i Pujol, Universitat Autònoma de Barcelona, Ctra. Del Canyet s/n, 08916 Badalona, Spain

ARTICLE INFO

Article history:
Received 28 July 2009
Revised 6 September 2009
Accepted 9 September 2009
Available online 13 September 2009

Keywords:
Protein–protein interactions
gp120–CD4 interactions
Phe43 cavity
Structure-based drug design
Pharmacophore modeling

ABSTRACT

A dynamic target-based pharmacophoric model mapping the CD4 binding site on HIV-1 gp120 was built and used to identify new hits able to inhibit gp120–CD4 protein–protein interactions. Two compounds showed micromolar inhibition of HIV-1 replication in cells attributable to an interference with the entry step of infection, by direct interaction with gp120. Inactivity of compounds toward a M475I strain suggested specific contacts with the Phe43 cavity of gp120.

Molecular events involved in HIV invasion and life cycle are accompanied by the formation of protein–protein complexes. A novel strategy to inhibit viral replication is based on the disruption of viral homomeric complexes by peptides or small molecules.[1] Moreover, taking into account the great number of host proteins that interact with viral proteins (as an example, the multi-tiered process of HIV entry into host cells involves critical protein–protein interactions),[2] blocking protein–protein interactions seems to be an even more compelling strategy.

In particular, in the field of HIV-1 infection, gp120–CD4, gp120–CCR5, and gp120–CXCR4 interfaces have been proposed and studied as antiviral targets, thus leading to the identification of a number of entry inhibitors.[3]

Although most of the potent gp120–CD4 inhibitors identified to date are proteins or peptides,[4] recent publications and patent disclosures indicate that there has been an increased effort to identify small-molecule gp120 inhibitors able to block gp120–CD4 interactions.[5] BMS-378806 (BMS-806), BMS-4880434, and #155, discovered through a cell-based screening assay, are potent (nanomolar range) small-molecule inhibitors that prevent the binding of gp120 to CD4 receptors.[6–8] NBD-556 and NBD-557, identified by an HIV syncytium formation assay on a drug-like small-molecule chemical library of 33,000 compounds, are inhibitors that block gp120–CD4 interactions, exhibit single digit micromolar potency against selected HIV-1 laboratory strains, and possess minimal cytotoxicity.[9]

The hydrophobic pocket of gp120, capped by CD4 Phe43, has been suggested as a good binding site for small molecules that block gp120–CD4 interactions either through a direct steric effect or through an allosteric mechanism. In fact, it has been proposed as the target of small-molecule entry inhibitors such as NBD-556 and similar analogues[10] and BMS-806, #155 and BMS-488043, although for the second series of compounds there is debate on their competitive[11] or non-competitive[12,13] mechanism of action. Mutation data suggest that analogues of BMS-806 interact with a subset of gp120 residues lining the Phe43 cavity. On the other hand, the same residues enclose a deep and narrow hydrophobic channel in the crystal structure of an unliganded simian immunodeficiency virus gp120 and BMS-806 has been suggested to interact with this channel, stabilizing the unliganded conformation.[14]

Moreover, the Phe43 cavity is highly conserved and has been hypothesized to be a site less prone to resistance-conferring mutations.[15] On the basis of such considerations, the Phe43 cavity was chosen as the target of our computational simulations, thus taking

* Corresponding author. Tel.: +39 0577 234306; fax: +39 0577 234333.
E-mail address: botta@unisi.it (M. Botta).
† Present address: Dipartimento di Scienze Farmaceutiche, Università degli Studi di Modena e Reggio Emilia, Via Giuseppe Campi, 183, 41125 Modena, Italy.

0960-894X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2009.09.029

up the challenge of rationally designing inhibitors binding to the gp120 Phe43 cavity.[16]

In fact, a computational protocol based on molecular dynamics (to account for protein flexibility), pharmacophoric model generation, and molecular docking, was applied to identify ligands targeting the Phe43 cavity and thus interfering with gp120–CD4 protein–protein interactions. Finally, the computational protocol resulted in novel chemical scaffolds able to inhibit HIV-1 replication in cell culture.

A 1 ns molecular dynamics (MD) simulation in the NPT ensemble with the AMBER 9 Package[17] (Amber2003[18] force field) was performed on the gp120 core structure (G chain), extracted from the crystal structure of the complex with CD4 and X5 antibody (pdb code 2B4C).[19] The system was equilibrated in a stepwise fashion by minimizing the solvent and then the whole system with a 10 kcal mol^{-1} $Å^{-2}$ force constant to constrain alpha carbons. Finally, a position restrained MD in the NVT ensemble was performed before the NPT run (see Supplementary data). Two thousand protein snapshots were saved and clustered (with the kclust tool of MMTSB)[20] on the basis of the root mean square deviation of a shell of gp120 residues within 10 Å from CD4 residues. Five frames, representing different configurations of the selected residues through the simulation, were selected (see Supplementary data). They accounted for the inherent flexibility of the uncomplexed protein under the influence of explicit waters.

Next, GRID[21] molecular interaction fields (MIFs) were computed on the above mentioned shell of residues for each selected frame to identify local minima of the resulting energy maps. DRY, N1 (neutral flat NH, e.g., amide), and O (sp^2 carbonyl oxygen) probes were chosen to mimic hydrophobic interactions, hydrogen bond donors and acceptors, respectively.

To reduce at minimum the arbitrariness in choosing probe minima to be used as a minimal recognition structure useful for pharmacophore building, the five selected frames and MIF minima for each probe were then overlaid and binding regions conserved over the course of MD were searched for. To be considered as conserved, the same binding region of at least two out of the five frames was required to show a minimum of favorable interaction with the same probe (with a 3 Å cutoff distance). As a result, four clusters of minima for the DRY probe (black circles 1, 2, 3, 4 wrapping pink spheres, Fig. 1a), one for the N1 probe (black circle 5 wrapping blue spheres) and three for the O probe (black circles 6, 7, 8 wrapping red spheres), were found as a consequence of the presence of conserved interaction zones. Moreover, the distribution of the DRY minima defined a hydrophobic channel into the Phe43 cavity (Fig. 1b), as previously observed in the crystal structure of gp120 core.[22]

The hydrophobic regions 1, 2, and 3, and the hydrogen bond acceptor region 8 were selected for the generation of the pharmacophoric model. The hydrophobic region 1 was accommodated between the alkyl side chains of CD4 Gln40 and Leu44. The hydrophobic region 2 was located in a pocket near Val430, a residue contacting CD4 Arg59 and able to interact with compounds such as BMS-806 and #155,[11] and NBD-556 analogues.[10] The hydrophobic region 3 was inserted between Met475 (a mutationally critical residue for the interaction with BMS-806 and NBD-556 analogues) and the zone where CD4 Phe43 is accommodated. The hydrogen bond acceptor region 8 was embedded between the hydroxy groups of Tyr384 and Ser375. Ser375 impedes the binding of BMS-806 when mutated to a tryptophan. Moreover, in a recent docking study,[23] BMS-806 indole NH has been proposed as a hydrogen bond donor to Ser375 OH, while Ser375 anchors a water molecule in the X-ray crystal structure of the gp120–CD4 complex.[22] Despite both the donor and acceptor nature of this site, many MIF minima retrieved in this region were derived from the O probe, thus suggesting us to consider it as a hydrogen bond acceptor binding site.

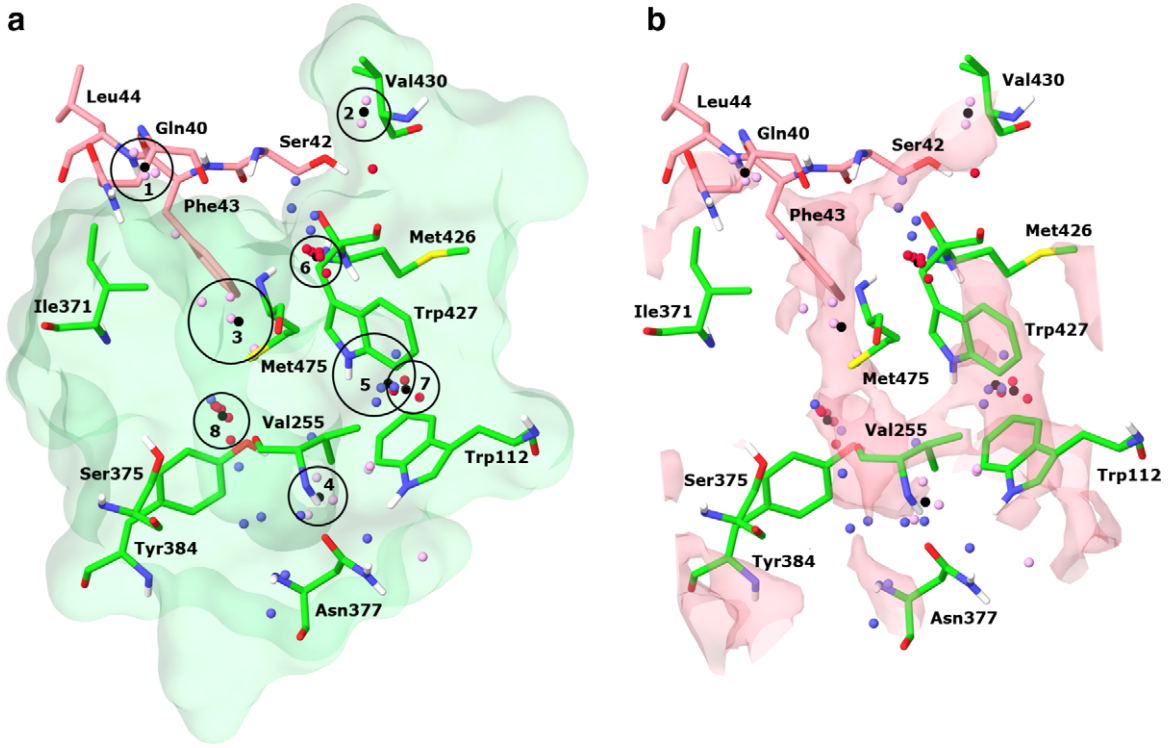


Figure 1. GRID MIF minima into and around the Phe43 cavity for the three probes (DRY, pink; N1, blue; O, red; centroids of each cluster, black) are shown together with selected CD4 (pink) and gp120 residues of one of the five frames extracted from the MD trajectory (green). (a) The molecular surface of the Phe43 cavity (green) and clusters of minima (black circled and numbered). (b) The hydrophobic surface (pink) displayed at an isovalue of −0.3 kcal/mol.

The hydrophobic region 4, deeply embedded at the bottom of the cavity, was rejected to avoid the construction of a pharmacophore with high interfeature distances and thus able to identify only medium or large size compounds. Polar binding regions close to Trp427 and Met426 (5, 6, and 7, Fig. 1a) flanked the entrance of or are embedded in a water channel and were discarded in order to generate a simpler pharmacophoric model. In fact, it is generally assumed that pharmacophores should consist of at least three recognition points (pharmacophoric features) to have selective database queries.[24] Still lacking structural data on known inhibitors able to target the Phe43 cavity, we decided to generate a non-restrictive model (with a limited number of features) and refine the results by docking calculations and visual inspection of binding modes. For each cluster of minima, a geometric centroid was defined and used as the center of a CATALYST[25] pharmacophoric spherical feature with a 1.5 Å radius (tolerance). The tolerance value was chosen in such a way to include all the identified MIF minima of each cluster into the pharmacophoric feature. Moreover, 10 excluded volumes, corresponding to some of the residues defining the boundaries of the cavity (Gln105, Glu370, Ile371, Tyr384, Trp427, Val430, Gly473, Asp474, Met475, Arg476), were incorporated into the resulting pharmacophore to discard compounds with steric clashes with the protein. Coordinates of all the heavy atoms of each selected residue, averaged on the five frames, contributed to calculate a geometric centroid where exclusion volume spheres were centered. The hydrogen bond acceptor feature corresponding to the cluster of minima of the O probe (8, Fig. 1a) was defined as a vectorial feature pointing toward Tyr384 OH group, although the geometric centroid of the cluster was equidistant from Ser375 and Tyr384 OH. Even though Ser375 OH and Tyr384 OH of the gp120 crystal structure were involved in hydrogen bonds with Glu370 oxygens, Tyr384 was chosen instead of Ser375 because Ser375 OH was always involved in a hydrogen bond with Thr257 OH during the course of MD, while Tyr384 OH was free to interact with a putative ligand group (see Supplementary data). As a result, a structure-based, four feature pharmacophore with three hydrophobics (HYD1-3), one hydrogen bond acceptor (HBA), and 10 excluded volumes, was obtained (Fig. 2).

The pharmacophore was then used as a three-dimensional query to perform a virtual screening of a database (the Asinex Gold Collection)[26] containing about 200,000 commercially available compounds. Moreover, a stepwise filtering protocol was applied to select a limited number of entries. In particular, compounds with a fit value lower than 1.5, with more than 10 rotatable bonds, and with more than one chiral center, were all removed from the hit list, as well as compounds with poor absorption and permeation, as calculated on the basis of a modified Lipinski's rule-of-five. On the basis of literature reports suggesting that ADME/Tox parameters for small-molecule inhibitors of protein–protein interactions should be modified,[27] Lipinski's rule-of-five[28] was slightly tuned increasing the molecular weight cutoff to 550. A Log *P* cutoff of 6 was allowed to take into account the hydrophobic nature of the Phe43 cavity.

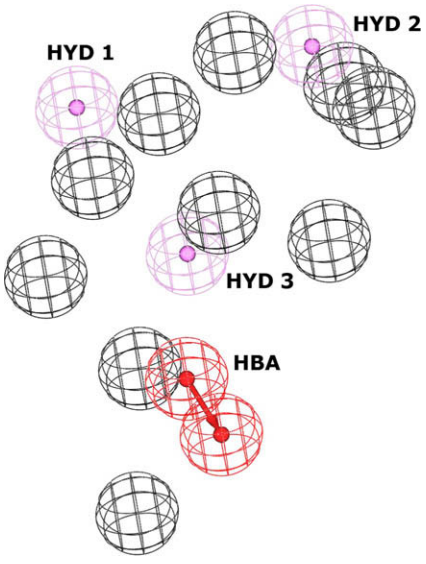


Figure 2. Graphical representation of the pharmacophoric model. Features are color coded: pink for hydrophobics (HYD1, HYD2 and HYD3), red for hydrogen bond acceptors (HBA), and black for excluded volumes.

As a result, a total number of 729 compounds, mapping the pharmacophore and satisfying drug-like requirements, were selected and submitted to a consensus docking protocol to be prioritized. Compounds were docked with both the Chemscore and Goldscore functions available in GOLD[29] into a refined gp120 core structure (pdb code 2B4C),[19] by focusing calculations on the Phe43 cavity and its surrounding residues (see Supplementary data). Compounds were ranked through a consensus protocol, following a rank-by-rank approach.[30] Next, the binding mode of the top 10% of ranked hits was visually inspected, finding most compounds allocated into the Phe43 cavity, while a few of them penetrated in the water channel beyond Met426 and Trp427. This result suggested that the Phe43 cavity was well coded by the pharmacophoric features, which represented possible small-molecule interaction sites, and by the excluded volumes, which avoided the selection of molecules with steric clashes.

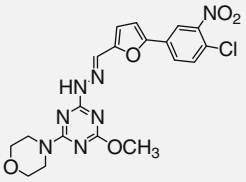
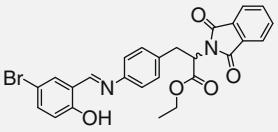
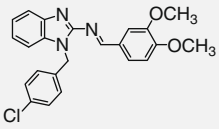
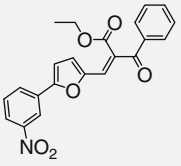
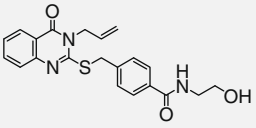
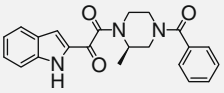
Hit compounds were prioritized by taking into account their binding mode, the match between the docking pose and the binding mode suggested by the pharmacophore, the overall match between the binding modes proposed by the two scoring functions, the average rank, and the commercial availability. Finally, five entries were purchased (Table 1).

Selected compounds were evaluated for their ability to inhibit the HIV-induced cytopathic effect in a human lymphocyte MT-4 cell culture infected with HIV-1 NL4-3 (wt), by the MT-4/MTT-assay.[31,32] Anti-HIV-1 activity and cytotoxicity were compared to those of #155, used as the reference drug (Table 1). Compounds **2** and **4** showed micromolar activity (EC_{50} = 22 and 9 μM, respectively) at subtoxic concentrations. Although their activity was significantly lower than that of #155 and its analogues, the two compounds had novel chemical scaffolds, never tested among small molecules targeting gp120 and able to inhibit HIV-1 replication in cell culture. Moreover, to assess their mechanism of action, they were also tested on MT-4 cells infected with a #155-resistant HIV-1 NL4-3 strain that contained the M475I mutation in the gp120 coding sequence. Compounds resulted completely inactive, thus suggesting a mechanism of action similar to that of #155 and its analogues. Inactivity of **2** and **4** toward the M475I mutant strongly suggested that such compounds are able to target the Phe43 cavity, residue 475 being part of this pocket.

Time of drug addition experiments[32] were also performed to identify the time/site of interaction of **4** in comparison to #155, AMD3100 (a CXCR4 antagonist), the gp41 fusion inhibitors T20 and C34, and a reverse transcriptase inhibitor (AZT). These results showed that, similarly to #155, **4** was active if already present at the time of infection and lost or began to lose activity if added any time post-infection (Fig. 3), thus confirming the initial hypothesis that it acted as an early inhibitor of the gp120–CD4 interaction.

Analysis of the binding mode of **4** showed steric and electronic complementarity between the ligand and the Phe43 cavity (Fig. 4c). In fact, the aromatic ring of the benzoyl moiety of **4** matched HYD1 and was surrounded by hydrophobic residues

Table 1
HIV-1 inhibitory activity and cytotoxicity of compounds **1**–**5** and #155[a]

Cmpd	Asinex BAS entry	Structure	EC_{50} wt[b]	EC_{50} M475I[c]	CC_{50}
1	0324372		>54		>54
2	0392821		22	>48	>48
3	0619235		>62		>62
4	0775946		9	>64	>64
5	3352720		>63		>63
#155			0.024	>2.66	>2.66

[a] Compounds tested at the maximum concentration of 25 μg/mL. EC_{50}: Effective concentration 50 or needed concentration to inhibit 50% HIV-induced cell death evaluated with the MTT method in MT-4 cells. CC_{50}: Cytotoxic concentration 50 or needed concentration to induce 50% death of non-infected cells evaluated with the MTT method in MT-4 cells.
[b] Toward NL4-3 wt (μM).
[c] Toward the Met475Ile mutant, #155-resistant NL4-3 (μM).

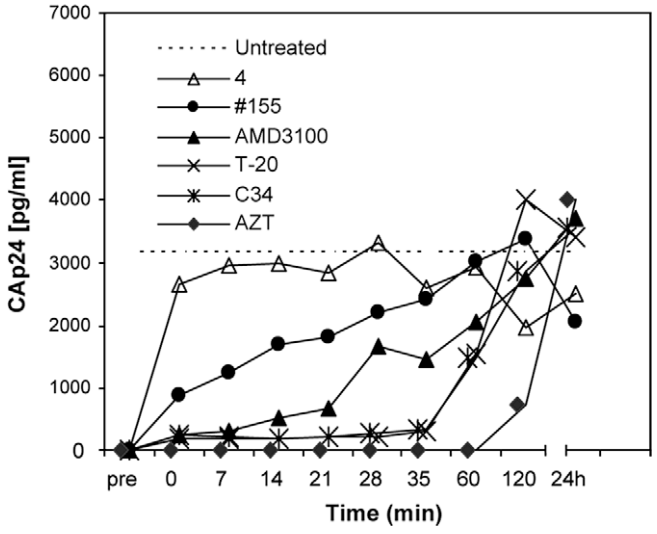


Figure 3. Time of drug addition in MT-4 cells of **4** in comparison to #155, AMD3100, T20, C34 and AZT. Virus growth was evaluated by measuring capsid p24 antigen (CAp24) in the supernatant of infected cells at 30 h post-infection.

(Ile371 and Gly473), while the ester ethyl chain mapped HYD2 and interacted with Val430. The furan ring corresponded to HYD3 and made contacts with Trp427, Gly473 and the alkyl chain of Glu370, while the nitro group accepted a hydrogen bond from the hydroxy group of Tyr384. The *p*-nitro aromatic ring was surrounded by hydrophobic residues (Val255, Trp427 and Met475). The preferential pose of **2** into the Phe43 cavity showed a surprising match between the compound moieties and CD4 residues Ser42, Phe43, and Leu44 (Fig. 4a and b, Supplementary data).

In conclusion, in silico screening of a commercially available compound library using a dynamic target-based pharmacophore as a search query led to the identification of two novel chemical scaffolds, different from all known gp120–CD4 inhibitors, able to inhibit HIV-1 replication in cell culture by interacting with gp120.

This study has reported the first pharmacophoric model based on the structure of gp120, taking up the challenge of rationally designing inhibitors targeting the Phe43 cavity.[16]

Time of addition experiments showed that **4** acted as an early inhibitor of gp120–CD4 interactions, while assays on a #155-resistant strain confirmed that both the compounds targeted the Phe43 cavity, supporting the binding mode proposed by docking calculations and pharmacophore modeling.

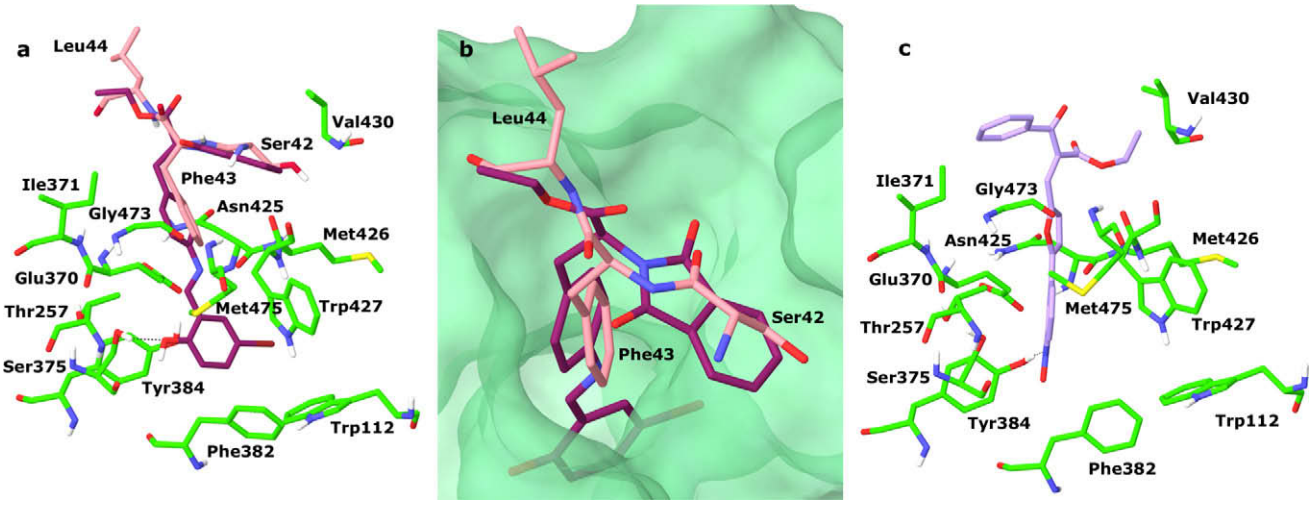


Figure 4. Binding mode of (*S*)-**2** (purple, a and b) and **4** (violet, c) to the gp120 core structure (green). CD4 residues are shown in pink (a and b).

However, the new scaffolds require further efforts in terms of reiterative hit optimization and may serve as a starting point for the development of more potent HIV entry inhibitors. Small sublibraries of the identified inhibitors will be synthesized and tested, possibly leading to the identification of better inhibitors, as well as to a further refinement of the pharmacophoric model.

Acknowledgements

This work was supported in part by the European Union THINC, HEALTH-2007-2.3.2-1, the Spanish Ministerio de Ciencia e Innovación (BFU2006-00966), FIS PI060624 and Asinex Ltd.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2009.09.029.

References and notes

1. Loregian, A.; Marsden, H. S.; Palu, G. *Rev. Med. Virol.* **2002**, *12*, 239.
2. Este, J. A.; Telenti, A. *Lancet* **2007**, *370*, 81.
3. Markovic, I. *Curr. Pharm. Des.* **2006**, *12*, 1105.
4. Vermeire, K.; Schols, D. *Expert Opin. Invest. Drugs* **2005**, *14*, 1199.
5. Kadow, J.; Wang, H. G. H.; Lin, P. F. *Curr. Opin. Investig. Drugs* **2006**, *7*, 721.
6. Lin, P. F.; Blair, W.; Wang, T.; Spicer, T.; Guo, Q.; Zhou, N.; Gong, Y. F.; Wang, H. G.; Rose, R.; Yamanaka, G.; Robinson, B.; Li, C. B.; Fridell, R.; Deminie, C.; Demers, G.; Yang, Z.; Zadjura, L.; Meanwell, N.; Colonno, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11013.
7. Si, Z.; Madani, N.; Cox, J. M.; Chruma, J. J.; Klein, J. C.; Schön, A.; Phan, N.; Wang, L.; Biorn, A. C.; Cocklin, S.; Chaiken, I.; Freire, E.; Smith, A. B., 3rd; Sodroski, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5036.
8. Ho, H. T.; Fan, L.; Nowicka-Sans, B.; McAuliffe, B.; Li, C. B.; Yamanaka, G.; Zhou, N.; Fang, H.; Dicker, I.; Dalterio, R.; Gong, Y. F.; Wang, T.; Yin, Z.; Ueda, Y.; Matiskella, J.; Kadow, J.; Clapham, P.; Robinson, J.; Colonno, R.; Lin, P. F. *J. Virol.* **2006**, *80*, 4017.
9. Zhao, Q.; Ma, L.; Jiang, S.; Lu, H.; Liu, S.; He, Y.; Strick, N.; Neamati, N.; Debnath, A. K. *Virology* **2005**, *339*, 213.
10. Madani, N.; Schön, A.; Princiotto, A. M.; LaLonde, J. M.; Courter, J. R.; Soeta, T.; Ng, D.; Wang, L.; Brower, E. T.; Xiang, S. H.; Kwon, Y. D.; Huang, C. C.; Wyatt, R.; Kwong, P. D.; Freire, E.; Smith, A. B., 3rd; Sodroski, J. G. *Structure* **2008**, *16*, 1689.
11. Guo, Q.; Ho, H. T.; Dicker, I.; Fan, L.; Zhou, N.; Friborg, J.; Wang, T.; McAuliffe, B. V.; Wang, H. G.; Rose, R. E.; Fang, H.; Scarnati, H. T.; Langley, D. R.; Meanwell, N. A.; Abraham, R.; Colonno, R. J.; Lin, P. F. *J. Virol.* **2003**, *77*, 10528.
12. Madani, N.; Perdigoto, A. L.; Srinivasan, K.; Cox, J. M.; Chruma, J. J.; LaLonde, J.; Head, M.; Smith, A. B., 3rd; Sodroski, J. G. *J. Virol.* **2004**, *78*, 3742.
13. Schön, A.; Madani, N.; Klein, J. C.; Hubicki, A.; Ng, D.; Yang, X.; Smith, A. B., 3rd; Sodroski, J.; Freire, E. *Biochemistry* **2006**, *45*, 10973.
14. Chen, B.; Vogan, E. M.; Gong, H.; Skehel, J. J.; Wiley, D. C.; Harrison, S. C. *Nature* **2005**, *433*, 834.
15. Xie, H.; Ng, D.; Savinov, S. N.; Dey, B.; Kwong, P. D.; Wyatt, R.; Smith, A. B., 3rd; Hendrickson, W. A. *J. Med. Chem.* **2007**, *50*, 4898.
16. Repik, A.; Clapham, P. R. *Structure* **2008**, *16*, 1689.
17. AMBER *Version 9*; University of California: San Francisco, CA, 2006.
18. Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, V.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T. *J. Comput. Chem.* **2003**, *24*, 1999.
19. Huang, C. C.; Tang, M.; Zhang, M. Y.; Majeed, S.; Montabana, E.; Stanfield, R. L.; Dimitrov, D. S.; Korber, B.; Sodroski, J.; Wilson, I. A.; Wyatt, R.; Kwong, P. D. *Science* **2005**, *310*, 1025.
20. *MMTSB*; The Scripps Research Institute: La Jolla, CA, 2004.
21. GRID *Version 22*; Molecular Discovery Ltd: Pinner, Middlesex, UK, 2004.
22. Kwong, P. D.; Wyatt, R.; Majeed, S.; Robinson, J.; Sweet, R. W.; Sodroski, J.; Hendrickson, W. A. *Structure* **2000**, *8*, 1329.
23. Kong, R.; Tan, J. J.; Ma, X. H.; Chen, W. Z.; Wang, C. X. *Biochim. Biophys. Acta* **2006**, *1764*, 766.
24. Fox, T.; Haaksma, E. E. J. *J. Comput. Aided Mol. Des.* **2000**, *14*, 411.
25. CATALYST *Version 4.10*; Accelrys: San Diego, CA, 2005.
26. *Asinex Gold Collection*; Asinex Ltd: Moscow, Russia, 2005.
27. Villoutreix, B. O.; Bastard, K.; Sperandio, O.; Fahraeus, R.; Poyet, J. L.; Calvo, F.; Déprez, B.; Miteva, M. A. *Curr. Pharm. Biotechnol.* **2008**, *9*, 103.
28. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug. Deliv. Rev.* **1997**, *23*, 3.
29. GOLD, *Version 3.0.1*; The Cambridge Crystallographic Data Centre: Cambridge, UK, 2006.
30. Verdonk, M. L.; Berdini, V.; Hartshorn, M. J.; Mooij, W. T.; Murray, C. W.; Taylor, R. D.; Watson, P. J. *Chem. Inf. Comput. Sci.* **2004**, *44*, 793.
31. Moncunill, G.; Armand-Ugon, M.; Pauls, E.; Clotet, B.; Este, J. A. *AIDS* **2008**, *22*, 23.
32. Moncunill, G.; Armand-Ugon, M.; Clotet-Codina, I.; Pauls, E.; Ballana, E.; Llano, A.; Romagnoli, B.; Vrijbloed, J. W.; Gombert, F. O.; Clotet, B.; De Marco, S.; Este, J. A. *Mol. Pharmacol.* **2008**, *73*, 1264.