The CA protein plays critical roles in the early and late phases of replication and has long been considered an attractive potential therapeutic target.[9] CA is originally synthesized as a 231 amino acid domain within the 55 kDa Gag precursor polyprotein. During the late phase of viral replication, the CA domain helps mediate the assembly of ~4000 copies of Gag into the immature virus particle.[10] Subsequent processing by the viral protease triggers conformational changes in CA that promote its assembly into the capsid, a conical protein shell composed of about 1500 CA molecules that encloses two copies of the viral genome and the viral enzymes essential for infectivity. Proper assembly of this core particle is critical for viral replication, and mutations that reduce or increase core stability lead to dramatic reductions in viral infectivity.[11] CA comprises two domains[12] that have distinct roles in stabilizing the viral capsid architecture.[13–15] In cylindrical *in vitro* assemblies that mimic viral capsids, the N-terminal domain (CA^N, residues 1–146) forms hexamers and the C-terminal domain (CA^C, residues 147–231), which is dimeric in solution,[16] links adjacent hexamers.[14] A crystal structure of the murine leukemia virus (MLV) CA^N protein has enabled atomic-level modeling of the HIV-1 CA^N hexamer.[17] Crystal structures are also available for the CA^C dimer,[16,18] although a domain-swapped dimer model has been proposed from analogy with the structurally related SCAN domain.[19,20] Finally, biochemical studies indicate that the N and C-terminal domains form intermolecular contacts in the mature capsid lattice.[21–24]

Disruption of capsid reorganization can be an effective approach to viral inhibition, with examples including targeting of capsid assembly in hepatitis B virus (HBV)[25] and blocking of capsid disassembly in picorona viruses.[26] Support for this approach for HIV-1 was provided by CAI, a peptide inhibitor of CA^C–CA^C interactions that inhibits immature and mature particle formation *in vitro*,[27,28] although CAI was unable to inhibit the release of HIV-1 particles when added to virus producing cells in cell culture or by peptide transfection. Similarly, betulinic acid appears to impede HIV maturation by binding to an assembled form of Gag and blocking processing of the C terminus of CA by the viral protease.[29] Small molecule inhibitors were also identified in a search for agents that bind directly to HIV-1 CA.[30] One of these compounds, *N*-(3-chloro-4-methylphenyl)-*N*′-{2-[({5-[(dimethylamino)-methyl]-2-furyl}-methyl)-sulfanyl]ethyl}urea (CAP-1) inhibits capsid assembly *in vitro*, HIV-1 infectivity *in vivo*, and leads to the production of poorly infectious virons with abnormal core morphologies.[30] NMR chemical shift mapping experiments revealed that CAP-1 binds to a site near the C-terminal end of CA^N,[30] and to date, more than two dozen additional compounds have been identified that bind specifically to this site (M. F. S., unpublished results).

In an effort to understand the molecular basis for inhibition by CAP-1, we determined the three-dimensional structure of the complex with CA^N using a combination of X-ray crystallography and NMR spectroscopy. The two approaches provide complementary information that was useful for addressing problems associated with relatively weak binding and poor ligand solubility. CAP-1 binds to a deep hydrophobic cavity that is formed upon extrusion of the Phe32 side-chain from a buried to an exposed position. Intermolecular interactions that appear important for binding, and additional elements that may be exploited to enhance binding affinity, are identified in the structure. The displacement of Phe32 appears to be facilitated by a strained main-chain conformation, which has implications for both the mechanism of capsid assembly and its inhibition by CAP-1.

Results and Discussion

Structure of CA^N crystallized in the presence of CAP-1

The HIV-1 CA^N protein (residues 1–146, wt $HIV\text{-}1_{NL4\text{-}3}$ sequence) was expressed recombinantly in *Escherichia coli*, purified, and crystallized in the presence of CAP-1 at concentrations above the binding constant. Although the structure was determined at 1.5 Å resolution and conformational changes that appear to open the CAP-1 binding site were observed (discussed below), density for CAP-1 was not visible. CAP-1 binds with modest affinity (~800 μM) and is poorly soluble, and our preferred explanation is that the CAP-1 that crystallized with the protein diffused out of the binding site and precipitated during crystal growth, while the ligand-bound CA^N conformation was retained because lattice contacts with a neighboring molecule occur in this region of the structure (Supplementary Data, Figure S1). It is possible that CAP-1 bound in multiple related conformations that reduce the interpretable electron density, although we do not see the residual density that might be expected at medium/low resolution for this situation. A third possibility is that the structure observed arose simply because of favorable lattice forces in the new crystal form, although this seems unlikely because we were unable to grow crystals from these conditions in the absence of CAP-1. Regardless of the reason for the missing CAP-1 density, the functional relevance of the crystal structure as the conformation that binds CAP-1 is indicated by independent NMR data that are discussed below.

Comparison of the crystal structure with other X-ray and NMR determinations of CA^N indicated that the majority of the protein residues were unaffected by the presence of CAP-1.[12,31,32] However, a significant conformational change was observed for residues that earlier NMR heteronuclear single quantum coherence (HSQC) titration experiments had indicated were in the vicinity of the