symmetry. The diffraction pattern could therefore be indexed either in the true space group or in the pseudo-space group in which the systematically weak reflections are ignored and the *c* axis length is halved. Both the true and pseudo-cells belong to space group *P*1. The structure was initially solved by molecular replacement in the pseudo-space group and the four molecules of that asymmetric unit were refined before confirming the true cell structure by molecular replacement and independent refinement of all eight molecules. As expected for large numbers of systematically weak reflections, some statistics for the pseudo-cell (R_{free} = 21.7%; resolution = 1.45 Å) appear better than for the true cell (R_{free} = 26.8%; resolution = 1.90 Å). Resolution is defined as the Bragg spacing at which half of the measured reflections are less than twice their estimated standard deviation. Nevertheless, refinement statistics for the true cell indicate that this structure is of good quality (Table 2).

NMR data collection and analysis

Samples of CA^N containing a C-terminal His-6 tag were prepared using a modification of a described protocol.[30] The gene encoding the N-terminal domain (NTD) of HIV-1 CA (residues 1 through 151) with a C-terminal His-6 tag was PCR-amplified from pNL4-3, sub-cloned into pET-11a (Novagen) and subsequently transformed into BL21 competent cells (Stratagene). Cells were grown in LB medium or M9 minimal medium containing $^{15}NH_4Cl$ and/or[^{13}C]glucose (Isotec) as its sole nitrogen and/or carbon source. Protein expression was induced in shake flasks at an A_{600} of 0.6 with 1 mM IPTG. The cells were harvested and lysed with a microfluidizer (Microfluidics) and the protein was purified to homogeneity by cobalt affinity (Talon) and cation exchange (Amersham) chromatographies.

NMR data were collected on Bruker 600 and 800 MHz instruments equipped with cryo-probes at 11 °C or 35 °C using protein samples of 100 μM –700 μM CA^N in 25 mM sodium phosphate (pH 7.0), 5 mM DTT, 10% ^2H_2O, and 5% DMSO-d_6. NMR signals of CAP-1 were assigned by standard 2D NMR methods.[57] Resonances of the protons attached to carbons in free CAP-1 were assigned using natural abundance ^1H,^{13}C HMQC and 2D homonuclear NOESY data collected in 100% DMSO-d_6 at 35 °C. The signals of these protons of CAP-1 in the complex were assigned in H_2O by titrating increasing amounts of CAP-1 (0–3 mM) into 700 μM CA protein and monitoring with 2D homonuclear NOESY (τ_{mix} = 120 ms) experiments. Intermolecular ^1H-^1H NOEs for the complex were obtained with 3D ^{13}C-edited HMQC-NOESY, ^{15}N-edited NOESY-HSQC and 2D homonuclear NOESY data. NMR data were processed with NMRPIPE[58] and analyzed with NMRVIEW.[59]

Joint NMR/X-ray structure determination of the CAP-1:CA^N complex

Refinement of the CAP-1:CA^N complex was performed using the AMBER-9 program package[60] by docking a CAP-1 model into the cavity of the X-ray structure and performing restrained molecular dynamics followed by

Table 2. Crystallographic and refinement data

	CA^N(A92E) real cell	CA^N(A92E) pseudo-cell	CA^N (+ CAP-1)
Data collection	NSLS X29	NSLS X29	NSLS X12-B
Space group	*P*1	*P*1	$C222_1$
Unit cell lengths: *a*, *b*, *c* (Å)	48.3, 59.0, 92.3	46.2 48.2 58.9	42.2, 62.8, 106.3
Unit cell angles: α, β, γ (°)	71.5, 88.1, 83.0	83.0, 71.3, 87.4	90, 90, 90
Wavelength (Å)	1.1000	1.1000	1.0000
Resolution (Å)	32.2–1.90 (1.97–1.90)	30.0–1.45 (1.50–1.45)	53–1.50 (1.54–1.50)
Number of observed reflections	1, 156, 031	579, 474	255, 759
Number of unique reflections	74, 959	84, 901	23, 354
Completeness (%)	96.5 (95.4)	91.0 (61.6)	95.2 (72.5)
R_{sym} (%)	5.9 (12.4)	5.6 (29.4)	4.1 (42.4)
Average $I/\sigma(I)$	10 (2)	12 (2)	10 (2)
Mosaicity (°)	0.40	0.35	0.50
Refinement			
Working R_{factor} (%)	0.204	0.170	0.163
R_{free} (%)	0.270	0.216	0.221
$R_{overall}$ (%)	0.208	0.173	0.166
No. non-hydrogen atoms	9946	5366	1,374
Number of water molecules	797	533	179
RMSD: Bond lengths (Å)	0.020	0.017	0.020
Bond angles (°)	1.724	1.664	1.873
ϕ/ψ angles, non-Gly/Pro res.			
Most favored regions (%)	94.5	93.8	93.4
Additional allowed (%)	5.3	5.8	5.8
Generously allowed (%)	0.2	0.2	0.8
Disallowed regions (%)	0.0	0.2	0.0
<*B*>: main-chain atoms (Å^2)	16.4	17.5	21.3
side chain atoms (Å^2)	19.2	20.4	24.4
Water molecules (Å^2)	26.3	29.8	36.4

Values in parentheses refer to the high-resolution shell.
$R_{sym} = \sum (I - \langle I \rangle) / \sum (I)$.
R-factor = $\sum_{hkl} | F_o - F_c | / | F_o |$. R_{free} is as for $R_{working}$ but calculated for a randomly selected 5% of reflections not included in the refinement. $R_{overall}$ is using all reflections ($R_{free} + R_{working}$).