Accepted Manuscript

Helical cyclic pentapeptides constrain HIV-1 Rev peptide for enhanced RNA binding

Rosemary S. Harrison, Nicholas E. Shepherd, Huy N. Hoang, Renée L. Beyer, Gloria Ruiz-Gómez, Michael J. Kelso, W. Mei Kok, Timothy A. Hill, Giovanni Abbenante, David P. Fairlie, Professor

PII: S0040-4020(14)01139-9

DOI: 10.1016/j.tet.2014.07.096

Reference: TET 25890

To appear in: *Tetrahedron*

Received Date: 2 May 2014

Revised Date: 25 July 2014

Accepted Date: 29 July 2014

Please cite this article as: Harrison RS, Shepherd NE, Hoang HN, Beyer RL, Ruiz-Gómez G, Kelso MJ, Mei Kok W, Hill TA, Abbenante G, Fairlie DP, Helical cyclic pentapeptides constrain HIV-1 Rev peptide for enhanced RNA binding, *Tetrahedron* (2014), doi: 10.1016/j.tet.2014.07.096.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

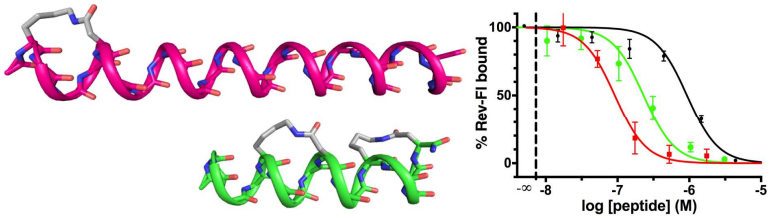
Graphical Abstract

Helical cyclic pentapeptides constrain HIV-1 Rev peptide for enhanced RNA binding

Leave this area blank for abstract info.

Rosemary S. Harrison, Nicholas E. Shepherd, Huy N. Hoang, Renée L. Beyer, Gloria Ruiz-Gómez, Michael J. Kelso, W. Mei Kok, Timothy A. Hill, Giovanni Abbenante, David P. Fairlie*

Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Tetrahedron
journal homepage: www.elsevier.com

Helical cyclic pentapeptides constrain HIV-1 Rev peptide for enhanced RNA binding

Rosemary S. Harrison, Nicholas E. Shepherd, Huy N. Hoang, Renée L. Beyer, Gloria Ruiz-Gómez, Michael J. Kelso, W. Mei Kok, Timothy A. Hill, Giovanni Abbenante, David P. Fairlie*

Division of Chemistry and Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia
*Correspondence to: Professor David Fairlie (Tel: +61-733462989, Email: d.fairlie@imb.uq.edu.au).

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
Cyclic peptide
Alpha helix
HIV rev
RNA
Helix induction

ABSTRACT

HIV-1 Rev is a 116 residue transporter protein that enters the host cell nucleus and uses its 17 amino acid segment (Rev_{34-50}) to bind and capture a specific piece of RNA, the Rev Response Element (RRE), for transport to the cytoplasm. This is critical for HIV replication. In isolation, Rev_{34-50} shows negligible structure in water, but is alpha helical in a mixture of water and 2,2,2-trifluoroethanol (TFE) or when bound to RRE. Here we report that helix-constrained cyclic pentapeptides, either appended to the N-terminus or incorporated within Rev_{34-50}, are efficient helix nucleators in water. They induce up to 90% alpha helicity for isolated Rev peptides in water and confer high RNA-binding affinity.

2009 Elsevier Ltd. All rights reserved.

1. Introduction

Designing drugs to selectively target specific protein-RNA interactions remains extremely challenging[1] due to the large and flexible interacting surfaces, the complex structural plasticity, and the negatively charged and solvent exposed surfaces of RNA. Nevertheless, for human immunodeficiency virus (HIV) and other retroviruses with small genomes (and thus few proteins for targeting by drugs), viral RNA and RNA-binding proteins are important targets for therapeutic intervention. Common to a number of RNA-binding proteins is their role as transporters.[2] For example, HIV-1 Rev is a 116 residue viral protein that plays a critical role in HIV replication[3,4] by binding to RNA (**Figure 1**) in the nucleus of a virus-infected host cell and transporting it to the cytoplasm for translation into viral structural and functional proteins that assemble into new viral particles.[3] Specific inhibitors of the binding of the HIV-1 Rev protein to its RNA partner might inhibit HIV replication, however very few compounds are known to potently antagonize this interaction.[4-8] Among known inhibitors are peptides,[4,5] aminoglycoside antibiotics (e.g. neomycin B),[6] nucleic acid aptamers,[7] and other small molecules.[8] All of these ligands have issues with either bioavailability, stability, selectivity or potency *in vivo*. Here we focus on promoting alpha helicity and enhancing affinity of the RNA-binding HIV-1 Rev peptide sequence by incorporating highly alpha helical cyclic pentapeptides[9] at the N-terminus and within the arginine-rich RNA-binding component of Rev. The effects of introducing these cyclic peptides on helix induction and affinity for the Rev-binding element RRE of RNA is reported.

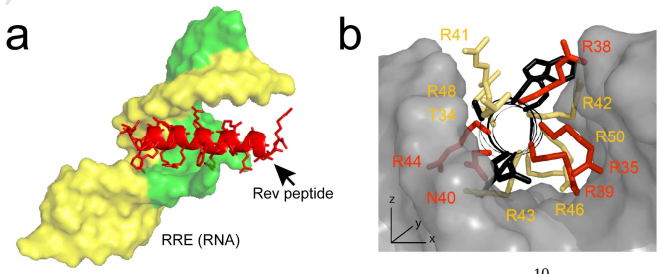


Figure 1. RNA-bound Rev_{34-50} peptide (PDB:1ETF).[10] (a) HIV-1 Rev residues 34-50 (red; peptide 1) bound to its double stranded RNA target sequence (green/yellow; Rev Response Element, RRE). (b) End view of the NMR-derived solution structure of H-Rev_{34-50}-OH peptide bound to the RRE. Residue numbers correspond to the Rev protein.

The Rev protein (18kDa) has two distinct functional domains identified through mutagenesis.[11] The N-terminal domain contains both the nuclear localisation signal and the RNA-binding domain (RBD), whereas the C-terminal domain contains the nuclear export signal. A 17-residue arginine-rich region in the N-terminal domain of Rev, residues 34-50, is important for both targeting to the nucleus[12] and specific binding to the major groove of the RNA segment termed the Rev Response Element (RRE).[4] When isolated, this short synthetic peptide sequence corresponding to HIV-1 Rev_{34-50} (H-TRQARRNRRRRWRERQR-NH_2; denoted from this point on as peptide Rev_{1-17} or peptide **1**) has negligible structure in water but becomes alpha helical when 2,2,2-trifluoroethanol (TFE) is added,[13] or when bound to the Rev Response Element[4,10] (Figure

ACCEPTED MANUSCRIPT

1). Alanine mutagenesis of this 17-residue peptide[4a] revealed that Thr1, Arg2, Arg5, Arg6, Asn7 and Arg11 were critical for high affinity binding to RRE and for specificity over other RNA segments, while Arg13, Arg15 and Arg17 also made important electrostatic interactions.

2. Results and discussion

Previously, we reported on a cyclic pentapeptide, Ac-(1,5-cyclo)-[KARAD]-NH_2, that formed a single α-helical turn with unusually high stability in water.[9a] This pentapeptide contained a crucial lactam bridge between the side chains of lysine (position 1) and aspartic acid (position 5), which promoted formation of three α-helix-defining hydrogen bonds within the peptapeptide.[9a] This lactam bridge was more effective in inducing alpha helicity within pentapeptides than any other lactam bridge[9a] or other hydrocarbon linkers between side chains of residues 1 and 5.[14a] Moreover, helicity was tolerant of many amino acids at positions 2, 3 and 4 in the sequence,[14a,b] helix-favouring amino acids in these positions in proteins also favoured helicity in the cyclic pentapeptide,[14b] and the cyclic pentapeptides could be combined as modules[9b] to confer alpha helicity in, and biological activity to, longer peptide sequences.[14b-f] We therefore decided to examine its influence here on alpha helicity when appended to the end, or within, the Rev peptide sequence corresponding to the RNA-binding domain.

Using the NMR solution structure of Rev_{1-17} (**1**) bound to RRE (pdb:1ETF),[10] we first modelled whether the cyclic pentapeptide attached to the N- or C-terminus might sterically clash with the RNA. There was insufficient space to permit attachment of the cyclic peptide to the C-terminus of **1** and, even when attached to the N-terminus of **1**, there were clashes with the RNA backbone bases of G47-G48 (**Figure 2**) that would be expected to negatively impact on binding affinity. To avert this potential problem, a series of models were created for Ac-(1,5-cyclo)-[KARAD]-$Ala_{(n)}$-Rev_{1-17}-NH_2, with different numbers of alanine residues (n = 0-4) as spacers with high helical propensity and small side chains to limit steric interference with the RNA backbone. We found that at least three alanine residues were needed to separate the Rev peptide from the cyclic peptide and thereby minimise steric clashes between the cycle and the RNA (**Figure 2**). The models suggested that the side chain lactam bridge may interact with the RRE backbone 64-GCU-66 when n = 2, but that n ≥ 3 would enable the side chain lactam bridge to be projected beyond the RNA backbone and thus would not be expected to interfere with binding (**Figure 2**). We therefore synthesized H-Rev_{1-17}-NH_2 (**1**), Ac-(1,5-cyclo)-[KARAD]-Rev_{1-17}-NH_2 (**2**), and Ac-(1,5-cyclo)-[KARAD]-Ala-Ala-Ala-Rev_{1-17}-NH_2 (**3**) to test for steric interference and for experimental comparison of helix induction and RNA binding affinities.

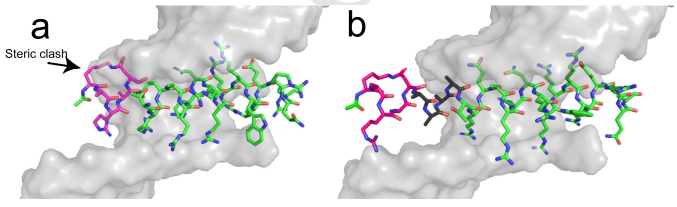


Figure 2. Models of N-terminally capped Rev peptides bound to the RRE. (**a**) compound **2;** (**b**) compound **3**. Rev_{34-50} is shown in green, Ac-(1,5-cyclo-[KARAD])- is in pink, and the -Ala-Ala-Ala- linker is in black.

Circular dichroism spectra (**Figure 3a**) showed that the Ac-(1,5-cyclo)-[KARAD]- unit was a very effective N-terminal helix nucleator, improving alpha helicity from 6% (compound **1**) to 54% (compounds **2** and **3**) in aqueous 10 mM phosphate buffer (pH 7.0, 22°C). Interestingly, helical induction by the N-terminal helix-constrained cyclic pentapeptide was the same with (compound **3**) or without (compound **2**) the Ala-Ala-Ala spacer between cyclic peptide and Rev peptide (**Figure 3a**). This induction of α-helicity using the N-terminal helix-constrained cyclic peptide also improved the binding affinity for RRE (**Figure 3b**) from IC_{50} 886 nM (**1**) to IC_{50} 260 nM (**2**). A further increase to IC_{50} 91 nM for **3**, despite comparable α-helicity to **2**, was attributed to more favourable fitting of **3** into RRE without the steric clash predicted earlier for **2** (**Figure 2a**).

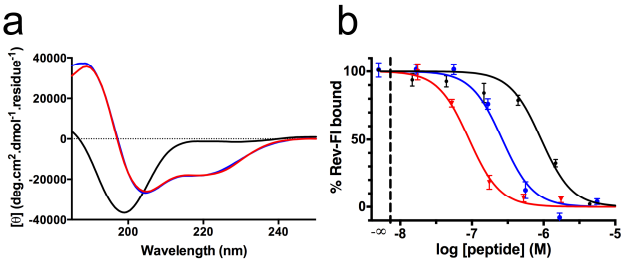


Figure 3. (a) Circular dichroism spectra of uncapped H-Rev_{1-17}-NH_2 peptide (**1**, black) versus N-terminally capped peptides Ac-(1,5-cyclo)-[KARAD]-Rev_{1-17}-NH_2 (**2**, blue) and Ac-(1,5-cyclo)-[KARAD]-Ala-Ala-Ala-Rev_{1-17}-NH_2 (**3**, red), measured in 10mM phosphate buffer (pH 7.0) at 22°C. **(b)** Competitive binding of N-terminal capped Rev peptides **1** (black), **2** (blue), **3** (red) for RRE using 10nM Suc-Rev_{1-17}-AAAAC(fluorescein)-NH_2 (K_d = 7 nM) and 7.5nM biotin-labelled RRE.

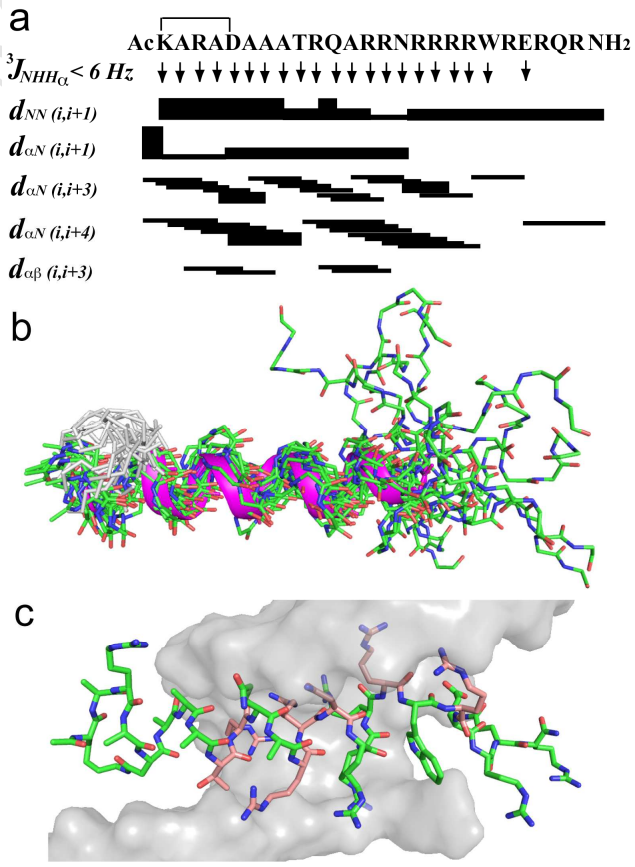


Figure 4. Structure of peptide **3** in water (N-terminal cycle on left). (**a**) Summary of sequential, short and medium range NOE data in 90% H_2O:10% D_2O at 298K. NOE intensities were strong

ACCEPTED MANUSCRIPT

(upper distance constraint 2.7Å), medium (3.5Å), weak (5.0Å), very weak (6.0Å) and proportional to bar thickness. $^3J_{NHCH\alpha}$ < 6Hz indicated by ↓. (**b**) 20 lowest energy structures with K(*i*)→D(*i*+4) lactam bridge in grey and backbone conformation in pink ribbon. Residues of the cyclic peptide, three spacer alanines, and Rev_{1-11} are in the helix, while Rev_{12-17} are disordered. (**c**) NMR structure of **3** modelled to bind in RRE. Residues important for binding and specificity are in orange (Thr1, Arg2, Arg5, Arg6, Asn7, Arg11, Arg13).

Using 2D ^{1}H-NMR spectroscopy, we determined a solution structure for **3** in water (**Figure 4**) that showed the N-terminal cyclic peptide plus the linker plus Rev_{1-11} in an alpha helical conformation, while Rev_{12-17} was disordered. This is consistent with CD spectra above where only 54% helix induction was found for **3** in water, NMR structures for **3** in 90% H_2O:10% D_2O at 298K (mixing time 200 ms) were calculated from 69 NOE-derived distance (short, medium, long range) and 21 $^3J_{NHH\alpha}$-dihedral angle restraints derived from NOESY spectra. Small $^3J_{NHH\alpha}$ coupling constants and several $d_{NN(i,i+1)}$, $d_{\alpha N(i,i+3)}$ and $d_{\alpha N(i,i+4)}$ NOEs were indicative of alpha helical structure for **3** in water (**Figure 4a**). The 20 lowest energy calculated structures (**Figure 4b**) showed no ϕ-dihedral angle (>5°) or distance (>0.2 Å) violations and relatively well-defined α-helical turns, frayed at the C-terminus with overall RMSD (residues 2 to 20) = 1.95 Å (Pymol software). The NMR-derived structure for **3** resembled Rev_{34-50} bound to RNA, with most of the important residues in **3** (Thr1, Arg2, Arg5, Arg6, Asn7, Arg11, Arg13; orange in **Figure 4c**) being in the alpha helix of **3** and corresponding to those in Rev_{34-50} that contact RNA.

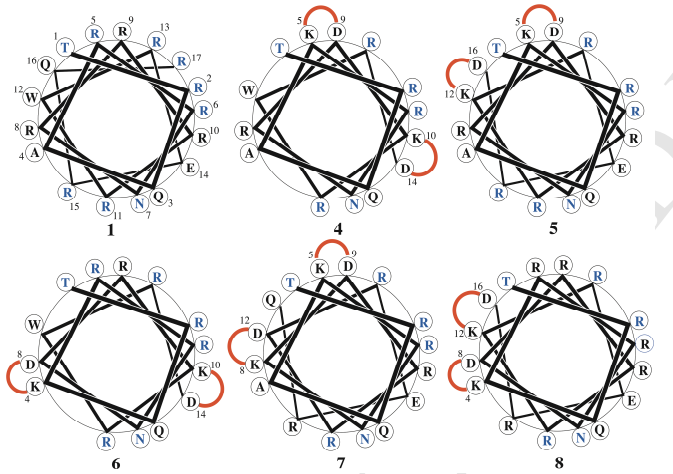


Figure 5. Helical wheel diagrams illustrating key residues at positions *i* and *i*+4 in the 17-residue sequence on the same face of an alpha helix. This was used along with Ala mutagenesis data[4a] to position K(*i*)→D(*i*+4) lactam bridges to constrain Rev peptides **4**-**8** into α-helices. Linkers in red. Residues in blue are important for binding and specificity for RRE.

An alternative to placing the helix-inducing cyclic pentapeptide constraint at the end of the Rev_{1-17} sequence is to incorporate it within the sequence. Because of the large interface between Rev_{1-17} and RRE (988Å^2 determined using the PISA server),[15] we were concerned that lactam bridges inserted into the Rev peptide might sterically impede binding to RRE. Modelling predictions suggested that this might be the case, but we proceeded anyway hopeful that peptide-RNA cooperativity might enable cyclic mimics to fit allowable space. Taking care to avoid replacing the important binding residues (Thr1, Arg2, Arg5, Arg6, Asn7 Arg11, Arg13, Arg15, Arg17), we designed (*i*, *i*+4) lactam bridges across positions 4-8, 5-9, 8-12, 10-14 and 12-16 (**Figure 5**). This led to five peptides (**4**-**8**) incorporating two differently spaced helix-constraining cyclic pentapeptide modules: Ac-(5,9; 10,14-cyclo)-TRQA[**K**RNR**D**][**K**RWR**D**]-NH_2 (**4**), Ac-(5,9; 12,16-cyclo)-TRQA[**K**RNR**D**]RR[**K**RER**D**]-NH_2 (**5**), Ac-(4,8; 10,14-cyclo)-TRQ[**K**RRN**D**]R[**K**RWR**D**]-NH_2 (**6**), Ac-(5,9; 8,12-cyclo)-TRQA[**K**RN[**KD**]RR**D**]RERQ-NH_2 (**7**) and Ac-(4,8; 12,16-cyclo)-TRQ[**K**RRN**D**]RRR[**K**RER**D**]-NH_2 (**8**). These all feature two cyclic pentapeptide modules, either back-to-back with no spacer (**4**) or spaced by one residue (**6**), two residues (**5**), three residues (**8**), or overlapping one another (**7**).

Circular dichroism spectra (**Figure 6a**) demonstrated that the cyclic pentapeptide dramatically increased α-helicity in all five compounds **4**-**8** (55-92%, **Table 1**) relative to **1** (6%), indicating effective induction of alpha helicity across and beyond the cyclised regions. The two most alpha helical peptides (**5**, **6**) had the highest competitive binding affinities for RRE (**Table 1**, **Figure 6b**), with IC_{50} 223 nM (**5**) and 440 nM (**6**). However, compounds **5** and **6** were not as potent as the N-terminal capped analogue **3**.

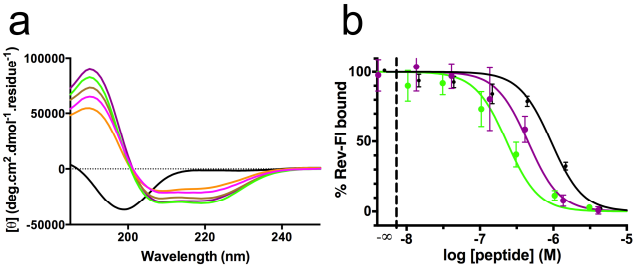


Figure 6. (a) Circular dichroism spectra of Rev_{34-50} peptide (**1**) and derivatives incorporating one or two -(1,5-cyclo)-[KXXXD**]**-motifs at various positions (**4**-**8**) measured in 10mM phosphate buffer (pH 7.0) at 22°C. Compounds **1** (black), **4** (magenta), **5** (green), **6** (purple), **7** (brown), **8** (orange**)**. (b) Binding isotherms for interaction with RRE-biotin. Competitive binding of peptides **1** (black), **5** (green), **6** (purple) for RRE using 10nM Suc-Rev_{1-17}-AAAAC(fluorescein)-NH_2 (K_d = 7 nM) and 7.5nM biotin-labelled RRE.

Compound **4**, Ac-(5,9-10,14-cyclo)-TRQA[**K**RNR**D**][**K**RWR**D**]-NH_2 had appreciable helicity (f_H = 0.64) but showed low binding affinity for RRE (IC_{50} 10 μM). Molecular modelling had suggested that the K(5) → D(9) and K(10) → D(14) lactam bridges in **4** are on two different faces of an α-helix (**Figure 5**) and potentially hinder interactions with RNA. To test this notion, we had synthesised compound **5**, Ac-(5,9-12,16-cyclo)-TRQA[**K**RNR**D**]RR[**K**RER**D**]-NH_2, where we altered the location of the second lactam bridge by placing two amino acids between the lactam bridges. This synchronizes the lactam bridges onto the same α-helical face in **5**, making them less likely to interfere with binding to RNA. This was evident from the 200-fold improvement in RRE-binding affinity for **5** (IC_{50} 223 nM) relative to **4** (IC_{50} 10 μM). Compound **6** with one residue between the cycles (IC_{50} 440 nM) was somewhat tolerated, but three alanines inserted between the cycles in compound **8** (IC_{50} 1660 nM) were not well tolerated. These observations were in agreement with modelling of these cyclic-constrained helices into the RNA, which revealed some steric clashes. Compound **7**, which overlaps with the cyclic regions, reduces the number of RNA-interacting residues and accordingly reduced affinity for RRE (IC_{50} 1820 nM).

The lower affinity for RRE of the 16-residue peptide **5** (IC_{50} 223 nM), containing two cyclic peptide constraints, over the 25-residue peptide **3** (IC_{50} 91 nM, **Table 1**) cannot be explained by differences in steric hindrance of lactam bridged side chains. Rather this difference is attributed to the required replacement of

ACCEPTED MANUSCRIPT

a key amino acid to cater for the lysine residue of the lactam bridge. In Ac-(5,9-12,16-cyclo)-TRQA[**K**RNR**D**]RR[**K**RER**D**]-NH_2 (compound **5**), Arg5 had to be replaced with Lys5 in order to synthesise the lactam bridge. An earlier alanine-scan study of Rev_{1-17} had shown that Arg5 has a specific role in binding to RRE and that the Arg5Ala mutant reduced specific binding 10-fold from the wild-type Rev_{1-17} peptide.[4a] Furthermore, the NMR structure of the Rev_{1-17}:RRE complex suggests that Arg5 makes bridging hydrogen bonds between U66 and G67 of the RRE.[10] This data is supported by earlier work showing that Arg5 cannot be functionally replaced by lysine.[4c] It therefore seems likely that replacement of Arg5 explains the reduction in binding affinity.

Table 1. Summary of competitive binding affinities and molar ellipticities for compounds **1**-**8** (10 mM PBS, pH 7.2, 298K).

Compound*	f_H	-log IC_{50} ± SEM	IC_{50} (nM)	ΔΔG
1	0.06	6.05 ± 0.05	886	--
2	0.54	6.59 ± 0.05	260	3.01
3	0.54	7.04 ± 0.06	91	-5.55
4	0.64	4.98 ± 0.07	10 000	5.94
5	0.92	6.65 ± 0.04	223	-3.83
6	0.87	6.36 ± 0.05	440	-1.72
7	0.79	5.74 ± 0.32	1 820	1.77
8	0.55	5.78 ± 0.07	1 660	1.54

*Maximum 10 μM tested. ΔΔG = -RTln[$IC_{50(1)}$/$IC_{50(2)}$] was determined in $kJ.mol^{-1}K^{-1}$ relative to compound **1**.

To confirm the extent and location of α-helicity in the cyclic-constrained peptide **5**, we determined its NMR-derived solution structure in water (**Figure 7**), calculated using a dynamic simulated annealing and energy minimisation protocol in X-PLOR, using 106 NOE-derived distance restraints (23 sequential, 67 medium-long range), 15 backbone ϕ-dihedral constraints and 10 hydrogen bonds. The structure of **5** confirmed its helical nature in water, as identified in the CD spectra, with Δδ αH chemical shifts all greater than -0.10 ppm and almost all residues in the cyclised region (residues 5-16) having Δδ αH values greater than -0.20 ppm, except for D9 and D16. We had previously seen this trend for aspartic acid residues in Ac-1,5-cyclo-[KXXXD]-NH_2 with decreased Δδ αH.[9,14] Furthermore, variable temperature data suggested that amide protons of Ala4, Lys6, Arg6, Arg8, Arg11, Lys12, Arg13, Arg15, Asp16 and the C-terminal amide were all involved in hydrogen bonds (**Figure 7**). Finally, the $^3J_{NHCH\alpha}$ coupling constants, with the exception of those for Asp16 and the terminal amide, were all less than 6 Hz. In combination, this data supports that all sixteen residues in compound **5** are in a helical conformation. 2D NOESY spectra were collected for compound **5** with sequential NOEs dominant, as well as significant medium-long range NOEs ($d_{\alpha N(i,i+3)}$, $d_{\alpha\beta(i,i+3)}$). Additional weaker NOEs ($d_{NN(i,i+2)}$, $d_{\alpha N(i,\ i+2)}$), and especially $d_{\alpha N(i,\ i+4)}$ were also observed throughout the peptide sequence, supporting the existence of a continuous α-helix (**Figure 7**).

The 20 lowest energy calculated structures (**Figure 8**) did not have any ϕ-dihedral angle (>5°) or distance (>0.3Å) violations and showed relatively well-defined α-helical turns with overall RMSD (residues 2 to 15) = 0.470 Å (Pymol software, 54 of 56 atoms used in calculation). The constraints are thus able to transfer α-helicity to those residues 1-4 and 10-11 outside of the cyclic regions. Compound **5** also closely aligns with the RRE-bound conformation of the Rev peptide, with its superposition on the lowest energy conformation of the RRE:Rev NMR structure (pdb: 1ETF) showing an overall RMSD = 0.755 Å (using residues 35-48 of pdb:1ETF and residues 2-15 of compound **5**).

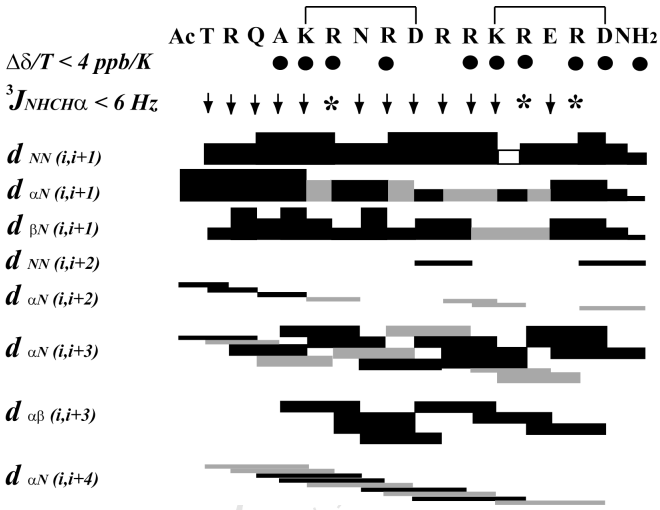


Figure 7. Summary of sequential, short and medium range NOE data for **5** in 90% H_2O 10% D_2O at 298K. NOE intensities were classified as strong (upper distance constraint 2.7Å), medium (3.5Å), weak (5.0Å), very weak (6.0Å) and are proportional to bar thickness; grey bars indicate overlapping signals and empty bar for NOE not observed due to proximity to diagonal. $^3J_{NHCH\alpha}$ < 6Hz are indicated by ↓ and ∗ for broad signals. Temperature insensitive amide NHs for which chemical shifts changed by < 4ppb/K are indicated by •.

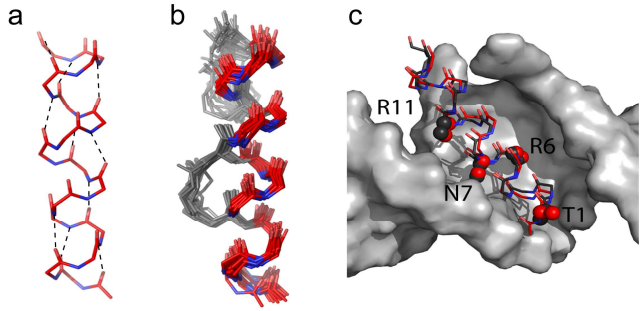


Figure 8. Lowest energy calculated structures for **5**. (**a**) Lowest energy structure showing proposed hydrogen bonds. (**b**) 20 lowest energy structures showing position of the K(*i*) → D(*i*+4) lactam bridges (grey). (**c**) Superimposition of compound **5** (red) on H-Rev_{1-17}-OH (black) bound to RRE (pdb 1ETF) showing the Cα-Cβ bonds of Thr1, Arg6, Asn7 and Arg11 as spheres.

In terms of free energy, the data in Table 1 suggest that at least 6 $kJ.mol^{-1}.K^{-1}$ can be gained by constraining the peptide into an α-helical conformation (this may be a maximum value), provided that the helix-inducing constraint is appropriately positioned. In this particular interaction, the RNA may also assist peptide folding through electrostatic interactions between the negatively charged phosphate backbone and positively charged arginine side chains.

3. Conclusion

In summary, we have used Rev-RNA structures to select appropriate locations for incorporating helix-constrained cyclic peptides into the RNA-binding domain of HIV-1 Rev peptide, the

ACCEPTED MANUSCRIPT

17-residue sequence TRQARRNRRRRWRERQR-NH_2. This peptide alone was not helical in water but has helical propensity, since it becomes helical when bound to the RNA segment known as RRE. We have demonstrated that helix-constrained cyclic pentapeptides are able to nucleate extensive alpha helicity in Rev peptides, even in the highly polar solvent water which competes very effectively for hydrogen bonding partners. When appended to the N-terminus of the isolated Rev peptide sequence, a cyclic pentapeptide was most effective in promoting binding to the RRE target when it was spaced three or more residues from the N-terminus. This spacing minimises steric clashes between the appended cyclic peptide helix and the RNA bases. Alternatively, the helix-constrained cyclic pentapeptides can be incorporated as modular helical templates within the Rev peptide sequence, also conferring increased alpha helicity and higher affinity binding to the RRE. Calculations suggest that this helix pre-organization in water confers an advantage of 2 to 6 $kJ.mol^{-1}.K^{-1}$ to the Rev peptide for binding to RNA.

4. Experimental section

Synthesis of peptides 1-8. Peptides were synthesized manually by standard solid phase methods using HBTU/DIPEA activation of Fmoc on Rink Amide resins (substitution 0.30-0.50 mmol/g, 0.2-0.5 mmol scale, 0.50-1.00 g resin). Four equivalents of Fmoc-protected amino acid, 4 equivalents of HBTU, and 4 equivalents of DIPEA were used in each coupling. Fmoc deprotection involved 2 × 3 min treatments with excess 1:1 piperidine:DMF. Coupling yields were monitored by quantitative ninhydrin assay. The phenylisopropyl ester of aspartic acid and methyltrityl group of lysine were removed by treating the peptide resin with 3% TFA in DCM (5 × 2 min). The resin was neutralised by washing with 5% DIPEA in DMF (2 x 3min). Cyclization was effected on-resin using BOP, DIPEA in DMF. The peptide resin was then washed with DMF, MeOH/DCM and DCM, dried under nitrogen with suction for 20 min. Peptides were cleaved using 95% TFA, 2.5% TIPS, 2.5% H_2O. Peptides were precipitated with diethyl ether and then decanted to give white solids that were re-dissolved in 1:1 acetonitrile/water and lyophilised. The crude peptides were purified by rp-HPLC (R_t:Vydac C18 column, 300 Å. 22 × 250mm, 214nm, Solvent A = 0.1% TFA in H_2O, Solvent B = 0.1% TFA, 10% H_2O in acetonitrile. Gradient: 0% B to 100% B over 30 mins).

H-TRQARRNRRRRWRERQR-NH_2 (**1**):. R_t: 14.25 min. MS $[M+H]^+$ = 2437.69 (found), 2436.41 (calc.). Ac-(1,5-cyclo)-[KARAD]TRQARRNRRRRWRERQR-NH_2 (**2**): R_t: 13.8 min. MS $[M+6H]^{6+}$ = 501.09 (found), 501.12 (calc.). Ac-(1,5-cyclo)-[KARAD]AAATRQARRNRRRRWRERQR-NH_2 (**3**): R_t: 15.5 min. MS $[M+6H]^{6+}$ = 536.58 (found), 536.64 (calc.). Ac-(5,9; 10,14-cyclo)-TRQA[KRNRD][KRWRD]-NH_2 (**4**): R_t: 17.2 min. MS $[M+2H]^{2+}$ = 946.00 (found), 946.02 (calc.). Ac-(5,9; 12,16-cyclo)-TRQA[KRNRD]RR[KRERD]-NH_2 (**5**) R_t: 13.3 min. MS $[M+2H]^{2+}$ = 1102.51 (found), 1102.61 (calc.). Ac-(4,8; 10,14-cyclo)-TRQ[KRRND]R[KRWRD]-NH_2 (**6**) R_t: 13.6 min. MS $[M+2H]^{2+}$ = 988.50 (found), 988.55 (calc.) . Ac-(5,9; 8,12-cyclo)-TRQA[KRN[KD]RRD]RERQ-NH_2 (**7**) R_t: 14.2 min. MS $[M+3H]^{3+}$ = 707.03 (found), 706.73 (calc.). Ac-(4,8; 12,16-cyclo)-TRQ[KRRND]RRR[KRERD]-NH_2 (**8**) R_t: 14.4 min. MS $[M+3H]^{3+}$ = 744.40 (found), 744.43 (calc.)

Circular Dichroism Spectroscopy. CD experiments were performed on a Jasco model J-810 spectropolarimeter calibrated with (1S)-(+)-10-camphorsulfonic acid. Spectra were recorded in a 0.1cm Jasco cell between 260-185 nm at 50 nm/min with a bandwidth of 1.0 nm, response time of 2 s, resolution step width 0.1 nm and sensitivity 20, 50 or 100 mdeg. Each spectrum is the average of 5 scans with smoothing to reduce noise. Peptide samples were dissolved in 1 mL of 18MΩ distilled water (~1-4 mg/mL). Each stock solution (400 µL) was then diluted 1:10 with phosphate buffer (10 mM, pH 7.2), while the remaining 600 µL was kept for concentration determination. Solutions were then prepared for each sample with final concentrations 40 – 400 µM in 10 mM sodium phosphate buffer (pH 7.2).

Concentrations were determined using the PULCON method.[16] Pulses (90^o) were accurately determined, then 1D spectra were acquired using standard Watergate sequence with ns = 32-64, rg = 128-512, d1 = 30 sec. Several fully resolved signals were integrated and used to calculate concentration from the equation:

$$c_U = c_R \frac{S_U\, T_U\, \theta^U_{360^\circ}\, n_R\, rg_R}{S_R\, T_R\, \theta^R_{360^\circ}\, n_U\, rg_U}$$

where *c* is concentration, *S* is integral (in absolute units/number of protons), *T* is temperature in Kelvin, θ_{360^o} is 360^o rf pulse, *n* is the number of scans, and *rg* is receiver gain used for measuring the reference (*R*) and unknown (*U*) samples.

CD ellipticity (in millidegrees) was converted to residue molar ellipticity ($deg.cm^2.dmol^{-1}.residue^{-1}$) using $[\theta] = \theta_{raw}/(10 \times C \times N \times l)$, where θ_{raw} is ellipticity in millidegrees, C is peptide concentration (mol/L), l is optical path length of cell (cm), and N is number of peptide units.

Percent Helicity of peptides was calculated from residue-molar elipticity at λ = 222 using eq: $f_{helix} = ([\theta]_{222} - [\theta]_0)/([\theta]_{max} - [\theta]_0)$. $[\theta]_{max}$ ($[\theta]_{max} = [\theta]_\infty(n - x)/n$) is the maximum theoretical mean residue ellipticity for a helix of n residues, $[\theta]_\infty$ is the mean residue ellipticity of an infinite helix, and x is an empirical constant that can be interpreted as the effective number of amides missing as a result of end effects, usually about 2.4-4 (we used x=3) and $[\theta]_\infty$ = (-44000 + 250T) (T is temperature of the peptide solution in ˚C). $[\theta]_0$ is the mean residue ellipticity of the peptide in random coil conformation and equals to (2220 – 53T) and $[\theta]_{222}$ ($[\theta]_{222}$ = 1/n . $[\theta_{obs}]/(10 \times l \times C)$) is the observed residue ellipticity of peptide at 222 nm. θ_{obs} = measured ellipticity in mdeg; n = number of peptide residues; C = sample concentration (mol/L); l = optical path length of the cell in cm.

NMR Spectroscopy. Peptides (1-2 mg) were dissolved in 600 µL of H_2O:D_2O (9:1) or H_2O:CD_3CN (1:1) at pH 5.0. 1D and 2D ^{1}H-NMR spectra were recorded on a Bruker Avance DRX-600 spectrometer. 2D ^{1}H-spectra were recorded in phase-sensitive mode using time-proportional phase incrementation for quadrature detection in the *t*1 dimension. The 2D experiments included TOCSY (standard Bruker mlevgpph pulse program), ROESY (standard Bruker roesygpph pulse program), NOESY and dqfCOSY (standard Bruker dqfcosygpph pulse program). TOCSY spectra were acquired over 6887 Hz with 4096 complex data points in *F2*, 512 increments in *F1* and 32 scans per increment. NOESY spectra were acquired over 6887 Hz with 4096 complex data points in *F2*, 512 increments in *F1* and 32 scans per increment. TOCSY and NOESY spectra were acquired with several isotropic mixing times of 80, 100 ms for TOCSY and 200-300 ms for NOESY. Water suppression was achieved using modified WATERGATE. For 1D ^{1}H NMR spectra acquired in H_2O/D_2O (9:1), the water resonance was suppressed by low power irradiation during the relaxation delay (1.5 to 3.0 s). The variable NMR experiments were performed over the range of 278-318K. Spectra were processed using Topspin (Bruker, Germany) software and NOE intensities were collected manually. The *t*1 dimensions of all 2D spectra were zero-filled to 1024 real data points with 90˚ phase-shifted QSINE bell window

ACCEPTED MANUSCRIPT

functions applied in both dimensions followed by Fourier transformation and fifth order polynomial baseline correction. ^{1}H chemical shifts were referenced to DSS (δ 0.00 ppm) in water. $^{3}J_{NHCH\alpha}$ coupling constants were measured from 1D ^{1}H NMR and dqf-COSY spectra using XPLOR program.

Structure Calculations. Distance restraints used in calculating the structure for **5** in water were derived from NOESY spectra (recorded at 298K) using a mixing time of 200 ms. NOE cross-peak volumes were classified manually as strong (upper distance constraint ≤ 2.7Å), medium (≤ 3.5Å), weak (≤ 5.0Å) and very weak (≤ 6.0Å). Standard pseudoatom distance corrections were applied for non-stereospecifically assigned protons. To address the possibility of conformational averaging, intensities were classified conservatively and only upper distance limits were included in the calculations to allow the largest possible number of conformers to fit the experimental data. Backbone dihedral angle restraints were inferred from amide $^{3}J_{NHCH\alpha}$ coupling constants in 1D spectra, ϕ was restrained to –65 ± 30° for $^{3}J_{NHCH\alpha}$ ≤ 6Hz and to -120 ± 30° for $^{3}J_{NHCH\alpha}$ ≥ 8Hz. There was clearly no evidence for *cis*-amides about peptide bonds (i.e. no CHα-CHα (*i, i+1*) NOEs) in the NOESY spectra so all ψ-angles were set to *trans* (ψ = 180°). Starting structures with randomised ϕ and ψ angles and extended side chains were generated using an *ab initio* simulated annealing protocol. The calculations were performed using the standard forcefield parameter set (PARALLHDG5.2.PRO) and topology file (TOPALLHDG5.2.PRO) in XPLOR-NIH with in house modifications to generated lactam bridges between lysine and aspartic acid residues. Refinement of structures was achieved using the conjugate gradient Powell algorithm with 4000 cycles of energy minimisation and a refined forcefield based on the program CHARMm.[17] Structures were visualised with Pymol and analysed for distance (>0.2Å) and dihedral angle (>5°) violations using noe.inp (in Xplor) files. Final structures contained no distance violations (>0.2Å) or angle violations (>5°).

Molecular modelling. Three-dimensional models of **2** and **3** were built by using Insight II[18] to covalently link the NMR structure of Ac-(1,5-cyclo)-[KARAD]-NH_2 [9b] to the linear Rev peptide (idealised alpha helix) with and without spacer sequences. All images of peptides incorporating cyclic constraints were visualised using MacPymol[19] or Oemga2 (OpenEye Scientific Software). [20]

RRE-Rev Affinity and Competitive Binding. K_d was determined by adapting a reported assay,[21] using 10nM fluorescein-labelled Rev (succinyl-TRQARRNRRRRWRER-QRAAAARC-fluorescein) and increasing concentrations of RRE-biotin 5'-GGU AUG GGC GCA GCG CAA GCU GAC GGU ACA GGC C-3'-biotin (Ambion, USA) in assay buffer (30mM HEPES pH 7.5, 100 mM KCl, 40 mM NaCl, 10 mM ammonium acetate, 10 mM guanidinium HCl, 2 mM $MgCl_2$, 0.5 mM EDTA, 50 µg/ml *E.coli* tRNA (Sigma) and 0.01% Igepal). 10µg AlphaScreen® FITC Acceptor beads (Perkin Elmer) were added and the plate was incubated for 30 minutes. 10 µg AlphaScreen® FITC Donor beads (Perkin Elmer) were added to each well and the plate was incubated for a further 30 minutes and read using the Envision® (Perkin Elmer). For competition assays, 10 nM suc-TRQARRNRRRRWRERQRAAAARC-fluorescein and 7.5 nM RRE-biotin were incubated with increasing concentrations of unlabelled peptides. Data was fitted using equation: Y= [bottom + (top – bottom)]/[1 + $10^{((\log EC_{50} - X)*hillslope)}$], where EC_{50} can also be IC_{50}, depending on experiment.

Free energy calculations (ΔΔG binding). Delta delta values of binding were calculated using the standard formulae ΔG = -RTlnKeq and ΔΔG = -RTln[$K_{eq(1)}/K_{eq(2)}$]. Although we do not have K_{eq} values for all our compounds, we were able to use the IC_{50} values obtained from the competitive binding data. We can use the assumption that the ratio of $K_{eq(1)} / K_{eq(2)} \approx IC_{50\,(1)} / IC_{50\,(2)}$ as illustrated in the literature.[22] The IC_{50} values are comparable to each other as they were obtained under the same experimental conditions. We and others[23] have used the IC_{50} values to calculate differences in the free energy of binding by combining equation to ΔΔG = -RTln[$IC_{50(1)}/IC_{50(2)}$].

Acknowledgments

We thank the Australian Research Council for fellowship (FF0668733 to DF) and grant (DP1096290) funding, and the Australian National Health and Medical Research Council for a Senior Principal Research Fellowship (1027369 to DF) and grant (511194) funding.

References and notes

(1) (a) Thomas, J. R.; Hergenrother, P. J. *Chem. Rev.* **2008**, *108*, 1171. (b) Guan, L.; Disney, M. D. *ACS Chem. Biol.* **2012**, *7*, 73. (c) Mehta, A.; Sonam, S.; Gouri, I.; Loharch, S.; Sharma, D. K.; Parkesh, R. *Nucleic Acids Research* **2013**, *42*, D132.

(2) Cook, A.; Bono, F.; Jinek, M.; Conti, E. *Annu. Rev. Biochem.* **2007**, *76*, 647.

(3) (a) Zapp, M. L.; Green, M. R. *Nature* **1989**, *342*, 714-6. (b) Daly, T. J.; Cook, K. S.; Gray, G. S.; Maione, T. E.; Rusche, J. R. *Nature* **1989**, *342*, 816. (c) Hung, L-W.; Holbrook, E. L.; Holbrook, S. R. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5107. (d) Fang, X.; Wang, J.; O'Carroll, I. P.; Mitchell, M.; Zuo, X.; Wang, Y.; Yu. P.; Liu. Y.; Rausch, J. W.; Dyba, M. A.; Kjems. J.; Schwieters, C. D.; Seifert, S.; Winans, R. E.; Watts, N. R.; Stahl, S. J.; Wingfield, P. T.; Byrd, R. A.; Le Grice, S. F. J.; Rein, A.; Wang, Y-X. *Cell* **2013**, *155*, 594.

(4) (a) Tan, R.; Chen, L.; Buettner, J. A.; Hudson, D.; Frankel, A. D. *Cell* **1993**, *73*, 1031. (b) Kjems, J.; Calnan, B. J.; Frankel, A.D.; Sharp, P.A. *EMBO J* **1992**, *11*, 1119. (c) Tan, R.; Frankel, A. D. *Biochemistry* **1994**, *33*, 14579.

(5) (a) Harada, K.; Martin, S. S.; Tan, R.; Frankel, A. D. *Proc. Natl. Acad. Sci. U S A* **1997**, *94*, 11887. (b) Tan, R.; Frankel, A. D. *Proc. Natl. Acad. Sci. U S A* **1998**, *95*, 4247. (c) Mills, N. L.; Daugherty, M. D.; Frankel, A. D.; Guy, R. K. *J. Am. Chem. Soc.* **2006**, *128*, 3496.

(6) (a) Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C. H. *J. Am. Chem. Soc.* **1997**, *119*, 3641. (b) Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. *Biochemistry* **2000**, *39*, 5630. (c) Kirk, S. R.; Luedtke, N. W.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 980.

(7) (a) Chaloin, L.; Smagulova, F.; Hariton-Gazal, E.; Briant, L.; Loyter, A.; Devaux, C. *J. Biomed. Sci.* **2007**, *14*, 565. (b) Held, D. M.; Kissel, J. D.; Patterson, J. T.; Nickens, D. G.; Burke, D. H. *Front. Biosci.* **2006**, *11*, 89.

(8) (a) Chapman, R. L.; Stanley, T. B.; Hazen, R.; Garvey, E. P. *Antiviral. Res.* **2002**, *54*, 149. (b) Shuck-Lee, D. *et al*. *Antimicrob. Agents Chemother.* **2008**, *52*, 3169. (c) González-Bulnes L, Ibáñez I, Bedoya LM, Beltrán M, Catalán S, Alcamí J, Fustero S, Gallego J. *Angew. Chem. Int. Edit. Engl.* **2013**, *52*, 13405.

(9) (a) Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. *J. Am. Chem. Soc.* **2005**, *127*, 2974. (b) Shepherd, N. E.; Abbenante, G.; Fairlie, D. P. *Angew. Chem. Int. Edit.* **2004**, *43*, 2687.

ACCEPTED MANUSCRIPT

(10) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, *273*, 1547.
(11) Pollard, V. W.; Malim, M. H. *Annu. Rev. Microbiol.* **1998**, *52*, 491.
(12) (a) Kubota, S.; Siomi, H.; Satoh, T.; Endo, S.; Maki, M.; Hatanaka, M. *Biochem. Biophys. Res. Commun.* **1989**, *162*, 963. (b) Perkins, A.; Cochrane, A. W; Ruben, S. M.; Rosen, C. A. *J. Acquir. Immune Defic. Syndr.* **1989**, *2*, 256.
(13) Scanlon, M. J.; Fairlie, D. P.; Craik, D. J.; Engelbretsen, D. R.; West, M. L. *Biochemistry* **1995**, *34*, 8242.
(14) (a) de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, G.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P. *Angew. Chem. Int. Edit.* **2014**, *53*, 6965. (b) Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gomez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P. *Proc. Natl. Acad. Sci.* **2010**, *107*, 11686. (c) Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. *J. Am. Chem. Soc.* **2009**, *131*, 15877. (d) Rao, T.; Ruiz-Gómez, G.; Hill, T. A.; Hoang, H. N.; Fairlie, D. P.; Mason, J. M. *PLoS One*, **2013**, *8(3)*, e59415. (e) Shepherd, N. E.; Harrison, R. S.; Fairlie, D. P. *Curr Drug Targets*, **2012**, *13*, 1348. (f) Harrison, R. S.; Ruiz-Gómez, G.; Hill, T. A.; Chow, S. Y.; Shepherd, N. E.; Lohman, R. J.; Abbenante, G.; Hoang, H. N.; Fairlie, D. P. *J. Med. Chem.* **2010**, *53*, 8400.
(15) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372*, 774.
(16) Dreier, L.; Wider, G. *Magn. Reson. Chem.* **2006**, *44*, S206.
(17) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(18) InsightII, Edit. v 2000 (Accelrys Inc., 2002).
(19) Delano, W. L. (DeLano Scientific LLC South San Francisco, CA, USA; 2006).
(20) Omega2, Version 2.3.2, OpenEye Scientific Software Inc., (Santa Fe, NM, USA), 2008; http://www.eyesopen.com/omega.
(21) Mills, N.L.; Shelat, A.A.; Guy, R. K. *J. Biomol. Screen.* **2007**, *12*, 946.
(22) Cheng, Y.; Prusoff, W.H. *Biochem. Pharmacol.* **1973**, *22*, 3099.
(23) Plount Price, M. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 9455.

ACCEPTED MANUSCRIPT