ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

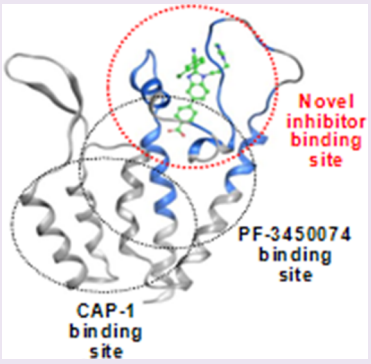
Novel Inhibitor Binding Site Discovery on HIV-1 Capsid N-Terminal Domain by NMR and X-ray Crystallography

Nathalie Goudreau,*[,†,§] Christopher T. Lemke,[†,§] Anne-Marie Faucher,[†,∥] Chantal Grand-Maître,[†] Sylvie Goulet,[†] Jean-Eric Lacoste,[†] Jean Rancourt,[†] Eric Malenfant,[†] Jean-François Mercier,[‡] Steve Titolo,[‡,⊥] and Stephen W. Mason*[,‡,#]

Departments of [†]Chemistry and [‡]Biological Sciences, Boehringer Ingelheim (Canada) Ltd, Research & Development, 2100 Cunard Street, Laval, Québec, Canada H7S 2G5

Supporting Information

ABSTRACT: The HIV-1 capsid (CA) protein, a domain of Gag, which participates in formation of both the mature and immature capsid, represents a potential target for anti-viral drug development. Characterization of hits obtained via high-throughput screening of an *in vitro* capsid assembly assay led to multiple compounds having this potential. We previously presented the characterization of two inhibitor series that bind the N-terminal domain of the capsid (CA_{NTD}), at a site located at the bottom of its helical bundle, often referred to as the CAP-1 binding site. In this work we characterize a novel series of benzimidazole hits. Initial optimization of this series led to compounds with improved *in vitro* assembly and anti-viral activity. Using NMR spectroscopy we found that this series binds to a unique site on CA_{NTD}, located at the apex of the helical bundle, well removed from previously characterized binding sites for CA inhibitors. 2D $^{1}H-^{15}N$ HSQC and ^{19}F NMR showed that binding of the benzimidazoles to this distinct site does not affect the binding of either cyclophilin A (CypA) to the CypA-binding loop or a benzodiazepine-based CA assembly inhibitor to the CAP-1 site. Unfortunately, while compounds of this series achieved promising *in vitro* assembly and anti-viral effects, they also were found to be quite sensitive to a number of naturally occurring CA_{NTD} polymorphisms observed among clinical isolates. Despite the negative impact of this finding for drug development, the discovery of multiple inhibitor binding sites on CA_{NTD} shows that capsid assembly is much more complex than previously realized.



The HIV-1 capsid protein (CA), which plays essential functions in both early and late stages of the viral replication cycle, represents a viable target for drug discovery.[1−4] CA is initially synthesized as a central domain of the 55-kDa Gag polyprotein, where it contributes key protein−protein interactions required for the assembly of immature viral particles.[2,5] During viral maturation, proteolytic cleavage of Gag releases CA, which reassembles to form a distinct fullerene cone-shaped structure called the capsid core that encapsidates the viral RNA genome along with specific viral enzymes and proteins necessary for virus infectivity.[1,6]

The capsid protein comprises two independently folded domains, the N-terminal domain (CA_{NTD}, residues 1−146) and the C-terminal domain (CA_{CTD}, residues 151−231) that are connected by a short flexible linker. The capsid protein and the assembled capsid core have been under intense structural characterization for more than a decade. Cryoelectron tomography, NMR, and X-ray structures of the isolated domains, as well as dimeric, pentameric, and hexameric assemblies, have been reported, thereby elucidating a detailed picture of the protein/protein interactions essential to the assembly and stability of the viral capsid lattice.[7−16] Proper capsid formation is required for infectivity, and CA mutations that affect the stability of the capsid core were found to impair virus infectivity.[5,17,18]

Inhibitors of HIV-1 CA assembly have been reported previously, providing evidence that CA may be a good anti-viral target.[19−27] Although several independent binding sites on CA have been defined and targeted by these different classes of inhibitors, the mechanism by which they all appears to operate is through interference with the formation of the CA/CA intermolecular interface. The first capsid assembly inhibitor reported was CAP-1, a small molecule that binds to a deep, induced-fit hydrophobic pocket at the base of the CA_{NTD} helical bundle (at the junction of α-helices 1, 2, 3, 4, and 7).[19,20] More recently, another small molecule inhibitor, PF-3450074, was also found to target CA_{NTD}.[21,22] Unlike CAP-1, this inhibitor did not induce a conformational change but instead bound in an adjacent pocket defined by helices 3, 4, 5, and 7.

Recently, using high-throughput screening (HTS), we also identified two structurally distinct classes of compounds, the benzodiazepines and the benzimidazoles, as capsid assembly inhibitors.[24−26] Optimization of these two series resulted in compounds with potent HIV-1 anti-viral activities (EC_{50} < 100 nM). These inhibitors were found to bind to the CA_{NTD} and induced the formation of a pocket at the bottom of its helical

Received: January 29, 2013
Accepted: March 15, 2013
Published: March 15, 2013

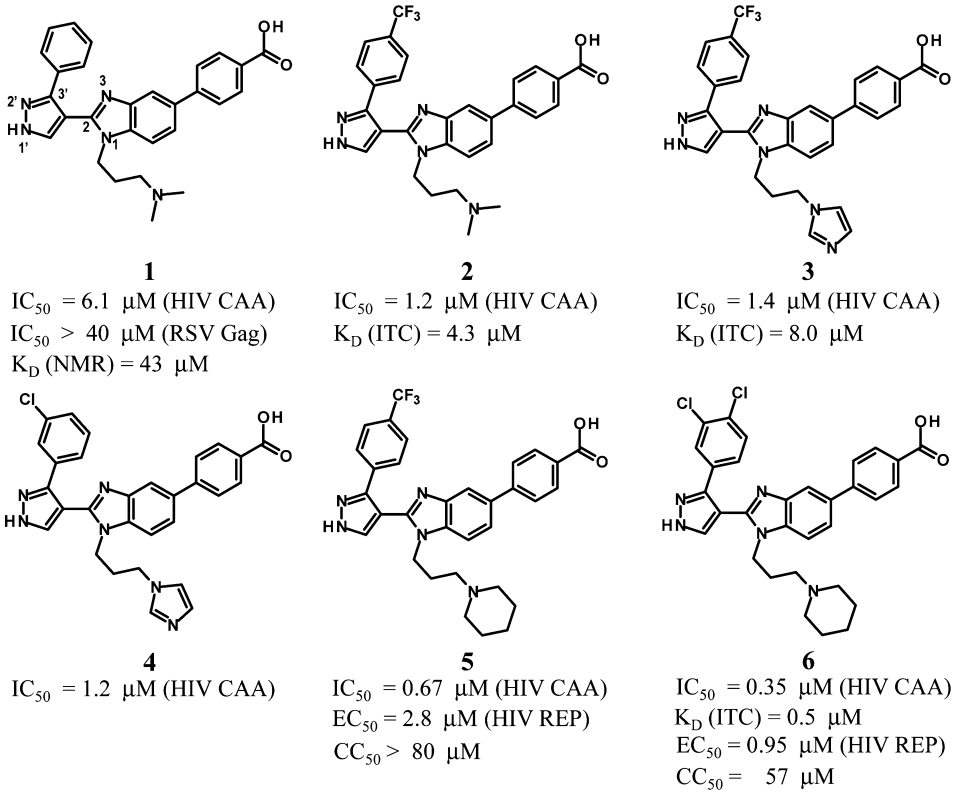


Figure 1. Stuctures and potencies in the HIV-1 capsid assembly assay (HIV CAA) of the initial hit (1) and related analogues. K_d (ITC) values were measured by isothermal titration calorimetry using the N-terminal domain of capsid (CA_{NTD}) as the protein. K_d (NMR) was measured as described in Figure 3. Anti-viral potencies (EC_{50}) were measured using an HIV-1 luciferase-based replication assay (HIV REP). Cytotoxicity concentration (CC_{50}) is defined as the concentration resulting in a 50% reduction in cell viability.

bundle that overlaps with the CAP-1 binding site.[24−26] Though both of these inhibitor series bind to an overlapping site and impair late stage virus replication, they had very different morphological effects on virus assembly.[25]

In this study, we report the identification of a novel capsid assembly inhibitor series. Using 2D $^{1}H-^{15}N$ HSQC NMR and X-ray crystallography we found that this series binds to a unique site located at the apex of the helical bundle on CA_{NTD} that is distinct from other previously described inhibitor binding sites. Despite the proximity of this new site to the cyclophilin A (CypA) binding loop, this novel inhibitor series does not prevent CypA binding. Both molecules can bind CA_{NTD} simultaneously. Using ^{19}F NMR, we were able to demonstrate that this novel inhibitor series can bind the CA_{NTD} in the presence of a previously identified CAP-1 binding site inhibitor (from the benzodiazepine series). Finally, while compounds of this series achieved promising *in vitro* assembly and anti-viral effects, they were also found to be quite sensitive to a number of naturally occurring CA_{NTD} polymorphisms observed among clinical isolates that precluded their development as broad spectrum HIV-1 antiretrovirals.

RESULTS AND DISCUSSION

Identification of a Novel Inhibitor Binding Site on CA_{NTD}. Previously, we described an *in vitro* capsid assembly assay (CAA) that was used in a HTS to identify potential inhibitors of capsid assembly.[25] Using this assay, we identified several structurally distinct families of compounds as hits. Triaging of these hits based on several criteria, including potency, lack of activity in a counter screen assembly assay using an analogous Gag CA-NC protein from Rous sarcoma virus (RSV),[28] cytotoxicity, and assessment of chemical tractability, left us with 121 hits representing more than 50 chemotypes. These surviving hits were then subjected to NMR chemical shift perturbation analysis using ^{15}N $CA_{W184A/M185A}$ (full length capsid dimerization defective mutant)[29,30] and/or ^{15}N CA_{NTD} to confirm the binding of compounds to CA and to identify potential binding sites. Of these 121 hits, only 6 compounds did not show binding to either of the above two proteins, while 111 compounds were found to bind CA_{NTD} at one of the two binding sites already described in the literature, the CAP-1 or the PF-3450074 binding sites.[19,21] This early profiling resulted in the prioritization of two distinct families of compounds for lead optimization, the benzodiazepines and the benzimidazoles, as described previously.[24−26]

Interestingly, the 4 remaining capsid binders belonged to a benzimidazole chemotype represented by inhibitor **1** (Figure 1, IC_{50} = 6.1 μM) that was different than our previously characterized benzimidazole series[24,25] and were found to behave distinctively from all the other hits in their binding to ^{15}N-labeled CA_{NTD}. Indeed, whereas all of the previous 111 CA_{NTD} binders were found to affect the same subsets of residues including E29, K30, A31, F32, S33, E35, V36, V59, G60, G61, H62, Q63, A65, M144, and Y145 for CAP-1 pocket binders and L56, N57, T58, G60, G61, H62, Q63, Q67, K70, E71, T72, T107, T108, I134, and L138 for PF-3450074 pocket binders (and sometimes a combination of both), inhibitor **1** was found to perturb a different subset of residues. Figure 2a shows an overlay of 2D $^{1}H-^{15}N$ TROSY spectra obtained for CA_{NTD} in the *apo* state and in the presence of inhibitor **1**.

dx.doi.org/10.1021/cb400075f | ACS Chem. Biol. 2013, 8, 1074−1082

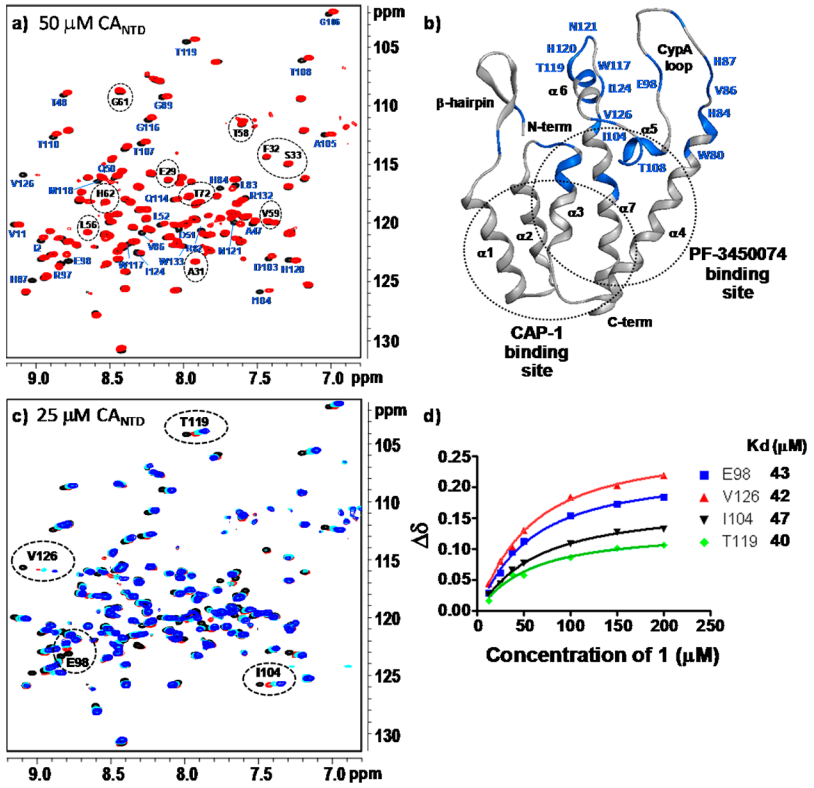


Figure 2. Identification of inhibitor binding site on CA_{NTD} and determination of its binding affinity. (a) Overlay of 2D $^{1}H-^{15}N$ TROSY spectra obtained for the HIV-1 N-terminal domain of capsid (CA_{NTD}, 1−146) upon addition of excess inhibitor **1** (black spectrum, free protein, 50 μM; red spectrum, protein + inhibitor, 1:1.3 ratio). Backbone NH crosspeaks that shift upon addition of inhibitor are labeled with corresponding residue numbers (highlighted in blue). Residues highlighted inside dashed circles are those that were unaffected by inhibitor **1** and belong to either the CAP-1 or PF-3450074 binding sites. (b) A ribbon diagram of the structure of CA_{NTD} highlighting in blue those residues that are affected upon saturation with inhibitor **1**, as indicated. These residues are located at the top of helix-4 and -7, on either side of the cyclophilin A binding loop, within helix-5 and helix-6 and the loop joining these two helices. Location of CAP-1 and PF-3450074 binding sites are also indicated with dash circles. (c) Overlay of a portion of 2D $^{1}H-^{15}N$ TROSY spectra obtained for CA_{NTD} (^{15}N CA_{NTD}, 1−146) upon addition of increasing amounts of inhibitor **1**. The color of spectra and corresponding CA_{NTD}:inhibitor **1** ratios were as follows: 1:0 (black); 1:0.5 (not shown); 1:1 (not shown); 1:1.5 (red); 1:2 (not shown); 1:4 (cyan), 1:6 (not shown); and 1:8 (blue) with 25 μM protein concentration. (d) Graphical representation of the ^{1}H and ^{15}N NMR chemical shift titration data for residues Glu98, Ile104, Thr119, and Val126. The data were fit to a 1:1 binding (K_d = 43 ± 2.9 μM).

Although most resonances were unaffected, there was a significant subset of resonances that shifted in the presence of this inhibitor, indicating site-specific binding. In Figure 2a, the NH crosspeaks that were perturbed by the addition of **1** have been labeled and correspond to I2, V11, A47, T48, Q50, D51, L52, W80, R82, L83, H84, V86, H87, G89, R97, E98, D103, I104, A105, G106, T107, T108, T110, Q114, G116, W117, M118, T119, H120, N121, I124, V126, R132, and W133. Most of these residues are located at the top of helix-4, the beginning of cyclophilin A (CypA) binding loop, within helix-5 and helix-6, the loops joining helix-5/helix-6 and helix-6/helix-7, and at the top of helix-7 (these residues are highlighted in blue in Figure 2b). The residues outside these regions that also experienced some chemical shift perturbations upon compound addition (I2, V11, A47, T48, Q50−L52, R132, W133) were probably affected indirectly as a result of rigidification of residues located at the compound binding site. Even though some of the perturbed residues were found to overlap with the PF-3450074 binding site (Figure 2b, in particular those belonging to helix-5), most were different and appeared to define a novel binding pocket, distinct from any of those previously described.[19−27] We will refer to this novel inhibitor pocket at the top of the helical bundle as the apical binding site of CA_{NTD}.

To further characterize this new inhibitor and estimate its binding affinity, a complete NMR titration was performed (Figure 2c). From an overlay of 2D $^{1}H-^{15}N$ TROSY spectra obtained for CA_{NTD} in the *apo* state and in the presence of increasing amounts of inhibitor **1**, the ^{1}H and ^{15}N chemical shift changes ($\Delta\delta$) for residues Glu98, Ile104, Thr119, and Val126 (among the largest) were chosen for plotting as a function of inhibitor concentration (Figure 2d). The resulting graph was fit to a 1:1 binding,[31] giving an average equilibrium dissociation constant (K_d) 43 ± 2.9 μM.

SAR of Benzimidazole Inhibitor Series. Preliminary SAR on this hit demonstrated that a CF_3 moiety could be introduced in the *para* position of the pyrazol C3′ phenyl ring, as exemplified by inhibitor **2** (Figure 1), producing at least a 5-fold improvement in potency. We also found that the N1 *N*-dimethyl moiety could be replaced by an imidazole ring without significant loss in potency (Figure 1, inhibitor **3**). Replacement of the *p*-CF_3 by a *m*-Cl substituent on the C3′ phenyl ring was also tolerated, as exemplified by inhibitor **4**. Another 2-fold improvement in potency was observed when the N1 *N*-dimethyl moiety of inhibitor **2** was replaced by a piperidine moiety, leading to compound **5**, the first inhibitor of this series demonstrating anti-viral activity in the HIV-1 replication assay. Finally, we discovered that combining both *m*-Cl and *p*-Cl

dx.doi.org/10.1021/cb400075f | *ACS Chem. Biol.* 2013, 8, 1074−1082

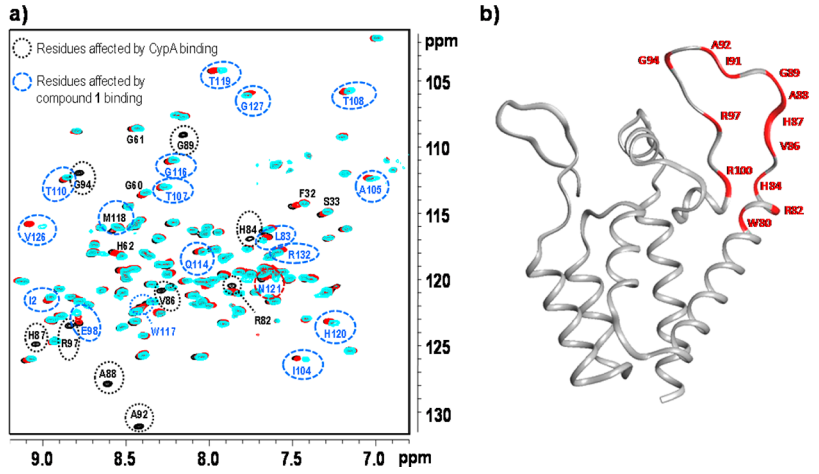


Figure 3. Inhibitor **1** does not prevent cyclophilin A from binding to CA_{NTD}. (a) Overlay of a portion of 2D $^{1}H-^{15}N$ TROSY spectra obtained for CA_{NTD} upon successive additions of excess CypA and inhibitor **1**. Black spectrum, free CA_{NTD}, 50 μM; red spectrum, CA_{NTD} + CypA, 1:1.4 ratio with 50 μM of CA_{NTD}; cyan spectrum, CA_{NTD} + CypA + inhibitor **1**, 1:1.4:1.5 ratio, 50 μM CA_{NTD}. Binding of CypA to CA_{NTD} resulted in broadening beyond detection of several residues present in the large loop between helices α4 and α5 (highlighted inside black dashed circles). Addition of excess inhibitor **1** to the CA_{NTD}/CypA complex resulted in the shift of several NH crosspeaks (highlighted inside blue dashed circles). The CA_{NTD} residues that shifted upon addition of inhibitor **1** were the same as those highlighted in Figure 2. (b) Structure of CA_{NTD} (ribbon diagram) highlighting residues (in red) affected by CypA addition.

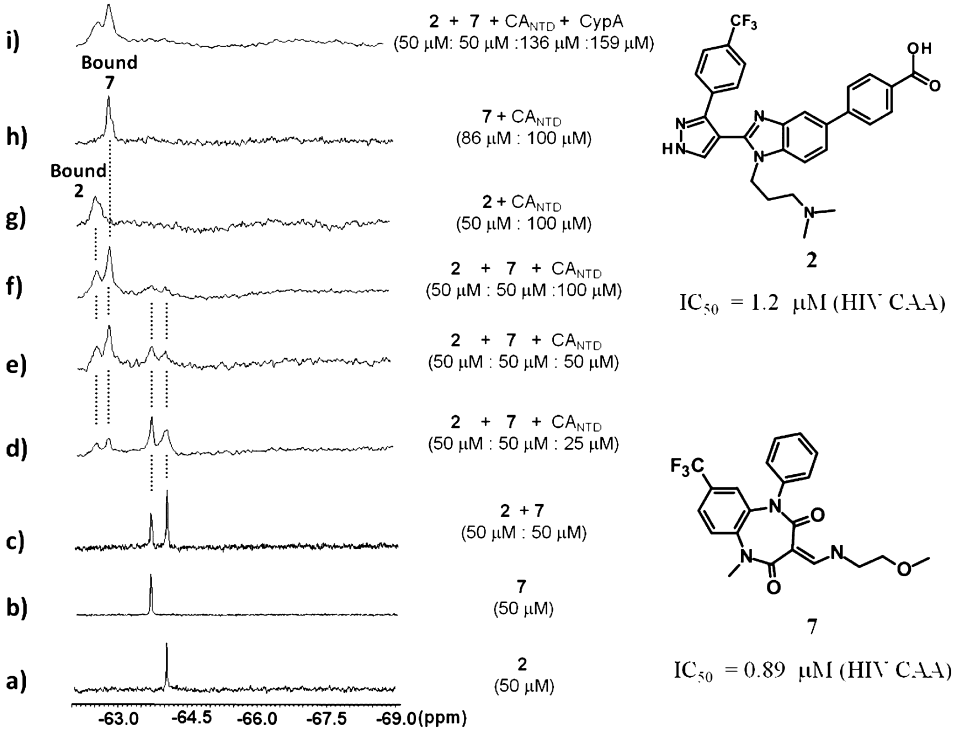


Figure 4. ^{19}F binding studies of inhibitors **2** and **7** to the CA_{NTD}. Portion of the 1D ^{19}F NMR spectra of inhibitors **2** and **7** in 20 mM Tris pH 7.5, 100 mM NaCl in 90% H_2O/10% D_2O mixture (400 MHz) recorded free and in the presence of CA_{NTD}. CF_3CO_2Na (50 μM) was added to all samples as an internal standard and referenced to −77 ppm (not shown). (a) Free inhibitor **2** (50 μM). (b) Free inhibitor **7** (50 μM); inhibitor **7** binds within the CAP-1 site at the bottom of CA_{NTD} helical bundle (Figure 2b). (c) An equimolar (50 μM:50 μM) mixture of inhibitors **2** and **7**. (d–f) A mixture of inhibitors **2** and **7** (50 μM:50 μM) in the presence of increasing amount of CA_{NTD} (25, 50, and 100 μM). (g) Inhibitor **2** in the presence of excess CA_{NTD} (50 μM:100 μM). (h) Inhibitor **7** in the presence of excess CA_{NTD} (86 μM:100 μM). (i) A mixture of inhibitors **2** and **7** (50 μM:50 μM) in the presence of excess CA_{NTD} and CypA (50 μM:50 μM:136 μM:159 μM).

substituents on the C3′ phenyl ring led to a further improvement in both *in vitro* and cell-based potencies (Figure 1, inhibitor **6**). This represented the first analogue of the series that had submicromolar anti-viral potency with a reasonable window with respect to cytotoxicity. Moreover, using a pseudotyped virus system described previously,[25] we found that the anti-viral activity of this inhibitor was manifest during the late stage of the viral replication cycle (data not shown), consistent with effects on capsid assembly.

CypA Is Not Competing wih Binding of the Benzimidazole Inhibitor Series to the CA_{NTD}. Several of the CA_{NTD} residues that were found to be perturbed by

dx.doi.org/10.1021/cb400075f | *ACS Chem. Biol.* 2013, 8, 1074–1082

inhibitor **1** were located within the CypA binding loop (Figure 2b). Since capsid is known to bind the human peptidyl prolyl isomerase CypA, an interaction known to be essential for viral infectivity,[32,33] we wondered if this series affected this binding. The mechanism through which CypA binding modulates viral infectivity is still not fully understood, but several possibilities have been discussed, including effects on capsid stability, viral uncoating, and the protection of viral cores from cellular restriction factors.[32,33] A crystal structure of CA_{NTD} in complex with human CypA has been reported previously and showed that CypA binds to the capsid extended loop formed by residues 85−93.[32] In order to verify if inhibitor **1** was disrupting CypA binding, additional NMR binding experiments were performed. Figure 3 shows an overlay of 2D $^{1}H-^{15}N$ TROSY spectra obtained for CA_{NTD} in the *apo* state and in the presence of CypA. As expected, and in agreement with the CA_{NTD}/CypA complex structure, binding of CypA to CA_{NTD} caused perturbations of residues almost exclusively within the loop joining helix-4/helix-5 (CypA binding loop). As highlighted in Figure 3a, most of these residues were broadened beyond detection, including residues R82, H84, V86, H87, A88, G89, I91, A92, G94, and R97 (residues W80 and R100 were also affected but not displayed in Figure 3a). The positions of these same residues are indicated in red in the ribbon diagram of CA_{NTD} shown in Figure 3b. As the next step, inhibitor **1** was then added to the same sample already containing the ^{15}N CA_{NTD}/CypA complex, and a new spectrum was recorded and overlaid as shown in Figure 3a. Addition of excess **1** resulted in the shift of the same CA_{NTD} residues as observed before in Figure 2a, while maintaining the broadening of those residues observed upon CypA binding. These data suggest that inhibitor **1** and CypA are able to bind CA_{NTD} simultaneously and do not compete with one another. Finally, this binding experiment was also repeated using a different order of addition (first binding of inhibitor **1** followed by CypA addition), resulting in the same conclusions (data not shown). Altogether these data demonstrate that the anti-viral mechanism of action of this novel inhibitor series is not through interference with CypA binding.

Simultaneous Binding of Two Inhibitor Series to CA_{NTD}. Having already identified potent capsid assembly inhibitors that bind CA_{NTD} in the CAP-1 binding site,[24−26] we wanted to verify if both inhibitory binding sites, the CAP-1 site at the bottom and this novel site at the top of the helical bundle, could be occupied simultaneously. Very recently, we have reported on the use of ^{19}F NMR to monitor binding of the previous benzodiazepine series (compound **7**, Figure 4); these studies were crucial to early hit confirmation and characterization.[34] Fluorine is a powerful nucleus for NMR as it has almost the same sensitivity as proton (^{1}H) and its ^{19}F isotope is 100% naturally abundant.[35−37] This high sensitivity, coupled with a virtual absence of background fluorine signal from most protein and buffer components, is a considerable advantage to monitor ligand binding. Importantly, the chemical shift of ^{19}F is very sensitive to its local environment and thus to changes that might occur upon binding.

Figure 4 displays subregions of the 1D ^{19}F spectra of free compound **2** (panel a), free compound **7** (a benzodiazepine inhibitor containing a CF_3 group and binding to the CAP-1 binding site, panel b), and a 50:50 mixture of both compounds **2** and **7** (panel c) recorded in the absence of CA_{NTD}. A distinct sharp peak was observed for each inhibitor, either alone or when mixed together. In panels d−f, increasing amounts of CA_{NTD} were added to the same mixture of inhibitors **2** and **7** (as in **c**) as indicated. This resulted in a significant broadening and a decrease in intensity of the free ^{19}F resonances for the compounds, while new distinct ^{19}F resonances appeared. The fact that we detect two distinct sets of broad resonances for the free and bound species is an indication that binding is occurring in the intermediate-to-slow exchange regime on the NMR time scale, consistent with the low micromolar/submicromolar activity measured for these inhibitors in the capsid assembly assay (IC_{50} = 1.2 and 0.89 μM for compounds **2** and **7**, respectively) and the somewhat large difference in chemical shifts observed between these two states ($\Delta\delta_{free-bound} \approx 1.4$ and 0.85 ppm for compounds **2** and **7**, respectively). Importantly, these data illustrate that both compounds can bind simultaneously. Moreover, when the same ^{19}F binding experiment was performed with each individual inhibitor (panels g and h), the binding behavior (exchange regime) and bound chemical shifts were found to be the same as those observed when the experiment was performed with the mixture of both inhibitors (panels d−f), suggesting that their respective CF_3 moieties are located in similar bound environments, in the presence or absence of the other inhibitor. Additionnal binding experiments performed with ^{15}N-labeled CA_{NTD} also confirmed that compounds **2** and **7** simultaneously bind at the two distinct binding sites based on different residue perturbations (Supporting Information). Interestingly, not only did the binding of the benzimidazole series to the apical site not affect binding of inhibitors to the CAP-1 site, it also positively impacted the success of co-crystallization studies of many other inhibitors with CA_{NTD} (PDB entries 4E91 and 4E92 and Lemke et al. *Acta Crystallogr., Sect. D*, accepted). Finally, addition of excess CypA to a mixture of inhibitors **2** and **7** in the presence of CA_{NTD} (panel i) did not displace either inhibitor, in agreement with the ^{15}N binding studies shown in Figure 2.

X-ray Co-crystal Structure of CA_{NTD} in Complex with a Benzimidazole Inhibitor. We also were successful in obtaining a co-crystal of CA_{NTD} in complex with inhibitor **4**, thus enabling a more detailed understanding of the binding of this novel benzimidazole inhibitor series (Figure 5). Figure 5a shows the structure of this complex with residues located within 5 Å from the inhibitor (colored in blue). In agreement with the ^{15}N NMR data, this co-crystal structure confirms that the inhibitor is binding to a novel site on CA_{NTD}, well-removed from the other two previously reported sites. Indeed, the CA_{NTD} residues that experienced the largest $^{1}H/^{15}N$ shifts were principally located within the 5 Å radius from the inhibitor (compare Figure 5a with Figure 2a).

The apical benzimidazole binding pocket is mainly composed of residues located at the base of the cyclophilin A binding loop, as well as those from the shorter loop that connects helices 6 and 7, and is surrounded by residues from helices 5 and 6 and the top of helix 7. The binding pocket is relatively featureless and shallow and extends approximately 20 Å along the surface. The inhibitor binds predominantly *via* hydrophobic interactions, especially with the anchoring of the *m*-Cl phenyl moiety in a small hydrophobic subpocket. The residues that are making significant hydrophobic contact with inhibitor **4** are Trp80, His84, Met96, Trp117, His120, Pro122, Ile124, Pro125, and Ile129 (Figure 5b). In addition to these hydrophobic interactions, the acid moiety of inhibitor **4** also forms a crucial bidentate salt-bridge with the side chain of Arg132, thus providing a second anchoring point to the inhibitor, located at the opposite end of the molecule. Full details about these

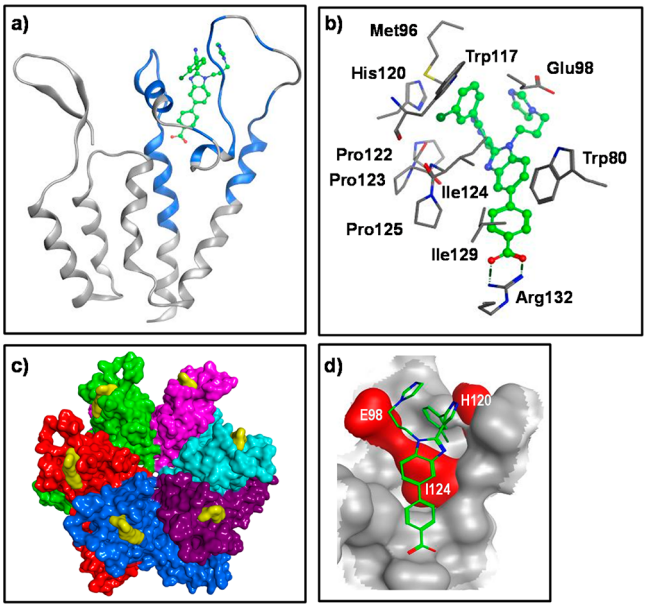


Figure 5. X-ray crystal structure of HIV-1 CA_{NTD} in complex with inhibitor **4** (1.7 Å resolution). (a) Ribbon representation of the structure of CA_{NTD} in complex with **4**. Residues highlighting in blue are those located within 5 Å of the inhibitor. (b) Detailed interactions of inhibitor **4** with CA_{NTD}. Side chains of interacting residues from CA_{NTD} are illustrated as gray sticks (with standard colors for noncarbon atoms). Hydrogen bonds are indicated with green lines. (c) Model of inhibitor **4** bound to hexameric CA (based on PDB entry 3MGE). The surface of each protomer is shown in a unique color with the surface of bound inhibitor **4** shown in yellow. The hexamer is depicted from an apical vantage point, tilted approximately 45° from the 6-fold axis. (d) Surface representation of inhibitor **4** binding site. The residues colored in red (Glu98, His120, and Ile124) represent residues with significant polymorphic variations based on 362 HIV-1 isolates in the Los Alamos National Laboratory HIV Database (http://www.hiv.lanl.gov).

interactions and comparison with the *apo* state will be published elsewhere (Lemke et al. *Acta Crystallogr., Sect. D*, accepted).

The mechanism by which most reported capsid assembly inhibitors operate appears to be through interference with the formation of the CA/CA intermolecular interfaces. Therefore, we also investigated possible effects on intermolecular capsid interactions by this novel inhibitor series. The reported cross-linked hexameric structure (PDB 3MGE) is believed to be a close representation of biologically relevant assembly for mature capsid hexamers.[38] In contrast to other reported capsid inhibitors, superposition of the current cocomplex structure onto the assembled hexamers (Figure 5c) did not place inhibitor **4** in proximity to any intermolecular CA_{NTD}/CA_{NTD} and CA_{NTD}/CA_{CTD} interfaces.[20,22,25,29] Thus, the exact molecular mechanism by which inhibitor **4** and its analogues elicit their observed anti-assembly and anti-viral effects was not clear based on the above structure. It does leave open the possibility that this apical binding site could constitute a platform for binding of host proteins (other than CypA) required for either capsid assembly or disassembly. However, alternate possibility comes from the recently reported 8 Å resolution cryo-electron microscopy structure of the Mason–Pfizer monkey virus (MPMV) immature Gag assembly, another retrovirus that is distantly related to HIV-1.[39] In that study, it was shown that the role of CA_{NTD} is to hold hexameric building blocks of the immature Gag lattice together and that the CA_{NTD}/CA_{NTD} interface involved in this interaction is primarily composed of helices 5–7. Interestingly, this interface matches our newly identified apical binding site. Albeit, MPMV is not HIV-1; nonetheless, the intriguing possibility exists that the current series of inhibitors exhibits its antiassembly and anti-HIV-1 effects by interfering with immature assembly.

Polymorphisms at Novel Inhibitor Binding Site. While clear improvements in the *in vitro* and anti-viral activities of this benzimidazole series were achieved during early hit optimization (Figure 1), unfortunately, it was found that this series also was sensitive to a number of naturally occurring CA_{NTD} polymorphisms observed among clinical isolates (Table 1).

Table 1. Effect on Compound Activity of Major CA Polymorphisms within the Apical Binding Site[a]

	CA-NC						
compd	WT	E98D	H120S	H120N	H120G	H120R	I124V
1	6.1	>25	>25	>25	>25	>25	>25
2	1.2	11.3	>25	3.2	>25	>25	17.6
3	1.4	>25	>25	4.8	>25	>25	18.4
5	0.67	7.2	>25	2.0	>25	>25	11.2

[a]The IC_{50} values (μM) for inhibitors **1**, **2**, **3**, and **5** against WT CA-NC in the fluorescence-based capsid assembly assay[25] are shown in column 2, and those for the indicated mutant proteins are in columns 3–8. All mutants demonstrated activity similar to that of WT in the assembly assay (data not shown).

On the basis of an alignment of capsid protein sequences from 362 clinical isolates (data not shown), it was realized that there were highly significant polymorphisms at position 120 of CA. In fact, histidine, the amino acid at CA position 120 in the HIV-1 NL4.3 molecular clone used in our experiments, is represented in only ~5% of clinical isolates, whereas serine (~50%), asparagine (~25%), and glycine (~12%) were the predominant amino acids found at that position. In addition, sequences at positions Glu98 and Ile124 were significantly polymorphic with ~7% aspartic acid and 28% valine observed at these two positions, respectively. Therefore, we felt that it would be extremely important to determine the susceptibility of these major capsid polymorphs to our novel capsid assembly inhibitor series; especially given the considerable impact that drug resistance can have on HIV-1 treatment options. For this purpose, we generated several CA-NC constructs and produced corresponding proteins containing mutations at position 98, 120, and 124 reflecting the above polymorphisms (Figure 5d and Table 1). In testing just a few inhibitors against these mutants in the CAA, it was realized that the IC_{50}'s were significantly shifted upward by at least 7-fold or greater against all polymorphic residues, with the exception of the H120N mutant for which smaller shifts, on the order of 2- to 3-fold, were observed with 3 of the 4 compounds. Therefore, given these results, the likelihood of delivering a broad spectrum HIV-1 antiretroviral drug with therapeutic value against this shallow, polymorphic pocket was considered rather low. Indeed as shown in Figure 5d, the polymorphic residues make up a good proportion of the binding site and are involved in direct contacts with the inhibitor. As a consequence, further efforts to optimize this novel inhibitor series were discontinued.

Conclusions. In this study, we demonstrated how protein-based 2D $^{1}H-^{15}N$ HSQC NMR was successfully used in the early validation phase of several hits identified from high-

dx.doi.org/10.1021/cb400075f | *ACS Chem. Biol.* 2013, 8, 1074–1082

throughput screening of a capsid assembly assay. In particular, this work allowed the discovery of a novel capsid assembly inhibitor series that binds to a unique site on CA_{NTD}, located at the apex of its helical bundle. Due to the proximity of this binding site to the cyclophilin A (CypA) binding loop, we also demonstrated that this novel inhibitor series does not prevent CypA binding and that both molecules can bind CA_{NTD} simultaneously. Using ^{19}F NMR, we also were able to demonstrate that this series can bind the CA_{NTD} in the presence of a previously identified CAP-1 binding site inhibitor. Co-crystallization of CA_{NTD} with this novel inhibitor series yielded well diffracting crystals and allowed for the determination of a complex structure. Finally, while compounds of this series achieved promising *in vitro* assembly and anti-viral effects, they were also found to be quite sensitive to a number of naturally occurring CA_{NTD} polymorphisms observed among clinical isolates. Despite this discouraging result, the discovery of another distinct inhibitor binding site on CA protein suggests that there are multiple possible mechanisms for inhibition of CA assembly and the process of capsid assembly is much more complex than previously realized.

METHODS

Inhibitor Synthesis and *in Vitro* Capsid Assembly Assay. The synthesis of the inhibitors presented in this manuscript is described in the Supporting Information. In addition, the capsid assembly assay (CAA) used to screen our compound collection and to measure IC_{50} values has been described in detail previously.[25] IC_{50} values reported are an average of at least two determinations.

Protein Expression and Purification. Expression and purification of HIV-$1_{NL4.3}$ CA-NC (WISP-98-68, Gag residues 133−432) and CA_{NTD} (WISP-96-19, CA residues 1−146) have been described previously.[25,34] For CA-NC mutant constructs (D94G/E98DS, D94G/H120S, D94G/H120N, D94G/H120G, D94G/H120R, and D94G/I124V), point mutations were introduced using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. Expression and purification of these CA-NC mutants was done as described previously for the wild-type protein.[25,34]

The cyclophilin A gene was subcloned into the pET29 plasmid (Novagen). The protein was expressed in BL21(DE3) *E. coli* cells (Novagen). Briefly, LB media was inoculated with overnight precultures and grown at 37 °C until mid log-phase (A_{600} ~0.6). Protein expression was induced with 0.5−1 mM isopropyl-β,D-thiogalactopyranoside (IPTG) for 4−6 h at 37 °C. Cells were harvested by centrifugation, and pellets were stored at −80 °C until purification. The bacterial pellet was resuspended in the extraction buffer [20 mM Tris-HCl pH 7.8, 1 mM PMSF, 1 mM TCEP] supplemented with Complete, EDTA-free protease inhibitors tablets (Roche), then sonicated, and centrifuged at 100,000*g* for 30 min to remove the cellular debris. The soluble CypA was precipitated with a 60% ammonium sulfate solution and then dissolved in 25 mM MES pH 6.5, 5 mM β-mercaptoethanol buffer. The protein was loaded on a HiTrapSP HP column (GE Healthcare) pre-equilibrated with the same buffer. Purified CypA protein was eluted from the column with the same buffer, supplemented with 500 mM NaCl. For NMR experiments, the protein was then buffer exchanged into a 25 mM sodium phosphate buffer, pH 5.5.

Isothermal Titration Calorimetry. Isothermal titration calorimetry was performed at 25 °C in 50 mM Tris pH 8.0, 350 mM NaCl, and 1% DMSO using a VP-ITC microcalorimeter (MicroCal Inc.; GE Health Sciences) and the Origin software and using the same protocol as described previously.[25]

NMR Sample Preparation. Unlabeled and ^{15}N-labeled samples of CA_{NTD} were provided at concentrations of 680 and 720 μM, respectively, in 25 mM sodium phosphate (pH 5.5), 2 mM DTT-deuterated. The CypA protein sample was provided at 636 μM in 25 mM sodium phosphate buffer, pH 5.5. $^1H-^{15}N$ NMR studies were conducted using 50 μM ^{15}N-labeled HIV-1 CA_{NTD} sample in 25 mM sodium phosphate (pH 5.5) in 90% H_2O/10% D_2O mixture. For the binding experiment (Figure 2), a small aliquot (~2 μL) of a 20 mM stock solution of inhibitor **1** in DMSO-d_6 was added to a protein sample, prepared as above such that a CA_{NTD}/compound ratio of 1:1.3 was achieved (a 2 μL aliquot of DMSO-d_6 was added for the reference spectrum of *apo* protein). For the titration experiment shown in Figure 2c, small aliquots of a 10 mM stock solution in DMSO-d_6 of compound **1** were successively added (0.75 μL, 0.75 μL, 0.75 μL, 0.75 μL, 3 μL, 3 μL) to a 25 μM ^{15}N-labeled HIV-1 CA_{NTD} sample in 25 mM sodium phosphate (pH 5.5) in 90% H_2O/10% D_2O mixture.

For the binding studies with cyclophilin A, 66 μL of the 636 μM concentrated stock of CypA was added to 42 μL of the 720 μM stock of ^{15}N-labeled CA_{NTD} to which were added 432 μL of 25 mM sodium phosphate buffer (pH 5.5), 2 mM DTT-deuterated, and 60 μL of D_2O. The sample was transferred into a 5 mm NMR tube. After recording a spectrum of this sample, 2.4 μL of the 20 mM stock solution of inhibitor **1** in DMSO-d_6 was added to the same NMR tube.

^{19}F NMR binding studies were conducted using a 50 μM solution of compound **2**, a 50 μM solution of compound **7**, or a 1:1 mixture solution (50 μM:50 μM) of compounds **2** and **7** to which was added increasing amount of HIV-1 CA_{NTD} such that compound(s)/NTD ratios of 50 μM:100 μM were achieved. All of these samples were prepared in 50 mM Tris buffer pH 7.0, 1 μM $ZnSO_4$ in 90% H_2O/10% D_2O mixture. A separate sample was also prepared containing 50 μM of **2**, 50 μM of **7**, 136 μM of CA_{NTD}, and 159 μM of CypA.

NMR Methods. $^1H-^{15}N$ NMR experiments were acquired on a 600 MHz Bruker AVANCE II spectrometer equipped with a 5 mm z-gradient triple resonance cryoprobe at T = 35 °C. Two-dimensional $^1H-^{15}N$ TROSY spectra were acquired using standard pulse sequences. Backbone resonances were assigned using a 450 μM $^{13}C,^{15}N$-labeled HIV-1 CA_{NTD} sample and were consistent with those reported in the literature.[13,34] ^{19}F NMR experiments were acquired on a 400 MHz Bruker AVANCE II spectrometer equipped with a 5 mm dual-$^1H,^{19}F$ probe at T = 37 °C. 1D ^{19}F spectra were acquired with proton decoupling. The number of scans varied between 3072 and 5120. All NMR data were processed and analyzed using Topspin (Bruker BioSpin). Binding constantfrom $^1H-^{15}N$ NMR titration experiment was calculated using Prism 4 software (GraphPad software). The ^1H and ^{15}N NMR chemical shift changes for residues Glu98, Ile104, Val126, and Thr119 were fit to a 1:1 binding and afforded an equilibrium dissociation constant (K_d) of 43 ± 2.9 μM.

X-ray Structure and Modeling. A full description of the structure solution of CA_{NTD} in complex with inhibitor **4** will be reported elsewhere (Lemke et al. *Acta Crystallogr., Sect. D*, accepted). Briefly, inhibitor **4** was co-crystallized in the presence of a second inhibitor, BD3 (an analogue of inhibitor **7** from benzodiazepine series), which binds to CAP-1 pocket. The resulting 1.7 Å resolution structure, PDB 4E91, clearly describes the binding of both inhibitor **4** as well as BD3. For model shown in Figure 5c, inhibitor **4** of the 4E91 structure was positioned onto the unliganded CA monomer of the hexameric CA structure (PDB 3MGE) by superposition of residues of the apical binding pocket (defined as less than 6 Å from inhibitor **4**). The hexamer of the superposed model was then generated by applying the appropriate crystallographic symmetry operations based on the hexagonal space group (P6) of 3MGE.

ASSOCIATED CONTENT

Supporting Information

This material is available free of charge *via* the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: RESGeneral.LAV@boehringer-ingelheim.com; Stephen.Mason@bms.com.

Present Addresses

[‖]Université de Montréal, Département de chimie, Montréal, Qc, Canada.

[⊥]AL-G Technologies, 201 Mgr. Bourget, Lévis, Québec, Canada.

[#]Bristol-Myers Squibb, Virology, 5 Research Parkway, Wallingford, CT 06492, USA.

Author Contributions

[§]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to thank W. Sundquist for the HIV-1 CA protein constructs as well as for very insightful discussions. We also would like to acknowledge V. Vogt for the RSV Gag ΔMBD ΔPR protein expression construct. We also extend our gratitude to E. Wardrop for viral replication assay data and M. Cartier for bioinformatics analyses. Finally we would like to thank P. Anderson, P. Bonneau, M. Bös, R. Bethell, and M. Cordingley for their leadership and guidance.

■ REFERENCES

(1) Sundquist, W. I., and H.-G. Kräusslich, H.-G. (2012) HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspect. Med. 2*, a006924.

(2) Ganser-Pornillos, B. K., Yeager, M., and Sundquist, W. I. (2008) The structural biology of HIV assembly. *Curr. Opin. Struct. Biol. 18*, 203−217.

(3) Prevelige, P. E., Jr. (2011) New approaches for antiviral targeting of HIV assembly. *J. Mol. Biol. 410* (4), 634−640.

(4) Waheed, A. A., and Freed, E. O. (2012) HIV type 1 Gag as a target for antiviral therapy. *AIDS Res. Hum. Retroviruses 28*, 54−75.

(5) von Schwedler, U. K., Stray, K. M., Garrus, J. E., and Sundquist, W. I. (2003) Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J. Virol. 77*, 5439−5450.

(6) Sundquist, W. I., and Hill, C. P.. (2007) How to assemble a capsid. *Cell 131*, 17−19.

(7) Gitti, R. K., Lee, B. M.., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science 273*, 231−235.

(8) Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell 87*, 1285−1294.

(9) Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science 278*, 849−853.

(10) Worthylake, D. K., Wang, H., Yoo, S., Sundquist, W. I., and Hill, C. P. (1999) Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 85−92.

(11) Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature 407*, 409−413.

(12) Byeon, I. J., Meng, X., Jung, J., Zhao, G., Yang, R., Ahn, J., Shi, J., Concel, J., Aiken, C., Zhang, P., and Gronenborn, A. M. (2009) Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. *Cell 139*, 780−790.

(13) Pornillos, O., Ganser-Pornillos, B. K., Kelly, B. N., Hua, Y., Whitby, F. G., Stout, C. D., Sundquist, W. I., Hill, C. P., and Yeager, M. (2009) X-ray structures of the hexameric building block of the HIV capsid. *Cell 137*, 1282−1292.

(14) Ganser-Pornillos, B. K., Cheng, A., and Yeager, M. (2007) Structure of full-length HIV-1 CA: a model for the mature capsid lattice. *Cell 131*, 70−79.

(15) Pornillos, O., Ganser-Pornillos, B. K., and Yeager, M. (2011) Atomic-level modelling of the HIV capsid. *Nature 469*, 424−427.

(16) Ganser-Pornillos, B. K., Yeager, M., and Pornillos, O. (2012) Assembly and maturation of HIV. *Adv. Exp. Med. Biol. 726*, 441−465.

(17) Forshey, B. M., von Schwedler, U., Sundquist, W. I., and Aiken, C. (2002) Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol. 76*, 5667−5677.

(18) Ganser-Pornillos, B. K., von Schwedler, U. K., Stray, K. M., Aiken, C., and Sundquist, W. I. (2004) Assembly properties of the human immunodeficiency virus type 1 CA protein. *J. Virol. 78*, 2545−2552.

(19) Tang, C., Loeliger, E., Kinde, I., Kyere, S., Mayo, K., Barklis, E., Sun, Y., Huang, M., and Summers, M. F. (2003) Antiviral inhibition of the HIV-1 capsid protein. *J. Mol. Biol. 327*, 1013−1020.

(20) Kelly, B. N., Kyere, S., Kinde, I., Tang, C., Howard, B. R., Robinson, H., Sundquist, W. I., Summers, M. F., and Hill, C. P. (2007) Structure of the antiviral assembly inhibitor CAP-1 complex with the HIV-1 CA protein. *J. Mol. Biol. 373*, 355−366.

(21) Blair, W. S., Pickford, C., Irving, S. L., Brown, D. G., Anderson, M., Bazin, R., Cao, J., Ciaramella, G., Isaacson, J., Jackson, L., Hunt, R., Kjerrstrom, A., Nieman, J. A., Patick, A. K., Perros, M., Scott, A. D., Whitby, K., Wu, H., and Butler, S. L. (2010) HIV capsid is a tractable target for small molecule therapeutic intervention. *PLoS Pathog. 6*, e1001220.

(22) Shi, J., Zhou, J., Shah, V. B., Aiken, C., and Whitby, K. (2011) Small-molecule inhibition of human immunodeficiency virus type 1 infection by virus capsid destabilization. *J. Virol. 85*, 542−549.

(23) Sticht, J., Humbert, M., Findlow, S., Bodem, J., Müller, B., Dietrich, U., Werner, J., and Kräusslich, H. G. (2005) A peptide inhibitor of HIV-1 assembly in vitro. *Nat. Struct. Mol. Biol. 12*, 671−677.

(24) Fader, L. D., Bethell, R., Bonneau, P., Bös, M., Bousquet, Y., Cordingley, M. G., Coulombe, R., Deroy, P., Faucher, A.-M., Gagnon, A., Goudreau, N., Grand-Maître, C., Guse, I., Hucke, O., Kawai, S. H., Lacoste, J.-E., Landry, S., Lemke, C. T., Malenfant, E., Mason, S., Morin, S., O'Meara, J., Simoneau, B., Titolo, S., and Yoakim, C. (2011) Discovery of a 1,5-dihydrobenzo[b][1,4]diazepine-2,4-dione series of inhibitors of HIV-1 capsid assembly. *Bioorg. Med. Chem. Lett. 21*, 398−404.

(25) Lemke, C. T., Titolo, S., von Schwedler, U., Goudreau, N., Mercier, J.-F., Wardrop, E., Faucher, A.-M., Coulombe, R., Banik, S. S., Fader, L., Gagnon, A., Kawai, S. H., Rancourt, J., Tremblay, M., Yoakim, C., Simoneau, B., Archambault, J., Sundquist, W. I., and Mason, S. W. (2012) Distinct effects of two HIV-1 capsid assembly inhibitor families that bind the same site within the N-terminal domain of the viral CA protein. *J. Virol. 86*, 6643−6655.

(26) Tremblay, M., Bonneau, P., Bousquet, Y., DeRoy, P., Duan, J., Duplessis, M., Gagnon, A., Garneau, M., Goudreau, N., Guse, I., Hucke, O., Kawai, S. H., Lemke, C. T., Mason, S. W., Simoneau, B., Surprenant, S., Titolo, S., and Yoakim, C. (2012) Inhibition of HIV-1 capsid assembly; optimization of the antiviral potency by site selective modifications at N1, C2 and C16 of a 5-(5-furan-2-yl-pyrazol-1-yl)-1H-benzimidazole scaffold. *Bioorg. Med. Chem. Lett. 22*, 7512−7517.

(27) Bocanegra, R., Rodriguez-Huete, A., Fuertes, M. A., Del, A. M., and Mateu, M. G. (2012) Molecular recognition in the human immunodeficiency virus capsid and antiviral design. *Virus Res. 169*, 388−410.

(28) Ma, Y. M., and Vogt, V. M. (2004) Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. *J. Virol. 78*, 52−60.

(29) Ternois, F., Sticht, J., Duquerroy, S., Kräusslich, H. G., and Rey, F. A. (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat. Struct. Mol. Biol. 12*, 678−682.

(30) Shin, R., Tzou, Y.-M., and Krishna, N. R. (2011) Structure of a monomeric mutant of the HIV-1 capsid protein. *Biochemistry 50*, 9457−9467.
(31) Hajduk, P. J., Boyd, S., Nettesheim, D., Nienaber, V., Severin, J., Smith, R., Davidson, D., Rockway, T., and Fesik, S. W. (2000) Identification of novel inhibitors of urokinase via NMR-based screening. *J. Med. Chem. 43*, 3862−3866.
(32) Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell 87*, 1285−1294.
(33) Matsuoka, S., Dam, E., Lecossier, D., Clavel, F., and Hance, A. J. (2009) Modulation of HIV-1 infectivity and cyclophilin A dependence by Gag sequence and target cell type. *Retrovirology 6*, 21.
(34) Goudreau, N., Coulombe, R., Faucher, A.-M., Grand-Maître, C., Lacoste, J.-E., Lemke, C. T., Malenfant, E., Bousquet, Y., Fader, L., Simoneau, B., Mercier, J.-F., Titolo, S., and Mason, S. W. (2013) Monitoring binding of HIV-1 capsid assembly inhibitors using ^{19}F ligand-based & ^{15}N protein-based NMR and X-ray crystallography: Early hit validation of a new benzodiazepine series. *ChemMedChem 8*, 405−414.
(35) Dalvit, C. (2007) Ligand- and substrate-based ^{19}F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc. 51*, 243−271.
(36) Dalvit, C., and Vulpetti, C, A. (2011) Fluorine−Protein interactions and ^{19}F NMR isotropic chemical shifts: An empirical correlation with implications for drug design. *ChemMedChem 6*, 104−114.
(37) Pellecchia, M., Becattini, B., Crowell, K. J., Fattorusso, R., Forino, M., Fragai, M., Jung, D., Mustelin, T., and Tautz, L. (2004) NMR-based techniques in the hit identification and optimization processes. *Expert Opin. Ther. Targets 8*, 597−611.
(38) Pornillos, O., Ganser-Pornillos, B. K., Banumathi, S., Hua, Y., and Yeager, M. (2010) Disulfide bond stabilization of the hexameric capsomer of human immunodeficiency virus. *J. Mol. Biol. 401*, 985−995.
(39) Bharat, A. M., Davey, N. E., Ulbrich, P., Riches, J. D., de Marco, A., Rumlova, M., Sachse, C., Ruml, T., and Briggs, J. A. G. (2012) Stucture of the immature retroviral capsid at 8 Å resolution by cryo-electron microscopy. *Nature 487*, 385−389.