Pergamon

Bioorganic & Medicinal Chemistry 8 (2000) 2219–2228

BIOORGANIC & MEDICINAL CHEMISTRY

The Structure of an HIV-1 Specific Cell Entry Inhibitor in Complex with the HIV-1 gp41 Trimeric Core

Genfa Zhou,[a,b,c,†] Marc Ferrer,[b,†] Rajiv Chopra,[b,c] Tarun M. Kapoor,[d] Tim Strassmaier,[e] Winfried Weissenhorn,[f] John J. Skehel,[g] Dan Oprian,[e] Stuart L. Schreiber,[c,d,h] Stephen C. Harrison[a,b,c] and Don C. Wiley[a,b,c,*]

[a]*Laboratory of Molecular Medicine, The Children's Hospital, 320 Longwood Avenue, Boston, MA 02115, USA*
[b]*Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA*
[c]*Harvard University, Howard Hughes Medical Institute, 7 Divinity Avenue, Cambridge, MA 02138, USA*
[d]*Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA*
[e]*Department of Biochemistry, Brandeis University, 415 South St., Waltham, MA 02454, USA*
[f]*EMBL, 6 rue Jules Horowitz, 38000 Grenoble, France*
[g]*National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, UK*
[h]*Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA*

Received 28 January 2000; accepted 18 April 2000

Abstract—The three-dimensional structure of the complex between an HIV-1 cell-entry inhibitor selected from screening a combinatorial library of non-natural building blocks and the central, trimeric, coiled-coil core of HIV-1 gp41 has been determined by X-ray crystallography. The biased combinatorial library was designed to identify ligands binding in nonpolar pockets on the surface of the coiled-coil core of gp41. The crystal structure shows that the non-peptide moiety of the inhibitor binds to the targeted cavity in two different binding modes. This result suggests a strategy for increasing inhibitor potency by use of a second-generation combinatorial library designed to give simultaneous occupancy of both binding sites.

Introduction

Synthetic peptides corresponding to segments of gp41 containing heptad repeats are potent inhibitors of HIV-1 fusion and entry.[1–3] The structure of the ectodomain of gp41 shows that the inhibitory peptides derive either from the inner α-helical coiled-coil or from the outer-layer α-helices of the trimeric, two-layered, rod-like molecule.[4–6] The observed conformation of gp41 is thought to result from a conformational change in the envelope glycoprotein, triggered by association with the CD4 receptor and the chemokine-binding co-receptors[7] (reviewed in ref 8). This conformational change, which activates the membrane-fusion and viral-entry activities of the glycoprotein, is required for infectivity. The peptide inhibitors of HIV-1 have been proposed to act on a transient intermediate conformation of gp41[4–6,9] adopted as the protein refolds into the double layered structure. This interpretation derives from an analogy with refolding events that have been observed in influenza virus HA.[10–12] Mutations identifying the central coiled-coil as the target of peptide inhibitors that mimic the outer layer are consistent with this proposal.[13–16]

As part of an effort to find small molecule inhibitors that bind to the coiled-coil core of gp41 and block membrane fusion, we have discovered a 445 Dalton synthetic moiety (Fig. 1a) that binds the coiled-coil when attached to the N-terminus of a 30-mer outer-layer peptide (Asn-125 to Lys-154).[17] The synthetic moiety comprises a terminal cyclopentylpropionic acid, a central ε-glutamic acid, and a linking *p*-(*N*-carboxyethyl)aminomethyl benzoic acid. A hybrid ligand, composed of this moiety linked to the peptide, inhibits HIV-1 envelope-mediated cell-cell fusion with an EC_{50} of 300 nM. The synthetic moiety was targeted to bind in a cavity in the central coiled-coil normally occupied by three side chains, Trp-117, Trp-120, and Ile-124, from two turns near the N-terminus of the outer-layer α helix of gp41. Targeting was achieved by synthesizing a combinatorial library of three building

*Corresponding author at Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA. Tel.: +1-617-495-1808; Fax: +1-617-495-9613; e-mail: dcwadmin@crystal.harvard.edu
†Equal contribution.

0968-0896/00/$ - see front matter
PII: S0968-0896(00)00155-3

blocks linked to the N-terminus of an outer-layer peptide lacking the first two α-helical turns. A library of 61,275 potential ligands was synthesized with all possible combinations of 50 building blocks at the first two positions and 25 different building blocks at the third, capping position (building blocks shown in Fig. 4 of ref 17). The highest affinity ligand was selected, while still attached to its solid-state-synthesis support bead, by its ability to capture a labeled, soluble form of the central coiled-coil of gp41.

Although the inhibitor is designed to interact with a transient intermediate of gp41, and is not expected to bind to the final folded form of gp41, the strategy of forming a hybrid molecule with a segment of the outer-layer α helix allows the structure of the bound inhibitor to be determined. An inhibitor complex was refolded in vitro from two components: an N-terminal gp41 peptide composed of residues 30 to 79 prefixed by a 31 residue trimeric-variant of GCN4,[18] and an outer-layer gp41 peptide composed of residues 125 to 154 N-terminally modified by the non-peptide moiety. The X-ray crystal structure of the inhibitor complex is reported here, with an analysis of two inhibitor binding modes that are observed and their consequences for designing higher affinity ligands. The nonpeptide moiety binds to the targeted cavity, as expected, burying many of the same atoms as the outer-layer gp41 α helix, but leaving some unoccupied positions that may be exploited in the design of improved inhibitors.

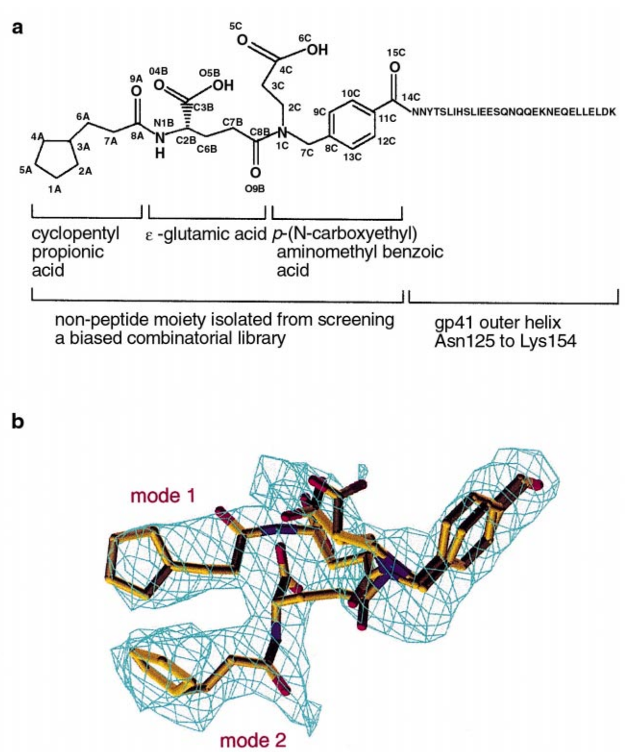

Figure 1. Structure of an inhibitor of HIV-1 viral entry. (a) Chemical structure of the three part, nonpeptide moiety derived from a combinatorial-library attached to the HIV-1 peptide, Asn-125 to Lys-154. The individual components from the combinatorial library are labeled in brackets. (b) A difference simulated-annealing omit map, contoured at 1.5 σ in the region of the synthetic substituent on the bound inhibitor, showing the electron density for the two binding modes, labeled mode-1 and mode-2. (Figure prepared with O.[35])

Results

Structure determination

The X-ray structure was determined to 3.0 Å resolution by molecular replacement using the model of the GCN4–gp41 structure.[6] The unit cell constants and space group (a = 35.8 Å, c = 165.7 Å; P321) suggested that the gp41 trimer was located on the crystallographic 3-fold symmetry axis, with only one monomer in the asymmetric unit. Furthermore, the length of the central coiled-coil, about 115 Å, suggests that the trimers must occupy the trigonal positions, 1/3, 2/3, z and 2/3, 1/3, z. The orientation and position of the model could then be determined by two dimensional grid search of the azimuthal orientation around the 3-fold axis (0 to 120°) and the z coordinate (0 to 0.5). After model building and refinement, the final structure has R_{work} of 0.24 and R_{free} of 0.295. Data and refinement statistics are shown in Table 1.

The full length of the non-peptide moiety is visible in electron density maps, but unexpectedly in two orientations, each with about 50% occupancy. The two binding modes share the same aminobenzoic acid position (right in Fig. 1b) but diverge at the two more distal building blocks (left in Fig. 1b). Although 29 of the 30 peptide residues are observed in their expected locations, the electron density for the amino acid at the connection to the non-peptide moiety is poor, suggesting disorder in the peptide linkage to the non-peptide moiety. Adjacent trimers pack in the crystal such that the nonpeptide moieties of one trimer are very close to those of another trimer, so that the non-peptide moiety may be able, in the crystal, either to bind to one trimer or to reach across to an adjacent trimer in the lattice. Because the linking peptide residue is disordered, the structure determination cannot distinguish what percent (from 0 to 100%) of the binding in the crystal is intra- or intermolecular. In any case, two ways that the non-peptide moiety can fit into

Table 1. Statistics for data collection and refinement

Space group	P321
Cell dimensions (Å)	35.8 35.8 165.7 90.0 90.0 90.0
Resolution (Å)	3.0
Unique reflections	2574
Completeness (%)	88.2
R_{sym}[a]	0.072
Refinement resolution (Å)	15.0–3.0
Sigma cut-off	2.0
No. reflections	2369
No. reflections (free set)	398
R_{work}[b]	0.240
R_{free}[b]	0.295
Protein atoms	890
Compound atoms	33
Rms deviations	
RMS bonds (Å)	0.010
RMS angles (°)	1.60

[a] $R_{sym} = \Sigma|I - \langle I \rangle|/\Sigma I$, where I is the observed intensity, and $\langle I \rangle$ is the average intensity of multiple observations of symmetry-related reflections.
[b] $R = \Sigma\|Fo|-|Fc\|/\Sigma|Fo|$, where R_{free} is calculated for a randomly chosen 10% of reflections, and R_{work} is calculated for the remaining 90% of reflections used for structure refinement.

the cavity on the central coiled-coil are clear (Fig. 2a) and only one residue of the peptide linker (Asn-125) is ambiguously positioned.

The two binding modes of the inhibitor

The two inhibitor binding modes, labeled mode 1 and mode 2 in Figure 2a, both have the cyclopentyl group inserted into the non-polar pockets normally occupied by Trp-117, Trp-120, and Ile-124 from the outer layer α-helix of gp41 (Fig. 3). Most of the same atoms that were earlier observed to be contacted by the outer layer α-helix are now contacted by atoms in the non-peptide part of the inhibitor (solid symbols in Fig. 3a), although the inhibitor makes fewer contacts (compare Fig. 3a with 3b). In binding mode-2, the cyclopentyl group of the inhibitor lies in approximately the same plane as Trp-117 and contacts some of the same target atoms as Trp-117 (square symbols in Fig. 3). In binding mode-1, the cyclopentyl group of the inhibitor is tipped on edge and contacts most of the same target atoms as Trp-120 and Ile-124 (circles and triangles in Fig. 3). In binding mode-1, the ϵ-glutamic acid constituents make a novel contact with Leu-54, whereas in mode-2 it contacts CD1 of Leu-57 that in the native gp41 structure contacts Ile-124. The *p*-(*N*-carboxyethyl)aminobenzoic acid building block, which links the other substituents to the Asn-125-Lys-154 α-helix (Figs 1b and 2a), makes no contacts to the inner coiled-coil. Instead its position is apparently stabilized by a number of contacts to Asn-124 and Asn-125 in the outer-layer peptide to which the combinatorial library was appended.

The observation that one binding mode (mode 2) of the inhibitor (Fig. 3c) approximately mimics Trp-117 and the other binding mode (mode 1) mimics both Trp-120 and Ile-124 of the outer-layer α helix, suggests that a double-headed inhibitor, incorporating two cyclopentylpropionic acid constituents or similar moieties, might have higher affinity.

Some of the building blocks chosen for use in the second and third positions of the combinatorial library contained carboxylate groups in an effort to mimic Asp-121 of the outer-layer α-helix,[17] which forms a hydrogen bond and salt bridge to Lys-63 on the central coiled-coil (Fig. 2a). The inhibitor selected from the library contains two carboxylate groups, one at the second and one at the third position (Fig. 1a). In the inhibitor complex structure determined here, only the carboxylate of the

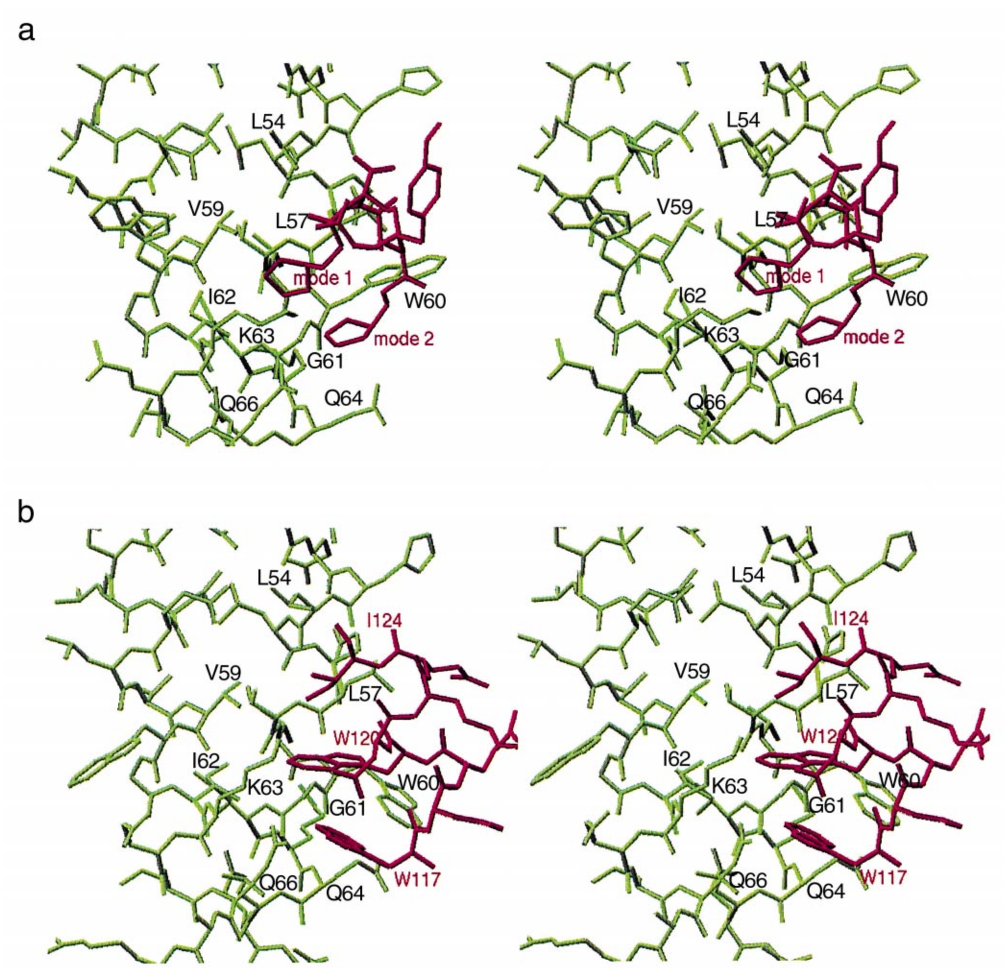

Figure 2. Comparison of the inhibitor complex of the gp41 coiled-coil core with the native structure containing the outer-layer α-helix of gp41: (a) stereo drawing of two α-helices of the core coiled-coil (green) and the two binding modes of the bound inhibitor (red); (b) stereo drawing of two α-helices of the core coiled-coil (green) and two turns of the outer-layer α-helix of gp41 (red). (Figure prepared with RIBBONS.[38])

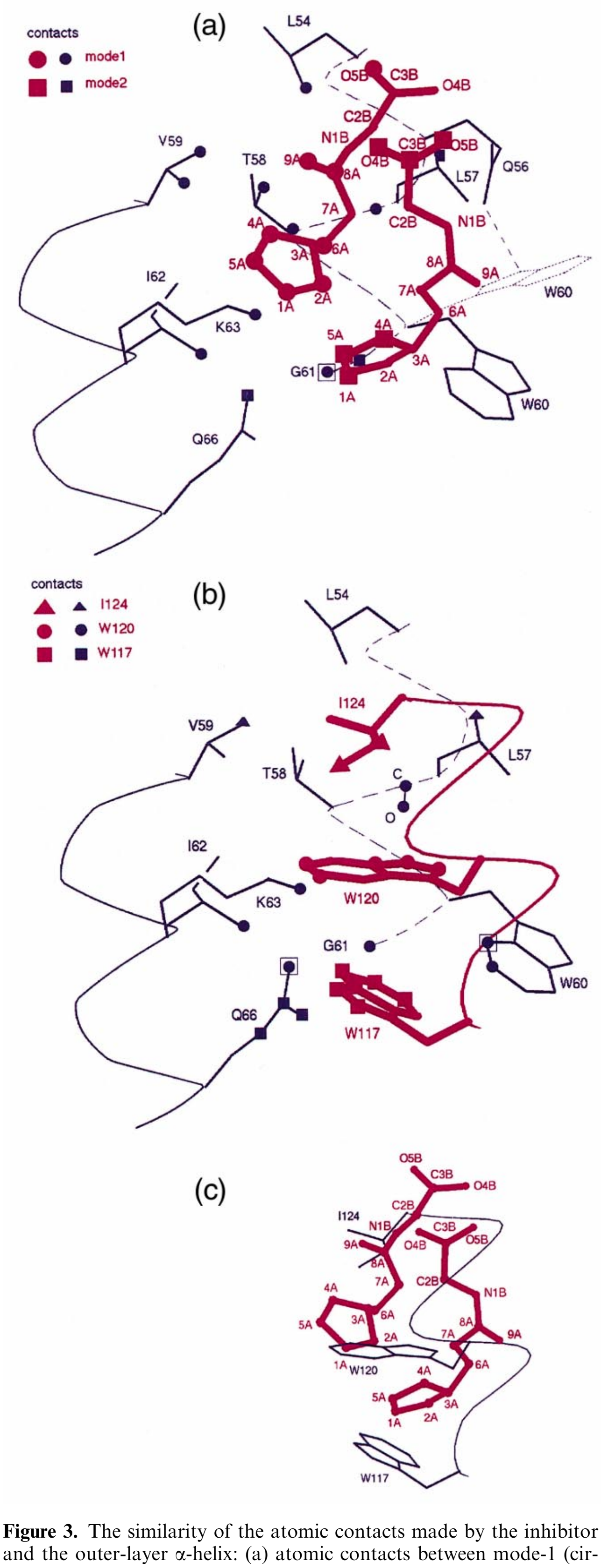

Figure 3. The similarity of the atomic contacts made by the inhibitor and the outer-layer α-helix: (a) atomic contacts between mode-1 (circles) and mode-2 (squares) of the inhibitor (red) and the core coiled coil (blue); (b) atomic contacts between Ile-124 (triangles), Trp-120 (circles), and Trp-117 (squares) of the outer-layer a helix (red) and the core coiled coil (blue); (c) superposition of binding mode-1 and -2 of the inhibitor with the N-terminal Trp-rich region of the outer helix of gp41. (Figure prepared with RIBBONS.[38])

third constituent, that of the *p*-(*N*-carboxyethyl)aminobenzoic acid moiety, is in a position to interact with Lys-63, but the electron density for the carboxylate group is weak, so that there may only be an electrostatic interaction rather than a hydrogen bond. The structure suggests that a longer linker, propyl instead of ethyl, to the carboxylate group might provide the desired hydrogen-bonded interaction.

Glu-123 of the outer-layer α helix also contacts the core coiled-coil, but only to make van der Waal contacts with Leu-57. The inhibitor also makes van der Waal contacts with Leu-57.

The inhibitor binds to the intended nonpolar cavity

The two binding modes of the non-peptide moiety of the inhibitor, if considered as one composite ligand, bury a total of 244 Å^2; the solvent accessible surface area of the cavity buried by outer-layer α helix in the native structure is 297 Å^2. Most of the same atoms in the target are covered as is evident from a comparison of Figure 4a and b, in which the radii of atoms are drawn proportional to their buried solvent accessible surface area. This comparison indicates that the strategy to direct the combinatorial ligands to the non-polar cavity by using a biasing peptide[17] was reasonably successful.

Approximately comparable solvent accessible surface areas on the inhibitor were buried in the inhibitor complex as were buried on the outer-layer α helix in gp41 (compare Fig. 4c and d).

A moving target

Trp-60 of the central coiled-coil (target) has rotated 120° about the Cα–Cβ bond in the inhibitor complex relative to gp41 (Fig. 3a). This movement displaces two atoms of the indole ring that had been part of the targeted cavity in the gp41 structure (circles on W60 in Fig. 3b). This conformational alteration is apparently stabilized by a new hydrogen bond in the inhibitor complex between the indole nitrogen of Trp-60 and the side chain of Gln-56 of the central coiled-coil (Fig. 3a). There is very little change in the solvent-exposed surface area of the indole, because it simultaneously exposes a surface of 159 Å^2 and buries a surface on its opposite edge of 162 Å^2. The effect of this motion, however, is to increase the apparent size of the nonpolar cavity by the removal of Trp-60 (see W-60 in Fig. 5). Future strategies for designing inhibitors could either aim to prevent the movement of Trp-60, or take advantage of the movement that provides an even larger target cavity for ligand binding.

Surface complementarity

Lawrence and Colman[19] have proposed a measure of surface complementarity, Sc, which depends both on the distance between two juxtaposed molecular surfaces and their local curvatures. Large values of Sc indicate a close fit and good shape complementarity between surfaces. Protein–protein interfaces in oligomeric proteins usually

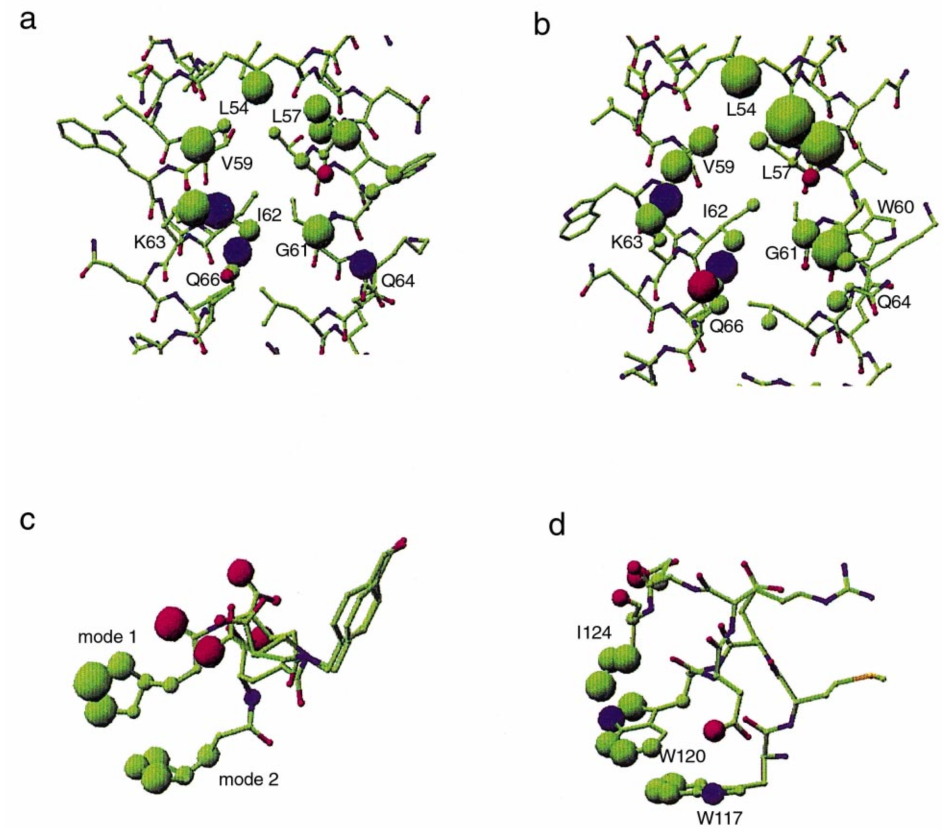

Figure 4. Atomic surfaces buried by interactions with the targeted cavity: (a) surfaces buried in the coiled coil cavity by the inhibitor; (b) surfaces buried in the coiled-coil cavity by the segment of the outer-layer α-helix from residues Trp-117 to Leu-124; (c) surfaces buried on the inhibitor in binding modes 1 and 2 when bound to the cavity in (a); (d) surfaces buried on the outer-layer α-helix from residues Trp-117 to Leu-124 when bound to the core coiled-coil in (b). (Spherical radii are proportional to solvent accessible surface area buried. Nitrogens are colored blue, oxygens are red, and carbons are green. Figure prepared with RIBBONS.[38])

have Sc values between 0.65 and 0.80. If the two binding modes of the inhibitor are considered as one composite ligand, then the shape complementary of the combined inhibitor with its target would be 0.64, comparable to the value of 0.63 for the shape complementarity of the first two turns of the outer-layer α-helix to the central coiled-coil. A plot of the shape complementarity of the apposed surfaces (Fig. 5) shows that although the inhibitor fits well in the center of the targeted cavity (Fig. 5a), the fit is not as favorable around the periphery of the cavity as that of the outer-layer α-helix (Fig. 5b).

Discussion

Since the discovery that a conformational change in viral glycoproteins is required to activate the membrane fusion/viral entry activity, inhibitors have been sought that prevent the conformational change, that trigger the change prematurely in order to inactivate, or that inhibit a step in the conformational refolding (reviewed in ref 20). The most successful viral entry inhibitors to date are synthetic peptides corresponding to HIV-1 gp41 sequences predicted to be α-helical based on the presence of heptad repeats of nonpolar residues and on similarities to the influenza HA2, where heptads follow a fusion peptide that lies immediately C-terminal to a cleavage site in the biosynthetic precursor.[21–25] One of these peptides (T20/DP-178) corresponds to residues 127–162 of the outer-layer of gp41. When administered to AIDS patients intravenously, it reduced circulating HIV-1 to undetectable levels, a result comparable to more conventional anti-HIV-1 therapies and an indication that viral entry can be inhibited in vivo.[3] These peptides appear to inhibit steps in the conformational change by interacting with a transient intermediate, mimicking an intramolecular interaction that occurs during refolding.[4–6,9] Apparently these inhibitors disrupt refolding by rerouting the process to a dead-end state. Otherwise, the normal intramolecular interaction would be expected to dominate at equilibrium due to the high local concentration of the intramolecular sequence relative to the exogenous peptide. Peptides corresponding to analogous sequences of the retrovirus HTLV-1 and the paramyxovirus SV5 membrane glycoproteins also inhibit membrane fusion by those viruses.[26–29]

A smaller molecular weight, non-peptide inhibitor might have better therapeutic value due to improved pharmacokinetics and oral bioavailability. Whether a low molecular weight inhibitor could be targeted to the transiently exposed core-coiled coil core of gp41 and whether it could block the formation of a large intramolecular interface were questions we sought to answer by selecting a small molecule from a biased combinatorial library. The structure reported here confirms that the small organic moiety selected by our strategy (Fig. 1a) binds in the targeted cavity, when presented linked to part of the outer-layer α helix of gp41, and that it contacts the same atoms that form an intramolecular interface in gp41 (Fig. 4). The small molecule alone had no inhibitory activity, however, presumably because it interacted too weakly,[17] but it increased the potency of the 30-mer peptide (EC_{50} of 6.6 μM) by about 20-fold (hybrid: EC_{50} of 300 nM).

An 18 residue D-amino acid peptide inhibitor, cyclized by a disulfide bond between positions 3 and 14, and with the same sequence as an L-amino acid phage-displayed peptide selected to bind a D-amino acid construct of the site targeted here, binds to the targeted cavity and inhibits HIV-1 glycoprotein mediated cell fusion (IC_{50}~10 μM).[30] This result demonstrates that blocking a relatively small targeted cavity is sufficient to inhibit the formation of the large intramolecular interface between the core coiled-coil and the outer-layer α helix of gp41. The inhibitory capacity of both of these targeted non-natural inhibitors, the cyclic D-peptide and the organic moiety–peptide hybrid, strongly suggest that the targeted cavity on the core coiled-coil is accessible at least transiently during the gp120/gp41 conformational change that precedes membrane fusion.

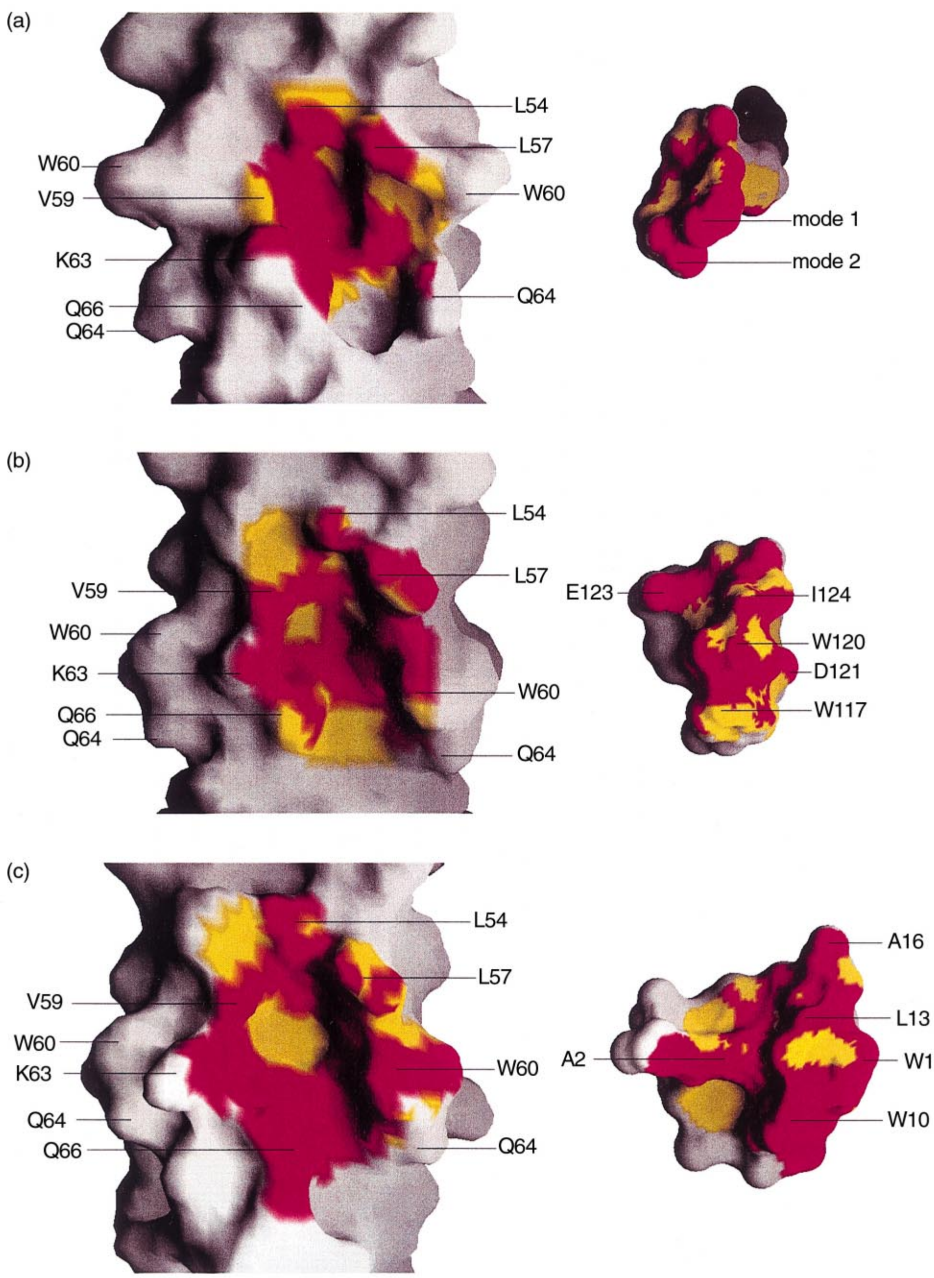

Figure 5. Surface complementarity of the inhibitor binding interaction compared with the intramolecular interaction in gp41. (a) The surface of the targeted cavity (left) and the bound inhibitor ligand (right), with both binding modes 1 and 2 shown. The interface is shown 'opened' like a book. The molecular surfaces are colored by the shape complementarity index.[19] Regions of high shape complementarity (0.5 < Sc < 1.0) are red, medium shape complementarity (0.0 < Sc < 0.5) are yellow. Surfaces with low complementarity or not in the interface are gray. (b) The surface of the targeted cavity (left) and the outer-layer α helix from residues Trp-117 to Leu-124 (right) as found in gp41, colored as in (a). (c) The surface of the targeted cavity (left) and the D-peptide inhibitor,[30] colored as in (a). (Surfaces drawn with GRASP.[39])

There is a striking similarity between the way the cyclopentyl moieties in the combinatorial ligand and two D-tryptophan indole rings of the cyclic D-peptide bind to the targeted cavity. Whereas Trp-117 and Trp-120 of the gp41 outer-layer α helix are inserted into the cavity horizontally, with their flat rings perpendicular to the coiled-coil axis (Fig. 3b), the cyclopentyl group in binding mode-1 (Fig. 3a) and the two D-tryptophan indole rings of the D-peptide insert into the cavity edge-on but vertically, with their flat rings oriented parallel to the coiled-coil axis. The cyclopentyl ring appears to fit more deeply into the cavity than do the indole rings, making a better fit (compare Fig. 5a, red at bottom of cavity, with 5b and c, yellow at bottom of cavity). This comparison suggests that smaller nonpolar rings like the cyclopentyl group presented on longer linkers, as found on the cyclopentylpropionic acid constituent (Fig. 1a), may be better probes for determining the depths of this nonpolar cavity than naturally occurring tryptophan rings. Table 2 compares the contacts made by the synthetic moiety, the original outer-layer L-peptide, and the D-peptide.

In the combinatorial inhibitor (Fig. 1a), part of the ε-glutamic acid constituent mimics Ile-124 of gp41 (Fig. 3a and b), and Trp-60 rotates out of the site, increasing the size of the cavity (Fig. 3a). D-Leu-13 of the cyclic D-peptide mimics Ile-124 of gp41, and D-Ala-2 of the cyclic D-peptide binds Trp-60, possibly helping to stabilize its native-like position (compare Fig. 5a and c). Future

Table 2. The interactions between the inner core pocket and the synthetic molecule, the L-peptide, and the D-peptide

Inner core		Synthetic[a]		L-Peptide		D-Peptide	
Residue	Atom	Mode	Atom	Residue	Atom	Residue	Atom
Helix-1							
V59	O					W12	CZ3,CH2
	CG1	MD1	4A			W12	CZ3
	CG2	MD1	4A	I124	CD1		
I62	CG2	MD1	1A,5A	W120	CH2	W12	CH2
K63	N					W12	CH2
	CA					W12	CZ2,CH2
	CB					W12	CZ3,CH2
	CE			D121	OD1		
	NZ	MD1	6A	W120	CZ3		
				D121	OD1		
Q66	CG			W117	CZ2		
	CD			W117	CE2,NE1	W12	NE1
					CZ2,CH2	E9	O
	OE1			W117	CE2,CD1	W12	CE2,NE1,CZ2
					NE1,CZ2	E9	O
	NE2	MD1	1A,5A	W117	CE2,CD1	E9	CG,C,O
		MD2	1A,5A		CZ2,CH2	W10	CE3,CD2,CE2
				W120	CZ3,CH2		NE1,CZ3,CZ2
							CH2
Helix-2							
L54	CD1	MD1	O5B,*A,9A			A16	C,O,NT
	CD2					A16	O
L57	C			W120	NE1		
	O	MD1	1A,2A	W120	CD1,NE1	L13	CD2
					CE2		
	CD1	MD1	C3B,O4B	E123	C,CB,O	L13	O
			O5B	I124	N,CG1	A16	NT
		MD2	C3B,O4B				
			O5B				
	CD2			E123	OE1	A2	CB
T58	CA	MD1	1A,5A				
	CG2	MD1	5A				
W60	CD2					A2	N
	CE2					G1	C,CA
						A2	N
	CE2			W120	CD1	A2	CB
	NE1					G1	CA
	CZ2					G1	C
						A2	N
	CZ3			W117	CZ3	A2	CA,CB
				W120	CD1	A10	CG,CD2,CE2
	CH2					A10	CG,CD1,NE1
						A2	N,CA
G61	N	MD1	4A				
		MD2	4A				
	CA	MD1	1A,4A,5A	W120	NE1		
		MD2	4A,5A				
Q64	CB					W10	CZ2
L65	CD2					W10	CH2

[a]MD1 stands for mode-1, and MD2 for mode-2.

ligands could target Trp-60 either in its native position, like the D-peptide, or in the rotated position with the increased cavity areas, as in the ligand studied here. Although the carboxylate of the *p*-(*N*-carboxyethyl) aminomethyl benzoic acid constituent (Fig. 1a) is positioned to make electrostatic interactions with Lys-63 of the core coiled-coil in the combinatorial ligand, the absence of a clear hydrogen bond to Lys 63 in that inhibitor or even an electrostatic interaction in the cyclic D-peptide inhibitor, suggests room for improvement. The solvent accessible surface areas buried on the core coiled-coil, 297 $Å^2$ by the gp41 outer-layer α-helix and 458 $Å^2$ by the cyclic D-peptide are greater than the 244 $Å^2$ buried by the combination of the two binding modes observed here. This comparison suggests that a second generation combinatorial scaffold that aims to include both binding modes observed here should also be longer and broader, to allow some new building blocks to bury more of the target surface (Fig. 5).

An inhibitor combining the two binding modes observed here for the combinatorial ligand would be expected to have a substantial higher affinity for the target and be a more potent membrane fusion/viral entry inhibitor. Such a double-headed inhibitor might be achieved by designing a linker with a program like CAVEAT[31,32] and might be combined with better mimics for Ile-124 in a future combinatorial library. Providing a more rigid scaffold for the pharmacophores than present on the current linear inhibitor, such as in the α helix in gp41 or the 18-mer cyclic D-peptide, could both decrease the entropic loss on binding by pre-organizing the binding elements and provide a larger binding footprint by contacting some of the atoms missed by this first lead (compare Fig. 5a–c).

There are at least two practical motivations for exploring the possibility of targeting the transient intermediate of a viral membrane fusion glycoprotein for inhibiting viral infection. One is that many enveloped viruses use similar entry mechanisms. Despite major differences in their glycoproteins, all apparently share a complex conformational change, during which a transient intermediate containing an exposed coiled-coil core may be available for interaction with an inhibitor (reviewed in ref 20). A second motivation is that such compounds might become part of multiple inhibitor antiviral therapies ('combination therapy'). Inhibitor induced resistance, as found in mono-inhibitor antiviral therapies employing enzymatic targets, suggests the need for multiple antiviral targets. Transient intermediates in a refolding process may be less compatible with resistance mutations, because the inhibitor target is part of an intramolecular interface requiring the tight fit between two distinct segments. A mutation in the target site that would prevent inhibitor binding might also often interfere with side-chain packing at the interface formed in the final conformation. In the case of the envelope glycoprotein of HIV-1, residues in the targeted cavity also pack in another, currently unknown, environment in the gp120/gp41 molecule before the receptor/co-receptor induced conformational change, and may have yet other interactions in the precursor gp160. The multiple interactions of this segment of the envelope protein polypeptide chain may be incompatible with mutations that would confer resistance. The overall fitness for infectivity of resistant mutant viruses is difficult to predict from structural considerations, however, and exploration of these issues will require biological experiments.

Experimental

Crystallization and data collection

The GCN4-gp41/inhibitor complex was refolded in vitro and purified as described previously.[17] Crystals were grown in hanging drops by combining 1 μL protein solution (10 mg mL^{-1} in 20 mM HEPES, pH 8.3, 75 mM NaCl) with 1 μL of reservoir solution (50 mM PIPES, pH 6.9, 0.5 M $MgAc_2$, 5% isopropanol). For data collection, crystals were soaked sequentially in reservoir solution with 10, 20 and 30% glycerol (30 min each, and twice in 30%), harvested, and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K using a MAR345 image plate detector and an Elliot GX-13 rotating anode source with mirror optics, and processed using the programs DENZO and SCALEPACK (HKL Research).[33] Crystallographic data statistics are in Table 1.

Molecular replacement

The structure of the gp41 ectodomain, determined by Weissenhorn et al.[6] comprises three copies each of two peptides: (1) 29 residues of a trimeric GCN4 fused to residues 30–77 of gp41, and (2) residues 117–154 of the outer-layer α-helix. In the inhibitor complex, an organic compound replaces residues 117–124 of the outer-layer α-helix. For molecular replacement, residues 117–124 of the gp41 model were omitted, and side chains that pointed into solution were replaced by alanines. The unit cell size suggested that only one gp41 monomer was present in the asymmetric unit of the P321 crystal and we concluded that the trimeric molecule must be located on a crystallographic 3-fold symmetry axis. The length of the c axis, 165.7 Å, suggested that the two molecules (each approximately 115 Å long) could not be on the 3-fold axes that are intersected by 2-fold symmetry axes and, therefore, must be located on the trigonal positions ($x=1/3$, $y=2/3$ and $x=2/3$, $y=1/3$). The rotation/translation search is thus limited to a 120° azimuthal search around the c axis and a z translation. A grid search of 10° and 1.66 Å intervals with a modified version of AMORE[34] located the molecule, which after rigid body refinement gave an R factor of 0.44 and correlation coefficient of 0.68.

Model building and refinement

A model was built into a 2Fo-Fc electron density map with the program of O[35] and refined using CNS.[36] Ten percent of the reflections were set aside for R_{free} calculations. The nonpeptide moiety (Fig. 1b) was modeled into interpretable density when the R_{free} was 0.34.

Buried surfaces

Buried solvent accessible surface areas were calculated by subtracting the surface areas of bound molecule from that of unbound molecule with program SURFACE[37] in CCP4. The radii of spheres in Figure 4 are the square root of the buried surface areas.

Acknowledgements

We thank members of the Schreiber, Harrison, and Wiley laboratories for their help and K. J. Smith and P. Colman for assistance with their Sc calculation program. M. F. acknowledges a fellowship from the Ministry of Education (Spain). This research was supported by the NIH program project (GM-39589), the Howard Hughes Medical Institute, and the Medical Research Council (UK). DCW, SCH, and SLS are investigators in the Howard Hughes Medical Institute.

References and Notes

1. Jiang, S.; Lin, K.; Strick, N.; Neurath, A. R. *Nature* **1993**, *365*, 113.
2. Wild, C. T.; Shugars, D. C.; Greenwell, T. K.; McDanal, C. B.; Matthews, T. J. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9770.
3. Kilby, J. M.; Hopkins, S.; Venetta, T. M.; DiMassimo, B.; Cloud, G. A.; Lee, J. Y.; Alldredge, L.; Hunter, E.; Lambert, D.; Bolognesi, D.; Matthews, T.; Johnson, M. R.; Nowak, M. A.; Shaw, G. M.; Saag, M. *S. Nat. Med.* **1998**, *4*, 1302.
4. Lu, M.; Blacklow, S. C.; Kim, P. S. *Nat. Struct. Biol.* **1995**, *2*, 1075.
5. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263.
6. Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. *Nature* **1997**, *387*, 426.
7. Sattentau, Q. J.; Moore, J. P. *J. Exp. Med.* **1991**, *174*, 407.
8. Wyatt, R.; Sodroski, J. *Science* **1998**, *280*, 1884.
9. Blacklow, S. C.; Lu, M.; Kim, P. S. *Biochemistry* **1995**, *34*, 14955.
10. Bullough, P. A.; Hughson, F. M.; Skehel, J. J.; Wiley, D. C. *Nature* **1994**, *371*, 37.
11. Chen, J.; Wharton, S. A.; Weissenhorn, W.; Calder, L. J.; Hughson, F. M.; Skehel, J. J.; Wiley, D. C. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12205.
12. Chen, J.; Skehel, J. J.; Wiley, D. C. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8967.
13. Rimsky, L. T.; Shugars, D. C.; Matthews, T. J. *J. Virol.* **1998**, *72*, 986.
14. Weng, Y.; Weiss, C. D. *J. Virol.* **1998**, *72*, 9676.
15. Judice, J. K.; Tom, J. Y.; Huang, W.; Wrin, T.; Vennari, J.; Petropoulos, C. J.; McDowell, R. S. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13426.
16. Furuta, R. A.; Wild, C. T.; Weng, Y.; Weiss, C. D. *Nat. Struct. Biol.* **1998**, *5*, 276.
17. Ferrer, M.; Kapoor, T.; Strassmaier, T.; Weissenhorn, W.; Skehel, J. J.; Oprian, D.; Schreiber, S.; Wiley, D. C.; Harrison, S. C. *Nat. Struct. Biol.* **1999**, *9*, 953.
18. Weissenhorn, W.; Calder, L. J.; Dessen, A. A.; Laue, T.; Skehel, J. J.; Wiley, D. C. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6065.
19. Lawrence, M. C.; Colman, P. M. *J. Mol. Biol.* **1993**, *234*, 946.
20. Skehel, J. J.; Wiley, D. C. *Annu. Rev. Biochem.* **2000**, *69*.
21. Gallaher, W. R.; Ball, J. M.; Garry, R. F.; Griffin, M. C.; Montelaro, R. C. *AIDS Res. Hum. Retroviruses* **1989**, *5*, 431.
22. Chambers, P.; Barr, J.; Pringle, C. R.; Easton, A. J. *J. Virol.* **1990**, *64*, 1869.
23. Chambers, P.; Pringle, C. R.; Easton, A. J. *J. Gen. Virol.* **1990**, *71*, 3075.
24. Wilson, I. A.; Skehel, J. J.; Wiley, D. C. *Nature* **1981**, *289*, 366.
25. Wiley, D. C.; Wilson, I. A.; Skehel, J. J. *Nature* **1981**, *289*, 373.
26. Rapaport, D.; Ovadia, M.; Shai, Y. *EMBO J.* **1995**, *14*, 5524.
27. Lamb, R. A.; Joshi, S. B.; Dutch, R. E. *Mol. Membr. Biol.* **1999**, *16*, 11.
28. Joshi, S. B.; Dutch, R. E.; Lamb, R. A. *Virology* **1998**, *248*, 20.
29. Sagara, Y.; Inoue, Y.; Shiraki, H.; Jinno, A.; Hoshino, H.; Maeda, Y. *J. Virol.* **1996**, *70*, 1564.
30. Eckert, D. M.; Malashkevich, V. N.; Hong, L. H.; Carr, P. A.; Kim, P. S. *Cell* **1999**, *99*, 103.
31. Bartlett, P. A.; Shea, G. T.; Telfer, S. J.; Waterman, S. *CAVEAT: A Program to Facilitate the Structure-derived Design of Biologically Active Molecules*; Roberts, S. M., Ed.; Royal Society of Chemistry: London, 1989; pp 182–196.
32. Bartlett, P. A.; Lauri, G. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 51.
33. Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307.
34. Navaza, J.; Saludjian, P. *Methods Enzymol.* **1997**, *276*, 581.
35. Jones, T. A.; Kjeldgaard, M. *Methods Enzymol.* **1997**, *277*, 173.
36. Brünger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J.-S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *CNS*, 0.5 ed.; Yale University: New Haven, 1998.
37. Handschumacher, M. D.; Richards, F. M. *Surface in CCP4*, 3.5 ed.; 1999.
38. Carson, M. *J. Appl. Crystallogr.* **1991**, *24*, 958.
39. Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 281.