



Small molecule inhibitors of HIVgp41 N-heptad repeat trimer formation [☆]



William J. Allen ^a, Hyun Ah Yi ^b, Miriam Gochin ^{c,d}, Amy Jacobs ^b, Robert C. Rizzo ^{a,e,f,*}

^a Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook, NY 11794, United States

^b Department of Microbiology and Immunology, State University of New York at Buffalo, Buffalo, NY 14214, United States

^c Department of Basic Sciences, Touro University—California, Mare Island, Vallejo, CA 94592, United States

^d Department of Pharmaceutical Chemistry, University of California San Francisco, CA 94143, United States

^e Institute of Chemical Biology & Drug Discovery, Stony Brook University, Stony Brook, NY 11794, United States

^f Laufer Center for Physical & Quantitative Biology, Stony Brook University, Stony Brook, NY 11794, United States

ARTICLE INFO

Article history:

Received 25 February 2015

Revised 16 April 2015

Accepted 20 April 2015

Available online 28 April 2015

Keywords:

HIV

gp41

Docking

Virtual screening

Footprint similarity

DOCK

ABSTRACT

Identification of mechanistically novel anti-HIV fusion inhibitors was accomplished using a computer-aided structure-based design approach with the goal of blocking the formation of the N-heptad repeat (NHR) trimer of the viral protein gp41. A virtual screening strategy that included per-residue interaction patterns (footprints) was employed to identify small molecules compatible with putative binding pockets at the internal interface of the NHR helices at the core native viral six-helix bundle. From a screen of ~2.8 million compounds using the DOCK program, 120 with favorable energetic and footprint overlap characteristics were purchased and experimentally tested leading to two compounds with favorable cell–cell fusion (IC₅₀) and cytotoxicity profiles. Importantly, both hits were identified on the basis of scores containing footprint overlap terms and would not have been identified using the standard DOCK energy function alone. To our knowledge, these compounds represent the first reported small molecules that inhibit viral entry via the proposed NHR-trimer obstruction mechanism.

© 2015 Elsevier Ltd. All rights reserved.

Human immunodeficiency virus-1 (HIV-1) is a significant global health threat, with upwards of 35 million people currently infected.¹ Drugs against a range of viral targets (e.g., reverse transcriptase and protease) and highly-active anti-retroviral therapies (HAARTs) have made a significant impact against HIV infection, but can suffer from problems of resistance and availability.^{2,3} Given the continued importance of this virus, it is essential to identify new mechanisms to halt infection and viral replication. There are several established mechanisms to inhibit fusion between HIV and its host, many of which target the viral protein gp41 during the pre-hairpin intermediate state.^{4–6} Peptide inhibitors derived from the C-heptad repeat (CHR) sequence of gp41, such as T20,^{7,8} C34,⁹ and Sifuvirtide,¹⁰ selectively bind the surface of the N-heptad repeat (NHR) trimer and block six-helix bundle formation (Fig. 1). Small molecules that bind a highly conserved hydrophobic pocket,^{6,11,12} also located on the outside surface of the NHR trimer, reportedly inhibit fusion via the same mechanism. Alternatively, peptides derived from the NHR sequence, such as

N36 and N36^{Mut(e.g.)}¹³ can block virus-host fusion by selectively binding in the groove formed by two NHR helices,^{13–17} thus obstructing association of the third NHR helix and subsequent six-helix bundle formation (Fig. 1).

Peptide inhibitors based on the NHR sequence have low bioavailability, suffer from aggregation,¹⁹ and therefore are not themselves viable drug candidates. However, their inhibitory capability provides important proof-of-concept for a mechanism of virus-host fusion inhibition that has not been well-explored. Further, to our knowledge, small molecule inhibitors that mimic the mechanism of NHR-derived peptides have not yet been reported. Drugs capable of disrupting NHR trimer formation would be mechanistically novel and a valuable addition to currently available treatments. The goal of the present work is the identification of small molecule leads that inhibit NHR trimer formation using a combined computational and experimental screening approach. Computational modeling methods, including virtual screening and molecular dynamics simulations, broadly termed ‘computer-aided structure-based design,’ are a useful approach given the advent of new structural information and increasing computing power. Specifically, we employed a computer-aided structure-based design strategy to identify key interaction energy patterns made at the NHR trimer interface (termed footprints),^{20,21}

[☆] This manuscript is dedicated to Professor Iwao Ojima on the occasion of his 70th birthday.

* Corresponding author. Tel.: +1 631 632 9340; fax: +1 631 632 8490.

E-mail address: rizzorc@gmail.com (R.C. Rizzo).

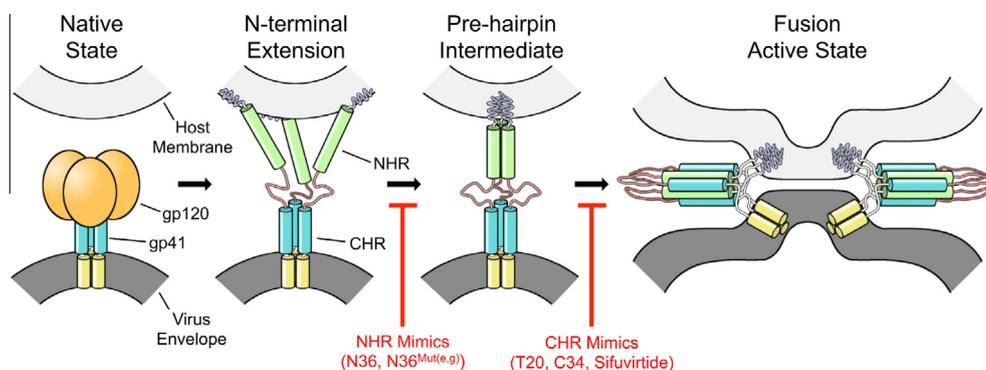


Figure 1. Model of HIV-host fusion. In the native state, viral proteins gp41 (blue/yellow) and gp120 (orange) are shown on the outside surface of the virus envelope (gp41 C-terminal domain not shown). Following receptor and co-receptor recognition by gp120, the N-terminal residues of gp41 extend and bind to the host cell membrane. The NHR helices trimerize to form the pre-hairpin intermediate. The fusion active state is reached following six-helix bundle formation, which leads to fusion of the two membranes. NHR mimics block transition from N-terminal extension to pre-hairpin intermediate states. CHR mimics block transition from pre-hairpin intermediate to the fusion active states. Figure adapted from Allen and Rizzo¹⁸ and Bewley et al.¹³

performed two virtual screens of approximately 1.4 million compounds each to enrich for small molecules capable of mimicking those interaction patterns, and experimentally evaluated 120 top compounds in terms of anti-fusion activity and cytotoxicity.

To identify which specific residues of the NHR are most important for trimer formation and stabilization, we employed the gp41 ectodomain model reported by McGillick et al.,²² which was based on earlier work by Caffrey (PDB 1IF3)²³ and includes all gp41 residues from the fusion peptide to the loop region (amino acids 1–94). Beginning from this construct, we treated one of the three NHR helices as a ‘ligand’ (Fig. 2, orange) and the other two assembled NHR helices as a ‘receptor’ (Fig. 2, gray). We then performed 20-ns all-atom molecular dynamics (MD) simulations in triplicate

using Gromacs 5.0.2²⁴ to compute time-averaged van der Waals steric packing (VDW) and electrostatic (ES) interactions on a per-residue basis across the ligand–receptor interface (Fig. 2). Briefly, MD simulations in the canonical NPT ensemble (constant number of atoms, pressure, and temperature) employed a 2-fs time step, the TIP3P force field for waters,²⁵ and the Amber99SB force field for protein.²⁶ Large negative peaks in the plots in Figure 2 indicate residues and regions on the gp41 NHR that interact favorably at the trimer interface.

The MD simulations suggest the fusion peptide (residues 1–16) and host-membrane proximal regions (residues 17–29) at the NHR interface include few significant inter-helical interactions, and particularly a lack of ES type interactions (Fig. 2). However, the strong favorable interactions beginning at amino acid 34 (Fig. 2, magenta and cyan brackets) likely indicate regions of importance with regards to NHR trimer formation. To determine whether these areas of interaction also support ‘pockets’ for potential small molecule binding, we removed one NHR helix from the trimer, then generated a surface of the remaining two NHR helices with the program DMS,²⁷ and finally identified cavities and areas of high local curvature with the DOCK utility sphgen.²⁸ When the outcome was compared with the energetic analysis in Figure 2, two putative, promising pockets emerged (Fig. 3A–B). The overlapping pockets are termed QLIQ (corresponding to region 1 in Fig. 2) and IQLT (corresponding to region 2 in Fig. 2), in which the nomenclature refers to the identity of the amino acid side chains from the NHR peptide that interpolate into each pocket. In addition to making favorable interactions within the newly-identified pockets, these specific groups of residues were chosen because they are spatially localized, thus they are able to act as a molecular reference to guide the virtual screens. Importantly, residues in these pockets are highly conserved across clinically relevant HIV strains,²⁹ and therefore are attractive regions to target.

Following identification of the QLIQ and IQLT pockets at the interface of two NHR helices, we performed large-scale virtual screens with the program DOCK 6.6.^{30,31} Following standard practices and protocols previously reported,^{12,32,33} we independently screened to each of the two pockets commercially-available small molecules from the ChemDiv vendor as listed on the ZINC database.³⁴ Docked molecules were scored with the standard DOCK grid energy (DGE) score, and then energy minimized and re-scored with the DOCK Cartesian energy (DCE) score. Finally, molecules were re-scored again with the DOCK footprint similarity (FPS) score,^{20,21} which quantifies the degree to which docked small molecules mimic the VDW and ES interactions between the receptor and a user-provided reference molecule. In the present study,

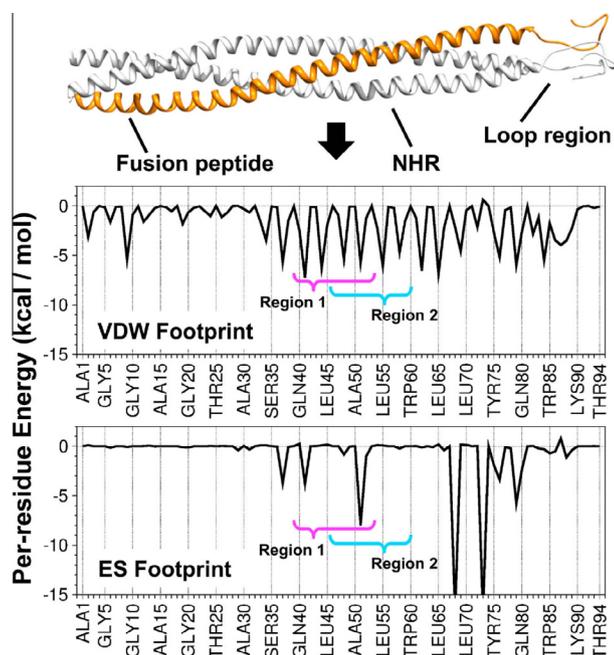


Figure 2. (Top) HIVgp41 N-terminal ectodomain model residues 1–94 (including the fusion peptide, NHR, and loop region) represented in the pre-hairpin state from McGillick et al.²² Orange peptide is treated as a ‘ligand’; gray peptides are treated as a single ‘receptor’. (Bottom) per-residue non-bonded interaction energies between individual amino acids from the ‘ligand’ and the ‘receptor’ as a whole. Interactions are divided into van der Waals (VDW) and electrostatic (ES) types based on the Lennard–Jones and Coulomb potentials, respectively. Data points averaged over 20-ns simulations and three replicates. Large negative peaks indicate areas of favorable interaction. Magenta and cyan brackets indicate approximate pocket regions.

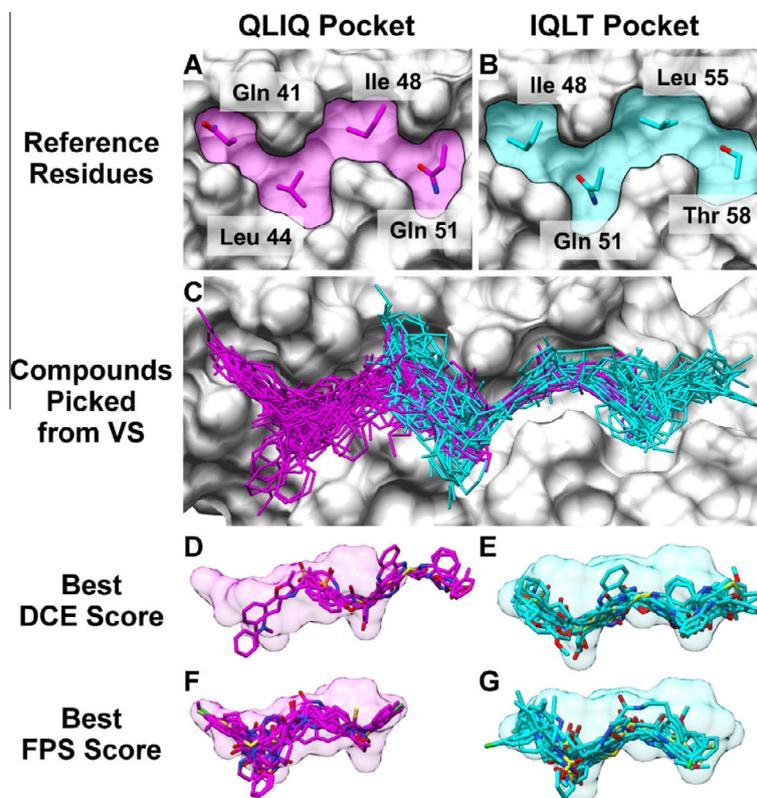


Figure 3. Two overlapping pockets at the interface of two NHR peptides are shown. (A) The QLIQ and (B) IQLT pockets are named for the amino acids of the substrate peptide that interpolate the space. The NHR surface is colored gray, the pocket regions are colored magenta (QLIQ) and cyan (IQLT). (C) All 120 molecules selected from the virtual screen shown overlaid in the two pockets. Molecules chosen from the QLIQ screen are shown in magenta, and molecules chosen from the IQLT screen are shown in cyan. Compounds chosen on the basis of their (D, E) DOCK Cartesian energy (DCE) score and (F, G) footprint similarity sum (FPS_{Sum}) score are shown overlaid with the reference peptide surface.

depending on the pocket the reference provided was either the QLIQ or IQLT native residues connected by a peptide backbone. We postulate that small molecules which mimic key interactions at the NHR trimer interface can disrupt the fusion process via disruption of NHR trimer formation, similar to the proposed mechanism for *N*-peptide based fusion inhibition.^{13–17}

To increase diversity, the 100,000 top-ranked docked and scored molecules (as determined by DCE score) in each of the pockets were clustered by similarity using Molecular Access System (MACCS) structural keys³⁵ as implemented in the Molecular Operating Environment (MOE) program.³⁶ Five ranked lists were generated from the cluster heads of each screen (10 lists total) based on the most favorable: (1) DCE scores, (2) VDW-type FPS (FPS_{VDW}) scores, (3) ES-type FPS (FPS_{ES}) scores, (4) the sum of both FPS scores (FPS_{Sum}), and (5) the sum of DCE and FPS_{Sum} scores (called TotalScore). Top compounds from each list were then visually inspected in the context of the corresponding pocket, and 120 compounds (between 6 and 18 from each scoring method) were purchased for experimental evaluation (Fig. 3C and Table 1). Importantly, some compounds appeared on multiple lists, meaning they were ranked favorably by multiple metrics, and some preference was given to those molecules over other molecules that only appeared on one list. Several descriptors were also computed using MOE including the predicted solubility, number of chiral centers, Lipinski-type properties,³⁷ among others, for consideration during the compound-picking phase. In general, compounds with extreme descriptors (e.g., >3 chiral centers, >3 Lipinski violations, >15 rotatable bonds, etc.) were eliminated although no firm rules were employed and molecules were considered on a case-by-case basis.

Presented in Table 1 are the average molecular weights, number of rotatable bonds, DCE score, and FPS_{Sum} score for the final group of molecules purchased from the virtual screens. As expected, those compounds that were chosen from the lists of best DCE scores had better-than-average DCE scores (approximately –50 kcal/mol for both pockets vs a range of –44 to –48 kcal/mol for compounds from other lists). Similarly, compounds picked based on the FPS_{Sum} score had more favorable versus average scores (Table 1, averages 4.7 and 6.8 kcal/mol) compared to those compounds picked from other lists. However, it is worth noting that although the DCE and FPS scores vary among lists, all chosen compounds fall within a range of scores that would be considered favorable; that is, there were no molecules in the final ensemble that scored poorly by any of the metrics. An advantage of choosing compounds based on different scoring metrics is that the final ensemble will typically contain greater diversity in terms of their physical properties. For example, molecules with favorable FPS scores tend to have significant steric overlap with the reference, and molecules with favorable DCE scores typically are larger in size.^{20,21} Interestingly, for the present case, the molecules chosen from different groups in this study were fairly close in size in terms of molecular weight (Table 1, average between 463 and 502 g/mol) and number of rotatable bonds (Table 1, average between 9.9 and 12.2 rotatable bonds), although those chosen based on the FPS_{Sum} score tended to overlap sterically with the reference peptide better than those chosen based on DCE score (Fig. 3, panels D–G). The QLIQ pocket molecules with the most favorable DCE scores extend outside the reference volume (Fig. 3D), whereas molecules docked to the same pocket with the most favorable FPS_{Sum} scores fit more

Table 1
Summary of properties of small molecules selected from virtual screens.

| Pocket | List | N ^a | MW ^b | RB ^c | DCE ^d | FPS _{Sum} ^e |
|----------------------|-------------------------------|----------------|-----------------|-----------------|------------------|---------------------------------|
| QLIQ | DCE | 6 | 483.6 ± 22.7 | 12.2 ± 1.2 | -50.3 ± 1.4 | 12.0 ± 2.6 |
| | FPS _{V_{DW}} | 14 | 499.0 ± 25.5 | 10.3 ± 1.7 | -44.9 ± 0.9 | 7.0 ± 1.0 |
| | FPS _{ES} | 15 | 471.7 ± 29.6 | 10.5 ± 1.8 | -45.1 ± 0.9 | 10.7 ± 1.5 |
| | FPS _{Sum} | 14 | 500.0 ± 25.8 | 10.7 ± 1.4 | -45.0 ± 1.0 | 6.8 ± 0.9 |
| | TotalScore | 17 | 502.5 ± 24.9 | 12.2 ± 1.4 | -47.4 ± 1.3 | 7.6 ± 1.3 |
| IQLT | DCE | 13 | 482.4 ± 22.3 | 11.6 ± 1.8 | -50.4 ± 1.3 | 14.6 ± 1.0 |
| | FPS _{V_{DW}} | 18 | 473.2 ± 25.4 | 9.9 ± 1.0 | -44.8 ± 0.5 | 5.5 ± 0.6 |
| | FPS _{ES} | 12 | 479.4 ± 32.0 | 10.0 ± 2.1 | -45.2 ± 0.9 | 6.4 ± 1.1 |
| | FPS _{Sum} | 15 | 463.1 ± 25.8 | 9.9 ± 1.5 | -44.7 ± 0.5 | 4.7 ± 0.1 |
| | TotalScore | 18 | 480.1 ± 24.1 | 10.3 ± 1.2 | -48.2 ± 0.9 | 6.4 ± 0.8 |
| Summary ^f | | 120 | 483.3 ± 29.1 | 10.7 ± 1.7 | -46.4 ± 2.2 | 7.8 ± 3.1 |

^a N = number of molecules in the group.

^b MW = average molecular weight.

^c RB = average number of rotatable bonds.

^d DCE = DOCK Cartesian energy score in units of kcal/mol (a larger negative number is more favorable).

^e FPS_{Sum} = footprint similarity score in units of kcal/mol (a number closer to 0 is more favorable).

^f Note that the sum of the number of molecules in all lists (142) is more than the total number of unique molecules purchased (120) due to some molecules appearing in multiple lists.

tightly within the reference volume (Fig. 3F). The same trend, although more subtle, was observed for molecules from the IQLT pocket as well (Fig. 3E, G).

To experimentally evaluate the efficacy of each selected compound in Table 1 for inhibiting fusion, a previously-established model of the HIV-host system using HL2/3³⁸ cells stably expressing the viral proteins and TZM-bl^{39–43} as the receptor cell line was employed. Briefly, the cell–cell fusion assay is based upon a luciferase reporter gene. The effector cell line (HL2/3) has high-level production of Gag, Env, Tat, Rev, and Nef proteins but does not produce any detectable reverse transcriptase or viral particles. When co-cultured with a target cell line (TZM-bl) that has the HIV receptors, CD4 and co-receptor, stably transfected, cell–cell fusion will occur. Tat from the effector cell line will transactivate the long terminal repeat that has been introduced in front of a luciferase gene in the target cell line. Fusion efficiency can be quantitated by measuring the relative light units emitted upon addition of the luciferase substrate (luminescence). Cytotoxicity will decrease the measured fusion efficiency because a dead cell will not express the luciferase enzyme resulting in decreased luminescence.

The TZM-bl cells were plated at an approximate density of 2.5×10^4 cells/well in a 96-well plate and, after 24 h, HL2/3 cells were added to the same plate at a density of 5.0×10^4 cells/well. Note that TZM-bl cells divide once every 24 h, therefore the receptor and infectious cells were present in a roughly 1:1 ratio at the time of the experimental measurement. Compounds from the virtual screen were diluted in dimethyl sulfoxide and added to the plates at the same time as the HL2/3 cells to a final concentration of 100 μ M. After 6 h of incubation, cell–cell fusion and cytotoxicity were measured on a Spectra Max M5 plate reader (Molecular Devices) using a combined reporter/cell viability assay following the protocol recommended by the manufacturer (ONE-Glo + Tox, Promega). All experiments were performed in sextuplicate. The results for the control conditions, including (1) TZM-bl cells alone, (2) TZM-bl cells + HL2/3 cells, and (3) TZM-bl cells + HL2/3 cells + 1 μ M known anti-fusion peptide C34,⁹ along with all experimental conditions (4) TZM-bl cells + HL2/3 cells + 100 μ M compound from virtual screen (top 25 compounds with the strongest anti-fusion signal are presented), are shown in Figure 4. Statistical significance was assessed by ANOVA with a post hoc pairwise Tukey's *t* test using GraphPad Prism 6 (GraphPad Software, Inc.). All *p*-values in this study were calculated relative to the fusion level of TZM-bl + HL2/3 without any inhibitor.

The luciferase reporter assay in Figure 4 (top panel) was designed such that the amount of signal directly correlates with the extent of cell–cell fusion. In the simplest control experiment, TZM-bl cells alone produced low, background level of luminescence signal. TZM-bl cells incubated with the effector HL2/3 cells produced a dramatic increase in luminescence, indicating the occurrence of cell–cell fusion. The peptide C34, a known potent inhibitor of HIV fusion,⁹ blocked cell–cell fusion as indicated by a return to background levels of luminescence ($p \leq 0.0001$). Encouragingly, at 100 μ M several of the small molecule compounds (450–500 molecular weight) appeared to inhibit cell–cell fusion (Fig. 4, top panel) at levels comparable to that of the much larger 34-amino acid peptide C34 (4248 molecular weight). All of the molecules in Figure 4 except B8, B5, and A8 presented statistically significant inhibition relative to the control (TZM-bl cells + HL2/3 cells without inhibitor) with $p \leq 0.0001$. In some instances, however, cytotoxicity at these test concentrations was higher than desired as shown in the bottom panel of Figure 4, where decreased signal correlates with increased cytotoxicity. Nevertheless, because small molecules may avoid many of the pharmacokinetic and pharmacodynamic drawbacks of peptide inhibitors, additional experiments to more fully characterize the hits were pursued.

To examine anti-fusion activity versus cytotoxicity in greater detail, dose–response experiments were subsequently performed over a range of concentrations for the top seven compounds (F8, C6, D10, A2, D9, I12, D7) shown in Figure 4 (blue bars) following a three-step protocol according to standard practices and the recommendation of the manufacturer.^{44,45} (1) The 'fraction of maximal effect' was computed by first dividing all cytotoxicity data points by the maximal obtained signal, thereby normalizing data to a 0 → 1 'fraction of maximal proliferation' scale. (2) All fusion data points were normalized to cell number by dividing by the fraction of maximal proliferation in the same well, thereby providing the fusion effect per unit cell. This accounts for the decrease in fusion efficiency due to cell death. (3) Background luminescence was removed from the fusion activity data points and normalized to achieve the fraction of maximal effect. Following this analysis, two compounds (D9 and A2) emerged as promising hits with good dose-dependent anti-fusion activity and reasonably low cytotoxicity as shown in Figure 5. Here, clear inflection points are observed for both compounds for the fusion activity (Fig. 5A–B, black lines) at approximately 70–80% cell health (Fig. 5, red lines). After curve fitting, the IC₅₀ was determined to be 58.6 μ M for D9 and 56.7 μ M

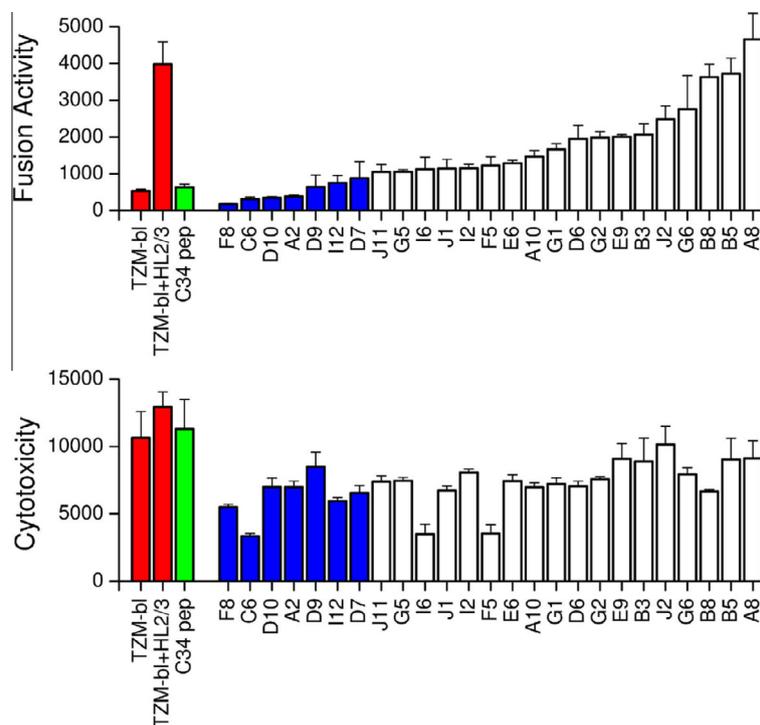


Figure 4. (Top) experimental fusion activity and (bottom) cytotoxicity for 25 compounds from the virtual screen as measured in a combined luciferase reporter and cell viability assay. The TQM-bl cell line was used to model the receptor cells, and the HL2/3 cell line was used to model HIV-1 and HIV-infected cells. The control inhibitor is peptide C34. Test compounds are labeled by alphanumeric code along the x-axis. Compounds colored in blue were investigated further with dose–response curves (see Fig. 5).

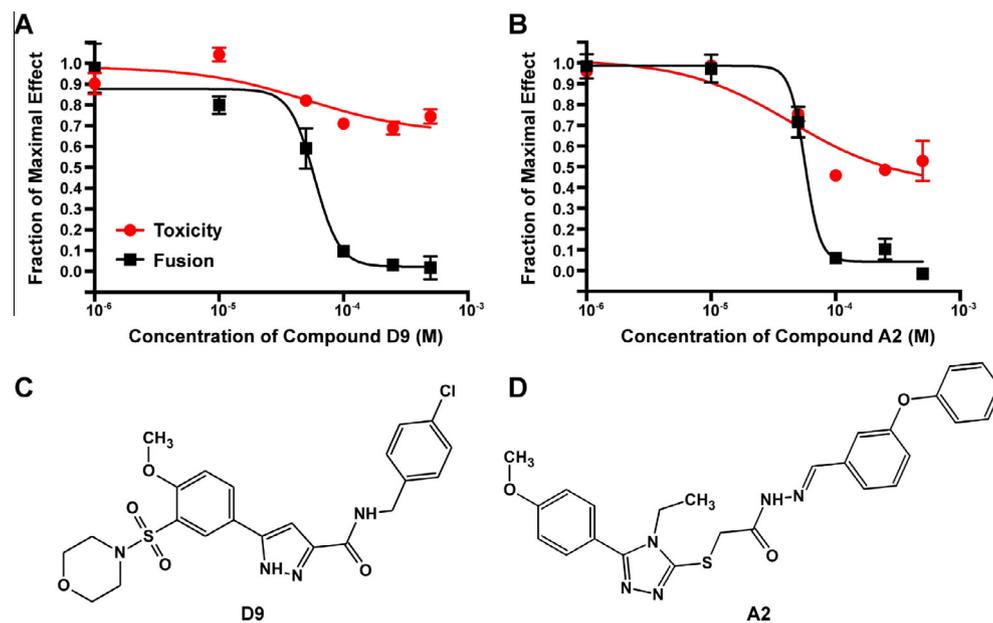


Figure 5. Dose–response anti-fusion curve (black) and cytotoxicity curve (red) for compounds (A) D9 and (B) A2. Chemical structures of compounds (C) D9 and (D) A2.

for A2. The corresponding CC_{50} values are estimated to be $>500 \mu\text{M}$ (D9) and $\sim 500 \mu\text{M}$ (A2). It is important to emphasize that these two compounds were chosen based on the FPS_{sum} (D9) or TotalScore (A2) scoring functions, highlighting the benefit of using multiple scoring functions when choosing compounds for experimental testing. While further work will be required to refine these compounds in order to achieve sub- μM anti-fusion activities, for an initial screen representing a mechanistically novel mode of blocking viral fusion, these are reasonable inhibition values recorded at an acceptable level of cytotoxicity.

Although the results for compounds D9 and A2 are encouraging, as of yet there is no direct evidence to demonstrate these compounds bind to the putative pockets at the interface of two gp41 NHR helices. The process of obtaining such evidence is hindered by a lack of a biochemical assay due to the high propensity for aggregation and precipitation of the NHR peptides alone.¹⁹ As additional indirect evidence for binding, however, we investigated whether these two compounds act by binding to a different conserved pocket formed at the outer interface of the NHR trimer (termed the ‘conserved hydrophobic pocket’), as reported for other

small molecules.^{6,11,12} Thus, we performed a FRET assay (described previously⁴⁶) which was designed specifically to measure binding to the conserved hydrophobic pocket. Importantly, the experiment produced negligible (A2) or relatively weak (D9) signal at a high compound concentration of 800 μM compared to a positive control bithionol at 40 μM (Fig. S1, Supplementary material), which suggests the hits are blocking cell–cell fusion by some other mechanism(s).

As an additional, physics-based approach to assess compatibility with the NHR inner pocket, we performed sextuplicate, explicit-solvent MD simulations of both compounds (A2 and D9) in their predicted docked geometries using the program Gromacs 5.0.2.²⁴ Prior studies from our laboratory³² showed that ligands in their native (i.e., crystallographic) binding geometry yield, on average, smaller root-mean-square deviations (RMSDs) during MD than do other low-energy decoys, and that such analysis can be a useful way to help evaluate pocket compatibility. The same simulation conditions described above were used here, with additional GAFF parameters⁴⁷ and AM1-BCC charges^{48,49} on the small molecule ligands and with position restraints ($1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$) on the protein backbone atoms. To maintain consistency with the original MD protocols used to define the references, as well as DOCK calculations, these follow-up MD simulations also employed a fixed partial charge model. Encouragingly, compound D9 showed very high stability during the simulations although A2 was not stable. Focusing on D9, negligible fluctuations in ligand RMSD were observed, relative to the initial pose, in three out of six 20-ns MD simulations in which the compound was unrestrained (Fig. 6A, black, red, and magenta lines), and small and short-lived fluctuations were observed in two out of six simulations (Fig. 6A, green and orange lines). And, although more significant RMSD fluctuations were observed in a final MD simulation, these fluctuations were transient and primarily localized to the terminal sulfonyl and morpholine group(s), and the ligand consistently returned to the initial docked position (Fig. 6A, blue line). In addition, MD simulations of 10 decoy ligands representing the top two scoring compounds from each of the five scoring metrics were performed in quintuplicate. Five out of ten decoys exhibited greater than 4 Å RMSD when averaged over all replicates, and eight out of ten decoys exhibited greater than 2 Å RMSD when averaged over all replicates (Table S1, Supplementary material). Notably, the average RMSD for active compound D9 (1.48 Å) was lower than all decoy compounds tested.

To better describe the specific interactions made by D9 in the pocket, we computed a time-averaged footprint and compared that to the footprint made by the reference peptide residues IQLT (Fig. 6B). Previously, we had decomposed one NHR helix as a ‘peptide ligand’ to determine which residues were most important for NHR trimer formation (see Fig. 2). Here, we took an opposite approach and decomposed the NHR ‘receptor’ residues to quantify the similarity of binding between the IQLT peptide reference and the D9 small molecule. As shown in Figure 6B (top), the time-averaged VDW footprint made by D9 (red line) overlaps well with that made by the IQLT peptide (black line), and the 14 significant steric-packing interactions made by both compounds with the receptor are nearly identical in terms of their position and magnitude. Figure 6C shows the position of these key residues in the context of the initial docked geometry (thick black line) as well as the tightly focused ligand sampling space (thin colored lines) during the MD simulation. The relative absence of structure in the accompanying ES footprint (Fig. 6B, bottom) for the IQLT reference (black line) emphasizes the importance of VDW-type interactions in the proposed pocket. Interestingly, the D9 small molecule adds two significant ES interactions at residue Gln 51 from either chain (Fig. 6B, bottom, red line) which form a pincer-like grip that may not only help to anchor the ligand into the proposed pocket

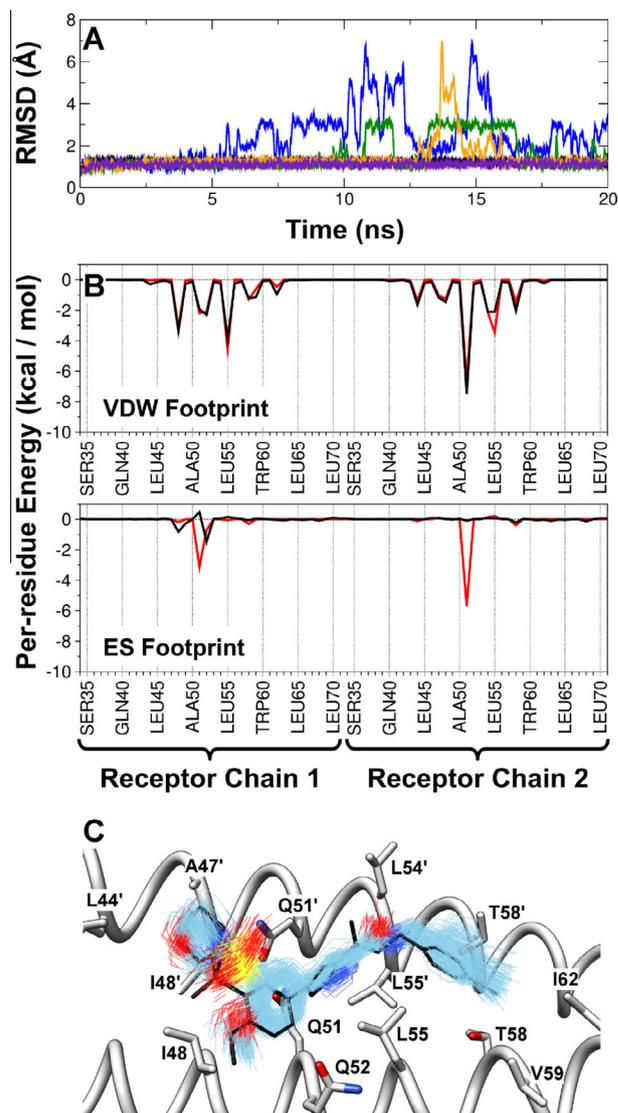


Figure 6. (A) RMSD for compound D9 in binding pocket from six replicate, 20-ns MD simulations. Lines shown are running average with window of 50 ps. (B) Footprint of reference peptide IQLT (black lines) compared to time-averaged footprints of compound D9 from MD simulations. Footprints are divided into VDW and ES types. (C) Compound D9 in the predicted binding geometry in the IQLT pocket (thick black line) overlaid with 200 evenly-spaced snapshots from one MD simulation (1 snapshot every 100 ps, thin colored lines). IQLT pocket shown with key residues labeled (gray ribbon/sticks). Amino acids from receptor chain 2 are marked with a single quote character (').

(Fig. 6C), but also add selectivity for the target. Taken together, these additional results (RMSD plots, time-averaged footprints, and structural overlays), help to provide an ensemble-based view that the docked pose of D9 is energetically and sterically compatible with the putative NHR inner pocket, thereby providing supporting evidence for the proposed mechanism of action.

In summary, we hypothesize that small molecules which mimic the activity and mechanism of *N*-peptide inhibitors that target HIVgp41 (e.g., N36) may avoid the pharmacological challenges associated with peptide-based inhibitors, yet still prevent virus-cell fusion by inhibiting NHR trimer formation. To that end, we computationally identified important interactions and areas of local curvature at the inner NHR interface of gp41, and performed high-throughput virtual screens with the program DOCK⁵⁰ to two putative binding pockets (termed QLIQ and IQLT) each consisting of approximately 1.4 million commercially-available small

molecules. Docked molecules were assessed by several different physics-based energetic and descriptive metrics, and 120 compounds were purchased for experimental testing. In dose–response curves, two compounds (D9 and A2) identified by footprint-based scoring emerged as promising candidates for probing anti-HIV fusion via this proposed mechanism. While additional work is required to conclusively demonstrate that these compounds disrupt NHR trimer formation, they potentially pave the way for a new class of small-molecule anti-HIV therapeutics, which to our knowledge has not yet been reported.

Acknowledgements

This work was funded in part by the Stony Brook University Office of the Vice President for Research and NIH grants F32GM105400 (to W.J.A.), R01GM087998 (to M.G.), R21AI102796 (to A.J.), and R01GM083669 (to R.C.R.). The following reagents were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HL2/3 from Dr. Barbara K. Felber and Dr. George N. Pavlakis, and TZM-bl from Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc. This research utilized resources at the New York Center for Computational Sciences at Stony Brook University/Brookhaven National Laboratory which is supported by the U.S. Department of Energy under Contract No. DE-AC02-98CH10886 and by the State of New York.

Supplementary data

Supplementary data (experimental data for the hydrophobic pocket binding assay, and RMSD values for active and decoy compounds during MD simulations) associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bmcl.2015.04.067>.

References and notes

- Global Update on the Health Sector Response to HIV, 2014; World Health Organization: Geneva, Switzerland, 2014.
- Shafer, R. W.; Schapiro, J. M. *AIDS Rev.* **2008**, *10*, 67.
- Johnson, V. A.; Calvez, V.; Gunthard, H. F.; Paredes, R.; Pillay, D.; Shafer, R.; Wensing, A. M.; Richman, D. D. *Top. Antivir. Med.* **2011**, *19*, 156.
- Eckert, D. M.; Kim, P. S. *Annu. Rev. Biochem.* **2001**, *70*, 777.
- Doms, R. W.; Moore, J. P. *J. Cell. Biol.* **2000**, *151*, F9.
- Cai, L.; Jiang, S. *ChemMedChem* **2010**, *5*, 1813.
- Wild, C.; Oas, T.; McDanal, C.; Bolognesi, D.; Matthews, T. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10537.
- Wild, C.; Greenwell, T.; Matthews, T. *AIDS Res. Hum. Retroviruses* **1993**, *9*, 1051.
- Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. *Cell* **1997**, *89*, 263.
- Wang, R.-R.; Yang, L.-M.; Wang, Y.-H.; Pang, W.; Tam, S.-C.; Tien, P.; Zheng, Y.-T. *Biochem. Biophys. Res. Commun.* **2009**, *382*, 540.
- Debnath, A. K.; Radigan, L.; Jiang, S. *J. Med. Chem.* **1999**, *42*, 3203.
- Holden, P. M.; Kaur, H.; Goyal, R.; Gochin, M.; Rizzo, R. C. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3011.
- Bewley, C. A.; Louis, J. M.; Ghirlando, R.; Clore, G. M. *J. Biol. Chem.* **2002**, *277*, 14238.
- Root, M. J.; Steger, H. K. *Curr. Pharm. Des.* **2004**, *10*, 1805.
- Ashkenazi, A.; Shai, Y. *Eur. Biophys. J.* **2011**, *40*, 349.
- He, Y.; Vassell, R.; Zaitseva, M.; Nguyen, N.; Yang, Z.; Weng, Y.; Weiss, C. D. *J. Virol.* **2003**, *77*, 1666.
- Gustchina, E.; Louis, J. M.; Bewley, C. A.; Clore, G. M. *J. Mol. Biol.* **2006**, *364*, 283.
- Allen, W. J.; Rizzo, R. C. *Biology* **2012**, *1*, 311.
- Lu, M.; Blacklow, S. C.; Kim, P. S. *Nat. Struct. Biol.* **1995**, *2*, 1075.
- Balius, T. E.; Mukherjee, S.; Rizzo, R. C. *J. Comput. Chem.* **2011**, *32*, 2273.
- Balius, T. E.; Allen, W. J.; Mukherjee, S.; Rizzo, R. C. *J. Comput. Chem.* **2013**, *34*, 1226.
- McGillick, B. E.; Balius, T. E.; Mukherjee, S.; Rizzo, R. C. *Biochemistry* **2010**, *49*, 3575.
- Caffrey, M. *Biochim. Biophys. Acta, Mol. Basis Dis.* **2001**, *1536*, 116.
- Abraham, M.; Hess, B.; van der Spoel, D.; Erik, L. *Gromacs User Manual version 5.0.2*; Royal Institute of Technology and Uppsala University: Sweden, 2014.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*.
- DMS, UCSF Computer Graphics Laboratory: San Francisco, CA, 2003.
- Desjarlais, R. L.; Sheridan, R. P.; Seibel, G. L.; Dixon, J. S.; Kuntz, I. D.; Venkataraghavan, R. *J. Med. Chem.* **1988**, *31*, 722.
- Foley, B.; Leitner, T.; Apetrei, C.; Hahn, B.; Mizrahi, I.; Mullins, J.; Rambaut, A.; Wolinsky, S.; Korber, B. *HIV Sequence Compendium 2014*; Los Alamos National Laboratory: Los Alamos, NM, 2014. 367.
- Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. *RNA* **2009**, *15*, 1219.
- Brozell, S. R.; Mukherjee, S.; Balius, T. E.; Roe, D. R.; Case, D. A.; Rizzo, R. C. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 749.
- Holden, P. M.; Allen, W. J.; Gochin, M.; Rizzo, R. C. *Bioorg. Med. Chem.* **2014**, *22*, 651.
- Berger, W. T.; Ralph, B. P.; Kaczocha, M.; Sun, J.; Balius, T. E.; Rizzo, R. C.; Haj-Dahmane, S.; Ojima, I.; Deutsch, D. G. *PLoS One* **2012**, *7*, e50968.
- Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. *J. Chem. Inf. Model.* **2012**, *52*, 1757.
- Brown, R. D.; Martin, Y. C. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 572.
- MOE, Chemical Computing Group: Montreal, Quebec, Canada, 2012.
- Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3.
- Ciminale, V.; Felber, B. K.; Campbell, M.; Pavlakis, G. N. *AIDS Res. Hum. Retroviruses* **1990**, *6*, 1281.
- Platt, E. J.; Wehrly, K.; Kuhmann, S. E.; Chesebro, B.; Kabat, D. *J. Virol.* **1998**, *72*, 2855.
- Derdeyn, C. A.; Decker, J. M.; Sfakianos, J. N.; Wu, X.; O'Brien, W. A.; Ratner, L.; Kappes, J. C.; Shaw, G. M.; Hunter, E. *J. Virol.* **2000**, *74*, 8358.
- Wei, X.; Decker, J. M.; Liu, H.; Zhang, Z.; Arani, R. B.; Kilby, J. M.; Saag, M. S.; Wu, X.; Shaw, G. M.; Kappes, J. C. *Antimicrob. Agents Chemother.* **2002**, *46*, 1896.
- Takeuchi, Y.; McClure, M. O.; Pizzato, M. *J. Virol.* **2008**, *82*, 12585.
- Platt, E. J.; Biliska, M.; Kozak, S. L.; Kabat, D.; Montefiori, D. C. *J. Virol.* **2009**, *83*, 8289.
- Schagat, T.; Paguio, A.; Kopish, K. *Cell Notes* **2007**, *17*, 9.
- Gopal, A., Recommendations for Normalizing Reporter Assays, <https://promega.wordpress.com/2012/03/19/recommendations-for-normalizing-reporter-assays/> (Accessed March 26, 2015).
- Gochin, M. *Assay Drug Dev. Technol.* **2012**, *10*, 407.
- Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132.
- Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623.
- Allen, W. J.; Balius, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. *J. Comput. Chem.* **2015**, *35*, 1132.