CryoEM Structure Refinement by Integrating NMR Chemical Shifts with Molecular Dynamics Simulations

Juan R. Perilla[1,*,†], **Gongpu Zhao**[2,3,†,‡], **Manman Lu**[3,4], **Jiying Ning**[2,3], **Guangjin Hou**[3,4], **In-Ja L. Byeon**[2,3], **Angela M. Gronenborn**[2,3], **Tatyana Polenova**[3,4], and **Peijun Zhang**[2,3,5,6,*]

[1]Department of Physics and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[2]Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA

[3]Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine, Pittsburgh, PA 15260, USA

[4]Department of Chemistry and Biochemistry, University of Delaware, Newark, DE 19716, USA

[5]Division of Structural Biology, University of Oxford, The Henry Wellcome Building for Genomic Medicine, Headington, Oxford OX3 7BN, UK

[6]Electron Bio-Imaging Centre, Diamond Light Sources, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK

Abstract

Single particle cryoEM has emerged as a powerful method for structure determination of proteins and complexes, complementing X-ray crystallography and NMR spectroscopy. Yet, for many systems, the resolution of cryoEM density map has been limited to 4–6 Å, which only allows for resolving bulky amino acids side chains, thus hindering accurate model building from the density map. On the other hand, experimental chemical shifts (CS) from solution and solid state MAS NMR spectra provide atomic level data for each amino acid within a molecule or a complex; however, structure determination of large complexes and assemblies based on NMR data alone remains challenging. Here we present a novel integrated strategy to combine the highly complementary experimental data from cryoEM and NMR computationally by molecular dynamics simulations to derive an atomistic model, which is not attainable by either approach alone. We use the HIV-1 capsid protein (CA) C-terminal domain as well as the large capsid assembly to demonstrate the feasibility of this approach, termed NMR CS-biased cryoEM structure refinement.

*Correspondence: Telephone: +44 (0)1865287567; pez7@pitt.edu (P.Z.), Telephone : +1 (217)-244-7403 ; jperilla@illinois.edu (J.R.P.).

†These authors contributed equally to this work

‡Current address: Van Andel Research Institute, 333 Bostwick Ave N.E., Grand Rapids, MI 49503, USA

Introduction

Remarkable advances in cryo-electron microscopes, direct electron detection devices and image data analysis software within the last few years have advanced cryoEM methodolgy to a stage where for many large proteins, macromolecular complexes and assemblies [1, 2] it has become the method of choice for structure determination, complementing NMR spectroscopy and X-ray crystallography [3, 4]. While some spectacular examples of cryoEM structures at atomic resolutions have been reported [5], for many systems it remains a challenge to reach atomic details, especially for every part of the molecule and certainly for flexible regions [6–8]. It is anticipated that a large number of cryoEM density maps will be reconstructed at 4–6 Å resolution level. At this level of resolution, the overall architecture of the system, secondary structures, as well as large side chains or bases will be clearly resolved, but it will remain challenging to refine every part of the structure. A number of methods have recently put forth to aid cryoEM structure refinement, such as those inspired by X-ray crystallographic refinement programs, Phenix [9, 10] and RefMac [11], or molecular dynamics simulations flexible fitting [12, 13] or a combination of molecular dynamics with Rosetta refinement [14, 15]. However, without real experimental data at the atomic level, the models derived from these refinement procedures fall short on accuracy and may contain ambiguities.

On the other hand, experimental chemical shift measurements from solution and solid state NMR spectra provide atomic level details for each amino acid within a molecule. Yet, for many large molecules, complexes and assemblies, it is necessary not only to obtain complete or almost complete resonance assignments but also a large number of distance and angular restraints. However, distance and angular restraint measurements and data analysis remain laborious both for solution and particularly for solid-state NMR. While great strides have been made to develop 3- and 4-dimensional experiments for resolution enhancements of magic angle spinning (MAS) NMR spectra [16–18], and ultrafast MAS in conjunction with ^{1}H detection has further enabled structure determination by solid state NMR [19, 20], experiments and data interpretation still take many months. Nonetheless, given that the data from cryoEM density maps and NMR chemical shifts are highly complementary, integrating the structural information from both methods, combined with all-atom molecular dynamics simulations, provides a powerful means to solving structures of systems that are otherwise not tractable using any single structural biology technique. Indeed, recent reports demonstrated successful incorporation of cryoEM data as constrains into the refinement of solution and solid state NMR structure using Xplor-NIH [21, 22], or Rosetta modelling [23, 24].

State of the art computational methodologies have reached the level where they can be used as powerful tools for the structural refinement of large macromolecular complexes and cellular systems [12, 13]. For example, computational techniques can be used to predict protein structure from sequence, e.g., using Rosetta [25] or MU-Fold [26]; refine structures by incorporating experimental data using molecular dynamics flexible fitting (MDFF) [12, 13] and study dynamic properties of proteins using molecular dynamics (MD) [27–29]. MD-based techniques can incorporate multiple experimental observables as external forces during the simulation thus obtaining atomistic details of protein structure and dynamics. Notable examples of large complexes that were determined by computer-guided methods and

experimental measurements include structures of the ribosome, the proteasome, virus capsids, the cytoplasm and the chemosensory array [30–35].

HIV-1 capsid is the protein shell that encloses the viral genome and plays critical roles in many steps of viral life cycle, including assembly, maturation, uncoating, reverse transcription, and trafficking to the cell nucleus. Capsid therefore represents an attractive target for anti-HIV drug development. A number of atomic structures for the HIV-1 capsid protein (CA) monomer, hexamer and pentamer have been determined [4, 36–39], but not in the assembled state. We previously determined a cryoEM structure of the HIV-1 capsid (CA) tubular assembly at 8 Å resolution [34]. Combining this structure with cryo-electron tomography of a native core and all-atom molecular dynamics simulations, we derived a computational model for the entire viral capsid [34]. We now report a high-resolution (5 Å) cryoEM density map of a tubular capsid assembly. The map shows densities for many bulky amino acid side chains, permitting the identification of intermolecular interactions in this assembly. In addition, well dispersed MAS NMR spectra recorded for the same assembly [40], permitted assignments of 96% of the backbone and side chain resonances. Since cryoEM density maps and NMR chemical shifts represent orthogonal structural data, a novel approach was developed, which combines the strengths of both methodologies, i.e., using the chemical shifts from MAS NMR spectra to guide the cryoEM structure model refinement. This was done in an exemplary fashion for a tubular capsid assembly. Since the capsid assembly is very large and computationally expensive, we first developed and validated our integrated method on the CA C-terminal domain (CA-CTD), for proof-of-principle. We further used the approach to refine the cryoEM structure of the CA hexamer in the tubular assembly, the main basic assembly unit for the viral capsid. We are confident that the NMR chemical shift-guided approach for cryoEM refinement will have broad impact to both cryoEM and NMR based structural analysis of large protein assemblies, extending beyond HIV-1 research.

Methods

Protein expression and purification

The NL4-3 A92E HIV-1 capsid protein (CA) was expressed in *E. coli* Rosetta 2 (DE3) cells, purified by ion exchange chromatography, as described previously, with a minor modification [41]. Briefly, cell lysate was loaded onto a 5 mL Hi-Trap SP column (GE Healthcare) equilibrated with a buffer containing 25 mM sodium phosphate, pH 5.8, 0.02% sodium azide and 1 mM DTT. The protein was eluted with a linear gradient of 0–1 M NaCl at a flow rate of 2 mL/min. The fractions containing CA protein were pooled and loaded onto a Hi-Load Superdex75 26/60 column (GE Healthcare) equilibrated with a buffer containing 25 mM sodium phosphate, pH 7.5, 0.02% sodium azide and 1 mM DTT at a flow rate of 2 mL/min. The protein was concentrated to 12 mg/ml using Amicon concentrators (Millipore) and stored at −80°C until used.

CryoEM specimen preparation and data collection

CA A92E protein (12 mg/ml) stock was diluted to 2 mg/ml in high salt buffer (1M NaCl, 50mM Tris pH 8.0) and incubated at 37°C for 1 hr for tubular assembly. The assembled

sample (2.5 μl) was applied to the carbon side of a glow discharged perforated Quantifoil grid (Quantifoil Micro Tools, Jena, Germany), followed by application of 3 μl of low salt buffer (100 mM NaCl, 50mM Tris pH 8.0) on the back side of the grid, and blotting, from the back side, with a filter paper, before plunge-freezing in liquid ethane, using a manual gravity plunger. Low dose projection micrographs of tubes embedded in a thin layer of vitreous ice were collected on a Tecnai Polara microscope (FEI Corp., OR.), operated at 300 kV at liquid nitrogen temperature. Images were recorded on a Gatan K2 summit direct electron detector camera at a nominal magnification of 31,000 ×, using super-resolution counting mode. The images were binned by 2 (bin 2) during image processing. The pixel size of bin 2 image is 1.22 Å. The dose rate was set to ~8 electrons per physical pixel per second. The total accumulative electron dose is ~41 electrons/$Å^2$ fractioned over 30 subframes with a total exposure time 6 seconds. The defocus values were from 0.7 to 2.2 μm.

Image processing and 3D reconstruction

Each tube was boxed out from the bin 2, and the whole micrograph was motion corrected using e2helixboxer program in EMAN2 [42]. All the tubes were sorted into different groups based on their diameters. Only the most abundant group was further analyzed, which mainly encompasses the (−12,11) helical symmetry group. These tubes were then boxed out from the unbinned 30 subframes movie with a fixed width of 1600 pixels to minimize interpolation. A total of 523 tubes with (−12, 11) symmetry were selected. An in-house script was used to generate the coordinate file for particles of 456×456 pixels with 95 percent overlap from the bin 2 movie. Each particle contained 6 asymmetric subunits. A total 40528 movies of individual particles were extracted from the bin 2 movies, using IMOD, motion corrected and exposure filtered using unblur, and normalized using RELION 1.4 [43]. All particles were subject to two rounds of reference-free 2D classification (with 80 classes) and bad classes containing ice contamination were discarded. Total 39712 particles were selected for auto-3D refinement using Relion 1.4 [43]. The reference density map was generated using frealign [44] and filtered to 15Å resolution. An asymmetric density map at ~6 Å resolution was reconstructed by 3D auto-refine. This density map was filtered to 15Å and used as a reference for 3D classification (with 4 classes), using only local angle searches with an angular step of 0.5° and angular range of 5°. The 3D classification generated three good classes with 42%, 28% particles and one bad class with 26% particles. The three good classes were individually refined using 3D auto-refine. Helical symmetry was imposed in real space by rotating and translating the map by Φ and z and averaged with itself. Two-fold symmetry was applied by rotating a hexamer by 180 degrees and averaged with itself. The final density map was reconstructed post processing by merging the two half maps in Relion, followed by sharpening and filtering. The final resolution was 5 Å, based on the Fourier Shell Correlation (FSC) value of 0.143.

MAS NMR sample preparation and data collection

Tubular assemblies of CA A92E were prepared by adding NaCl solution to 32 mg/ml protein solution to a final salt concentration of 1 M, followed by one-hour incubation at 37 °C. Tubular assemblies were collected by centrifugation at 10,000 g, and the pellet was packed into a Bruker 3.2 mm thin-wall rotor.

MAS NMR experiments were acquired on a Bruker 19.9 T narrow bore AVIII spectrometer outfitted with a 3.2 mm BioMAS HXY probe. Larmor frequencies were 850.400 MHz for ^{1}H, 213.855 MHz for ^{13}C, and 86.181 MHz for ^{15}N. The MAS frequency was maintained at 13.000 ± 0.001 kHz by a Bruker MAS controller. The sample temperature was calibrated using KBr as a temperature sensor, and was controlled at 4 ± 0.1 °C throughout the experiments by a Bruker temperature controller.

Typical 90° pulse lengths were 3.3 μs for ^{1}H and 3.8 μs for ^{13}C. ^{1}H–^{13}C cross-polarization was performed with contact time of 1 ms, and a linear amplitude ramp of 80–100% on ^{1}H, with the ^{1}H rf field of 80 kHz matched to the second spinning sideband. Small phase incremental alternation (SPINAL-64) ^{1}H decoupling (80 kHz) [45] was applied during the evolution and acquisition periods. The ^{1}H field strength during DARR was 13 kHz, with a DARR mixing time of 50 ms. States-TPPI [46] was used for frequency discrimination in the indirect dimension.

MDFF of CA hexamers

The hexameric structure of CA (PDB: 4XFX) was docked into the cryoEM density using Chimera. Crystallographic waters were kept for MDFF. Protonation states of titratable groups and the coordinates of hydrogens belonging to the CA hexamer and crystallographic waters were calculated and added using PDB2PQR [47]. Subsequently, using CIONIZE [48] in VMD [49], neutralizing ions were added to the CA hexamer and the crystallographic waters. The simulation system was then fully solvated in a water box of dimensions $126 \times 133 \times 133$ Å. In addition, bulk ions were added to a concentration of 1M NaCl. Molecular dynamics flexible fitting (MDFF) simulations were carried out using the r-RESPA integrator available in NAMD2.10 [50]. Long-range electrostatic force calculations employed the PME (particle-mesh Ewald) method, with a grid spacing of 1.0 Å and 4th order interpolation and a 1.2 nm cutoff. Simulations used the CHARMM36 force field [51] with the TIP3P water model [52], using a time-step of 2 fs, with non-bonded interactions evaluated every 2 fs and electrostatics updated every 4fs. All hydrogen bonds were constrained with the SHAKE algorithm. Secondary structure and chiral restraints were applied with a force constant of 1 kcal/A. MDFF was performed in two steps: 1) the backbone atoms were coupled to the density with a coupling constant ramping from 0.05 to 1.0 over 5ns. 2) bulky residues, in particular phenylalanine, tryptophan isoleucine, leucine, tyrosine, arginine as well as the backbone of the protein were coupled to the density with a coupling constant ramping from 0.05 to 1.0 over 5 ns. A second model was produced by optimizing the MDFF-derived structure using FastRelax in Rosetta. The FastRelax algorithm optimizes the model by applying a gradient minimization to all torsional degrees of freedom (ϕ, ψ, ω, and χ) followed by a simulated annealing rotamer search.

SHIFTX2 and model prediction

SHIFTX2 [53] was employed to predict the backbone and side chain ^{13}C chemical shifts using two MDFF-derived models of hexameric CA. The SHIFTX2 calculation was carried out in the multi-chain mode, in which the hexameric CA was treated as a single long chain. The temperature was set to 277 K, and the pH was set to 6.0.

Density map-based, CS-biased MD model refinement

All NMR chemical-shift biased MD (CSMD) simulations were performed using NAMD 2.9, with an integration timestep of 1 fs, bonded interactions computed for every timestep, and electrostatics updated every 4 fs. Particle-mesh-Ewald was used for long-range electrostatics with a grid size of 1 Å. In addition, the CHARMM22 force-field [54] with CMAP correction [55] and the generalized Born implicit solvent model as implemented in NAMD were used for all CSMD simulations [56]. One of the monomers in the crystal structure of the C-terminal domain of CA (PDB code: 1A43) was used as the starting structure. Chemical shifts corresponding to the dimeric form of the C-terminal domain of CA in solution were used in the following manner: At every MD time-step, chemical shifts were calculated for the current structure using the program Almost (CS2BACKBONE and CH3SHIFTS) [57]. Subsequently, the difference between the predicted and measured CSs was incorporated as a linear potential centered around 0.0 and a slope of 12.0 using the software PLUMED [58] and NAMD, over 1 ns [50]. Backbone heavy atoms ($C\alpha$, $C\beta$, C and N), as well as backbone hydrogens ($C\alpha$-H and N-H) for residues 1, 3, 4, 6, 18, 24, 25, 27, 30, 34, 39, 41, 42, 43, 44, 47, 53, 54, 55, 57, 58, 62, 63, 64, 69 and 70, were employed for the CS-biasing of the CTD.

The MDFF-derived model of hexameric CA was used subjected to CS-biased refinement. In particular, MAS chemical shifts for assembled tubes of CA were used as a bias for the backbone heavy atoms ($C\alpha$, $C\beta$, C and N). Therefore, chemical shifts were calculated for each time-step using the program Almost (CS2BACKBONE). The difference between the predicted and measured CSs was incorporated as a linear potential centered around 0.0 and with a slope of 1.0 using the software PLUMED and NAMD, over 1 ns.

Results and Discussion

Overview of CS biasing strategy

CryoEM densities and NMR chemical-shifts can be readily used in molecular-dynamics simulations for refinement of protein structure. Typically, these measurements are integrated as an external potential during the simulation. For instance, cryoEM densities are incorporated as a grid-based biasing-potential in MDFF [13]. NMR chemical-shifts, on the other hand, can be included in the form of a knowledge-based potential [59]; from such potential, forces are calculated based on the deviations in chemical shifts between the atomic coordinates and the NMR measurements [57–59]. Remarkably, successful in-silico folding of proteins, based solely on CS, has been reported in the literature [60], albeit this approach so far is limited to small, globular proteins. Here we propose and present an approach that combines data from complementary experiments in order to obtain a holistic picture of molecular structure.

In our approach, an experimentally or computationally derived structure is refined by incorporating cryoEM and NMR restraints in an integrative fashion using a molecular dynamics force-field (Figure 1). Since a complete structure is required as a starting point, missing loops and medium-sized structural motifs need to be generated *de novo* and added to the structure by using specialized software (e.g., Modeler, MU-fold, and Rosetta) [25, 26]. Subsequently, using 4–6 Å resolution cryoEM density maps, all backbone and side chains

atoms that are clearly defined in the density are fitted into the map using the MDFF protocol [12, 13]. In order to allow for variation in secondary structure (e.g., helix bending or helix kinks), the MDFF-refined model is further refined adding NMR chemical shift derived backbone restraints. These NMR restraints are incorporated during each time-step of the simulation. For subsequent cross-validation of the final model, a subset of chemical shifts is omitted from the biasing step. The resulting model can be used as the new starting model for further iterative refinement, including fitting to the experimental density map and optimization of side-chain rotamers using knowledge-based potentials [31].

CS-biased model refinement of HIV-1 CA-CTD

We tested the workflow and algorithms using a small model system with a known NMR structure, since such a model system permits testing the approach and validating the derived structural model by comparing it to the 'ground truth', the solution NMR structure. HIV-1 CA-CTD, was selected as the model system since its NMR solution structure was previously solved by us (PDB code: 2KOD) (Figure 2A, orange) and its relevance to our target, the full-length CA assembly [4]. A simulated density map was generated at 5 Å resolution in order to mimic the cryoEM density map (Figure 2B). A different CA-CTD model determined by X-ray crystallography (PDB code: 1A34) (Figure 2C grey) [61], exhibiting a heavy atoms backbone RMSD of 1.5 Å to the NMR solution structure (Figure 2C orange), was used for MDFF fitting into the simulated density map. Using this starting MDFF model, we applied chemical shifts from the solution NMR spectrum (methyl-containing residues) to C and H atoms as a biasing potential during MDFF model refinement. The resulting refined model, shown in blue in Figure 2D, clearly deviates from the initial X-ray structure (Figure 2D grey), and more closely resembles the 'ground truth' solution NMR structure (Figure 2D orange), with heavy atoms backbone RMSDs of 1.4 Å and 0.9 Å, respectively. The CS-biased model refinement converges rapidly from an heavy atoms backbone RMSD of 1.5 Å to less than 1.0 Å in less than 0.5 ns (Figure 2E). Note, that due to limitations in Almost (CH3SHIFTS) [57] only a subset of chemical shifts corresponding to residue types which contain methyl-groups, namely Ala, Ile, Leu, Met, Thr, and Val (listed in the methods section, Figure 2F with "*"), were included in the CS-biased simulation. To measure the performance of this approach, we further calculated predicted chemical shifts from the resulting model using SHIFTX2 [53], and compared these to the experimental NMR chemical shifts for every residue in the CA-CTD molecule (Figure 2F). The overall chemical shift deviations are 0.1 ppm for Cα-H, 0.4 ppm for Cα, and 0.3 ppm for Cβ, respectively.

CryoEM structure of the CA tubular assembly

Previously we investigated the NL4-3 CA A92E tubular assembly and obtained a 3D density map at 8 Å resolution, using image data collected on photographic film [34]. Taking advantage of the K2 summit direct electron detecting device, which allows for drift-correction using the movie mode capturing and electron counting in super-resolution mode, we recorded a large number of high-resolution movie stacks on the Polara electron microscope. A total of 523 tubes of a single helical family (−12,11), based on their Fourier transforms (for helical symmetry see reference [62]) were selected and boxed out into small segments that were treated as single particles for structural analysis. Using an iterative real-space helical 3D reconstruction method [63], about 40,000 helical segments were processed to derive the final

density map at ~5 Å resolution (Figure 3), with a Fourier shell correction of 0.143, calculated between two independent half maps. The resulting map at this resolution clearly resolves all secondary structures, in this case α helices. Each turn of the α helices is clearly visible in the density map (Figure 3C) and densities for some large amino acid side chains are also observed (Figure 3C). However, even at this resolution, it is still very challenging to derive an atomic structure, since density for all the residues is not resolved.

The CA tubular assembly is composed of CA hexamers arranged in a hexagonal lattice on a curved cylindrical surface (Figure 3A). Interestingly, the CA hexamer itself in this arrangement is also curved, and does not exhibit perfect six-fold symmetry, in contrast to the planar structure of the crystallographic wild type CA hexamer (PDB code 4XFX) [36]. As a result, rigid body docking of the crystal structure does not yield a satisfactory fit to the cryoEM density. MDFF was employed, using the density as a guide during the molecular dynamics simulation, resulting in a final density map with a cross-correlation of 0.8 (Figure 3B&C). Gratifyingly, MDFF places the large amino-acid side chains into the corresponding densities in the cryoEM map (Figure 3D). Note that the fit is still relatively poor in the CypA-binding loop region, since only weak density was observed given the intrinsic flexibility of this loop [64]. The above described MDFF-derived model was used as the starting model for the NMR CS-biased structure refinement, as detailed below.

MAS NMR of CA- assembly

We previously reported that the NL4-3 CA A92E tubular assembly yields MAS NMR spectra of remarkable resolution [40], which permitted the *de novo* assignments of backbone and side chain carbon and nitrogen atoms for 96% of residues (Table S1 of the Supporting Information). As illustrated in Figure 4, the spectral resolution for the WT NL4-3 CA tubular assembly is excellent, on par with that of the A92E mutant. Interestingly, we see a number of ^{13}C chemical shift and intensity differences between the spectra of the two variants (Figure 4A–F). Most of these changes are confined to the CypA loop region in which the amino acid change resides; however, several are detected outside the loop, also within the CTD. While these changes are small and do not exceed 1 ppm for ^{13}C nuclei and hence do not signalize a change in secondary structure, they do indicate distinct changes in the local structure, comparing wild type and the A92E mutant assemblies. This finding represents a further example for the presence of remarkable conformational plasticity in CA, which can adopt numerous slightly different conformations, thereby enabling formation of the varied curvature in the assembled mature capsids [4, 36, 38, 40, 65]. The fact that point mutations and/or changes in the experimental conditions for *in vitro* assembly affect local structure is not surprising, given the protein's plasticity. However, overall no changes in secondary structure are seen for the A92E and wild type tubular assemblies, and the only notable differences pertain to the CypA loop region.

Comparison of initial modeling with MDFF

To test whether in the assembled CA carbon side chain chemical shifts can be used to reliably predict local conformation, we examined the agreement between experimental ^{13}C sidechain shifts for Ile residues (reporting the different rotamers) and those calculated from the MDFF and the rotamer-optimized MDFF model using SHIFTX2. We selected

isoleucines for this test since a large number is present, fifteen in total, and they are distributed over the entire CA sequence, residing in rigid and flexible regions of the NTD and CTD. Shifts were calculated for each molecule in each hexamer, and the results are summarized in Figure 5. Figure 5A depicts the experimental $C\gamma1$, $C\gamma2$, and $C\delta1$ shifts for each residue, together with the predicted shifts from the MDFF models, color coded according to the rotameric state, and in Figure 5B the torsion angles dependence for $\chi1$ and $\chi2$ of the experimental and predicted shifts for each carbon type, $C\gamma1$, $C\gamma2$, and $C\delta1$ are plotted. For the majority of the isoleucines, with the exception of I15, I91, and I150, MDFF predicts that one predominant rotamer is present and, if a second rotamer is predicted, this is only seen in one or two of the twelve chains. Remarkably, for eleven residues (I2, I15, I73, I91, I115, I124, I129, I134, I135, I141), experimental shifts for at least two carbons are in agreement with the predictions for the predominant rotameric state. Interestingly, for the remaining four residues (I37, I104, I153, I201), the experimental shifts for one of the carbons agree with the prediction for one rotamer, while for the other two atoms the experimental shifts do not match any of the predicted ones.

These initial results are encouraging in several ways. First, the agreement between the experimental and predicted shifts for the majority of carbons of eleven of the fifteen residues demonstrates that experimental ^{13}C shifts could be used potentially be used for refinement, in particular for selecting rotameric states, as was also demonstrated previously on other systems [66]. More importantly, these findings will guide the next steps for improving our current work. The current MDFF models span a relatively limited set of torsion angles and, for a subset of residues, the poor agreement between the experimental and calculated shifts (outside of the error limit) clearly indicates that the MDFF models do not cover the experimentally observed rotamers. This needs to be addressed through a systematic analysis of conformational space, which is beyond the scope of this manuscript and will be pursued in a further study.

CS-biasing of backbone refinement

The MDFF-derived model of a single CA hexamer was refined using CS by MD simulations. CS restraints were applied only to the backbone heavy atoms. As expected, the backbone structure of well-defined helices in the cryoEM density only marginally changed by the CSMD simulation (< 0.5 Å). However, residues in flexible loops, particularly in the CypA-binding loop and the loop connecting helix 8 and 9 exhibited larger changes (Figure 6A). Importantly, the differences between predicted chemical shifts of the model and the experimental NMR shifts are relatively small (Figure 6B). In addition, the resulting trajectory reaches a relatively steady phase after 0.7 ns, indicating convergence of the simulation (Figure 6C). We should point out, however, that the use of time- and ensemble-averaged biasing forces should produce more reliable results since more than one conformer can result in a particular chemical shift [60, 67]. This sampling problem, in addition to the lack of a biasing-force for atoms in the amino acid's side-chains currently limits our approach and we will address this limitation in further work.

Conclusions

We determined a cryoEM density map of a tubular HIV-1 capsid assembly at 5 Å resolution and combined the electron density data with MAS NMR chemical shift data from the same capsid assembly. The cryoEM density map reveals the overall architecture of the capsid assembly, the secondary structures and positions of large amino acid side chains, while NMR chemical shifts are sensitive probes for the chemical environment for each amino acid. We thus combined the data from both cryoEM and MAS NMR for structure refinement, using molecular dynamics simulations. The derived models of CA-CTD and the CA hexameric assembly clearly demonstrate the feasibility and potential of this approach. At the present time, however, we only used chemical shifts from backbone carbons of the CA assembly, limiting our conclusions to some degree. Ongoing work will incorporate side chain chemical shifts as ensemble- and time-averaged restraints in the MD protocols, greatly improving the resulting structures through joint refinement algorithms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Bartesaghi A, Merk A, Banerjee S, Matthies D, Wu X, Milne JL, Subramaniam S. 2.2 Å Resolution cryo-EM Structure of Beta-Galactosidase in Complex with a Cell-Permeant Inhibitor. Science. 2015; 348:1147–1151. [PubMed: 25953817]

2. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 Ion Channel Determined by Electron Cryo-Microscopy. Nature. 2013; 504:107–112. [PubMed: 24305160]

3. Vinothkumar KR, Henderson R. Single Particle Electron Cryomicroscopy: Trends, Issues and Future Perspective. Q Rev Biophys. 2016; 49:e13. [PubMed: 27658821]

4. Byeon IJ, Meng X, Jung J, Zhao G, Yang R, Ahn J, Shi J, Concel J, Aiken C, Zhang P, et al. Structural Convergence Between Cryo-EM and NMR Reveals Intersubunit Interactions Critical for HIV-1 Capsid Function. Cell. 2009; 139:780–790. [PubMed: 19914170]

5. Merk A, Bartesaghi A, Banerjee S, Falconieri V, Rao P, Davis MI, Pragani R, Boxer MB, Earl LA, Milne JL, et al. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. Cell. 2016; 165:1698–1707. [PubMed: 27238019]

6. Zalk R, Clarke OB, des Georges A, Grassucci RA, Reiken S, Mancia F, Hendrickson WA, Frank J, Marks AR. Structure of a Mammalian Ryanodine Receptor. Nature. 2015; 517:44–49. [PubMed: 25470061]

7. Fan G, Baker ML, Wang Z, Baker MR, Sinyagovskiy PA, Chiu W, Ludtke SJ, Serysheva. Gating Machinery of InsP3R Channels Revealed by Electron Cryomicroscopy. Nature. 2015; 527:336–341. [PubMed: 26458101]

8. Yan Z, Bai XC, Yan C, Wu J, Li Z, Xie T, Peng W, Yin CC, Li X, Scheres SH, et al. Structure of the Rabbit Ryanodine Receptor RyR1 at near-Atomic Resolution. Nature. 2015; 517:50–55. [PubMed: 25517095]

9. Wang Z, Hryc CF, Bammes B, Afonine PV, Jakana J, Chen DH, Liu X, Baker ML, Kao C, Ludtke SJ, et al. An Atomic Model of Brome Mosaic Virus using Direct Electron Detection and Real-Space Optimization. Nat Commun. 2014; 5:4808. [PubMed: 25185801]

10. Baker ML, Hryc CF, Zhang Q, Wu W, Jakana J, Haase-Pettingell C, Afonine PV, Adams PD, King JA, Jiang W, et al. Validated Near-Atomic Resolution Structure of Bacteriophage Epsilon15 Derived from cryo-EM and Modeling. Proc Natl Acad Sci U S A. 2013; 110:12301–12306. [PubMed: 23840063]

11. Brown A, Long F, Nicholls RA, Toots J, Emsley P, Murshudov G. Tools for Macromolecular Model Building and Refinement Into Electron Cryo-Microscopy Reconstructions. Acta Crystallogr, Sect D: Biol Crystallogr. 2015; 71:136–153. [PubMed: 25615868]

12. Goh BC, Hadden JA, Bernardi RC, Singharoy A, McGreevy R, Rudack T, Cassidy CK, Schulten K. Computational Methodologies for Real-Space Structural Refinement of Large Macromolecular Complexes. Annu Rev Biophys. 2016; 45:253–278. [PubMed: 27145875]

13. Trabuco LG, Villa E, Mitra K, Frank J, Schulten K. Flexible Fitting of Atomic Structures into Electron Microscopy Maps using Molecular Dynamics. Structure. 2008; 16:673–683. [PubMed: 18462672]

14. DiMaio F, Song Y, Li X, Brunner MJ, Xu C, Conticello V, Egelman E, Marlovits TC, Cheng Y, Baker D. Atomic-Accuracy Models From 4.5-Å Cryo-Electron Microscopy Data with Density-Guided Iterative Local Refinement. Nat Methods. 2015; 12:361–365. [PubMed: 25707030]

15. Lindert S, McCammon JA. Improved cryoEM-Guided Iterative Molecular Dynamics-Rosetta Protein Structure Refinement Protocol for High Precision Protein Structure Prediction. J Chem Theory Comput. 2015; 11:1337–1346. [PubMed: 25883538]

16. Franks WT, Kloepper KD, Wylie BJ, Rienstra CM. Four-Dimensional Heteronuclear Correlation Experiments for Chemical Shift Assignment of Solid Proteins. J Biomol NMR. 2007; 39:107–131. [PubMed: 17687624]

17. Huber M, Hiller S, Schanda P, Ernst M, Bockmann A, Verel R, Meier BH. A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination. Chemphyschem. 2011; 12:915–918. [PubMed: 21442705]

18. Linser R, Bardiaux B, Andreas LB, Hyberts SG, Morris VK, Pintacuda G, Sunde M, Kwan AH, Wagner G. Solid-State NMR Structure Determination From Diagonal-Compensated, Sparsely Nonuniform-Sampled 4D Proton-Proton Restraints. J Am Chem Soc. 2014; 136:11002–11010. [PubMed: 24988008]

19. Huber M, Bockmann A, Hiller S, Meier BH. 4D Solid-State NMR for Protein Structure Determination. Phys Chem Chem Phys. 2012; 14:5239–5246. [PubMed: 22402636]

20. Andreas LB, Jaudzems K, Stanek J, Lalli D, Bertarello A, Le Marchand T, Cala-De Paepe D, Kotelovica S, Akopjana I, Knott B, et al. Structure of Fully Protonated Proteins by Proton-Detected Magic-Angle Spinning NMR. Proc Natl Acad Sci U S A. 2016; 113:9187–9192. [PubMed: 27489348]

21. Sborgi L, Ravotti F, Dandey VP, Dick MS, Mazur A, Reckel S, Chami M, Scherer S, Huber M, Bockmann A, et al. Structure and Assembly of the Mouse ASC Inflammasome by Combined NMR Spectroscopy and Cryo-Electron Microscopy. Proc Natl Acad Sci U S A. 2015; 112:13237–13242. [PubMed: 26464513]

22. Gong Z, Schwieters CD, Tang C. Conjoined use of EM and NMR in RNA Structure Refinement. PLoS One. 2015; 10:e0120445. [PubMed: 25798848]

23. Demers JP, Habenstein B, Loquet A, Kumar Vasa S, Giller K, Becker S, Baker D, Lange A, Sgourakis NG. High-Resolution Structure of the Shigella Type-III Secretion Needle by Solid-State NMR and cryo-Electron Microscopy. Nat Commun. 2014; 5:4976. [PubMed: 25264107]

24. Loquet A, Sgourakis NG, Gupta R, Giller K, Riedel D, Goosmann C, Griesinger C, Kolbe M, Baker D, Becker S, et al. Atomic Model of the Type III Secretion System Needle. Nature. 2012; 486:276–279. [PubMed: 22699623]

25. Rohl, CA., Strauss, CEM., Misura, KMS., Baker, D. Methods in Enzymology. Vol. 383. Academic Press; 2004. Protein Structure Prediction Using Rosetta; p. 66-93.

26. Zhang J, Wang Q, Barz B, He Z, Kosztin I, Shang Y, Xu D. MUFOLD: A New Solution for Protein 3D Structure Prediction. Proteins: Struct, Funct Bioinf. 2010; 78:1137–1152.

27. Mori T, Miyashita N, Im W, Feig M, Sugita Y. Molecular Dynamics Simulations of Biological Membranes and Membrane Proteins using Enhanced Conformational Sampling Algorithms. Biochim Biophys Acta. 2016; 1858:1635–1651. [PubMed: 26766517]

28. Chong LT, Saglam AS, Zuckerman DM. Path-Sampling Strategies for Simulating Rare Events in Biomolecular Systems. Curr Opin Struct Biol. 2016; 43:88–94. [PubMed: 27984811]

29. Perilla JR, Goh BC, Cassidy CK, Liu B, Bernardi RC, Rudack T, Yu H, Wu Z, Schulten K. Molecular Dynamics Simulations of Large Macromolecular Complexes. Curr Opin Struct Biol. 2015; 31:64–74. [PubMed: 25845770]

30. Cassidy CK, Himes BA, Alvarez FJ, Ma J, Zhao G, Perilla JR, Schulten K, Zhang P. CryoEM and Computer Simulations Reveal a Novel Kinase Conformational Switch in Bacterial Chemotaxis Signaling. Elife. 2015; 4:e08419. [PubMed: 26583751]

31. Schweitzer A, Aufderheide A, Rudack T, Beck F, Pfeifer G, Plitzko JM, Sakata E, Schulten K, Förster F, Baumeister W. Structure of the Human 26S Proteasome at a Resolution of 3.9 Å. Proc Natl Acad Sci U S A. 2016; 113:7816–7821. [PubMed: 27342858]

32. Villa E, Sengupta J, Trabuco LG, LeBarron J, Baxter WT, Shaikh TR, Grassucci RA, Nissen P, Ehrenberg M, Schulten K, et al. Ribosome-Induced Changes in Elongation Factor Tu Conformation Control GTP Hydrolysis. Proc Natl Acad Sci U S A. 2009; 106:1063–1068. [PubMed: 19122150]

33. Wang X, Xu F, Liu J, Gao B, Liu Y, Zhai Y, Ma J, Zhang K, Baker TS, Schulten K, et al. Atomic Model of Rabbit Hemorrhagic Disease Virus by Cryo-Electron Microscopy and Crystallography. PLoS Pathog. 2013; 9:e1003132. [PubMed: 23341770]

34. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, et al. Mature HIV-1 Capsid Structure by Cryo-Electron Microscopy and All-Atom Molecular Dynamics. Nature. 2013; 497:643–646. [PubMed: 23719463]

35. Yu I, Mori T, Ando T, Harada R, Jung J, Sugita Y, Feig M. Biomolecular Interactions Modulate Macromolecular Structure and Dynamics in Atomistic Model of a Bacterial Cytoplasm. eLife. 2016; 5:e19274. [PubMed: 27801646]

36. Gres AT, Kirby KA, KewalRamani VN, Tanner JJ, Pornillos O, Sarafianos SG. X-ray Crystal Structures of Native HIV-1 Capsid Protein Reveal Conformational Variability. Science. 2015; 349:99–103. [PubMed: 26044298]

37. Pornillos O, Ganser-Pornillos BK, Yeager M. Atomic-Level Modelling of the HIV Capsid. Nature. 2011; 469:424–427. [PubMed: 21248851]

38. Pornillos O, Ganser-Pornillos BK, Kelly BN, Hua Y, Whitby FG, Stout CD, Sundquist WI, Hill CP, Yeager M. X-ray Structures of the Hexameric Building Block of the HIV Capsid. Cell. 2009; 137:1282–1292. [PubMed: 19523676]

39. Perilla JR, Gronenborn AM. Molecular Architecture of the Retroviral Capsid. Trends Biochem Sci. 2016; 41:410–420. [PubMed: 27039020]

40. Han Y, Hou G, Suiter CL, Ahn J, Byeon IJ, Lipton AS, Burton S, Hung I, Gor'kov PL, Gan Z, et al. Magic Angle Spinning NMR Reveals Sequence-Dependent Structural Plasticity, Dynamics, and the Spacer Peptide 1 Conformation in HIV-1 Capsid Protein Assemblies. J Am Chem Soc. 2013; 135:17793–17803. [PubMed: 24164646]

41. Yang R, Aiken C. A Mutation in alpha Helix 3 of CA Renders Human Immunodeficiency Virus Type 1 Cyclosporin A Resistant and Dependent: Rescue by a Second-Site Substitution in a Distal Region of CA. J Virol. 2007; 81:3749–3756. [PubMed: 17267487]

42. Ludtke SJ, Baldwin PR, Chiu W. EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. J Struct Biol. 1999; 128:82–97. [PubMed: 10600563]

43. Scheres SH. RELION: Implementation of a Bayesian Approach to Cryo-EM Structure Determination. J Struct Biol. 2012; 180:519–530. [PubMed: 23000701]

44. Grigorieff N. FREALIGN: High-Resolution Refinement of Single Particle Structures. J Struct Biol. 2007; 157:117–125. [PubMed: 16828314]

45. Bräuniger T, Wormald P, Hodgkinson P. Improved Proton Decoupling in NMR Spectroscopy of Crystalline Solids Using the SPINAL-64 Sequence. Monatsh Chem. 2002; 133:1549–1554.

46. Marion D, Ikura M, Tschudin R, Bax A. Rapid Recording of 2D NMR Spectra without Phase Cycling. Application to the Study of Hydrogen Exchange in Proteins. J Magn Reson. 1989; 85:393–399.

47. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an Automated Pipeline for the Setup of Poisson–Boltzmann Electrostatics Calculations. Nucleic Acids Res. 2004; 32:W665–W667. [PubMed: 15215472]

48. Stone JE, Phillips JC, Freddolino PL, Hardy DJ, Trabuco LG, Schulten K. Accelerating Molecular Modeling Applications with Graphics Processors. J Comput Chem. 2007; 28:2618–2640. [PubMed: 17894371]

49. Humphrey W, Dalke A, Schulten K. VMD: Visual Molecular Dynamics. J Mol Graphics. 1996; 14:33–38.

50. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. Scalable Molecular Dynamics with NAMD. J Comput Chem. 2005; 26:1781–1802. [PubMed: 16222654]

51. Best RB, Zhu X, Shim J, Lopes PEM, Mittal J, Feig M, MacKerell AD. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain $\chi 1$ and $\chi 2$ Dihedral Angles. J Chem Theory Comput. 2012; 8:3257–3273. [PubMed: 23341755]

52. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983; 79:926–935.

53. Han B, Liu Y, Ginzinger SW, Wishart DS. SHIFTX2: Significantly Improved Protein Chemical Shift Prediction. J Biomol NMR. 2011; 50:43–57. [PubMed: 21448735]

54. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. J Phys Chem B. 1998; 102:3586–3616. [PubMed: 24889800]

55. Mackerell AD, Feig M, Brooks CL. Extending the Treatment of Backbone Energetics in Protein Force Fields: Limitations of Gas-Phase Quantum Mechanics in Reproducing Protein Conformational Distributions in Molecular Dynamics Simulations. J Comput Chem. 2004; 25:1400–1415. [PubMed: 15185334]

56. Tanner DE, Chan KY, Phillips JC, Schulten K. Parallel Generalized Born Implicit Solvent Calculations with NAMD. J Chem Theory Comput. 2011; 7:3635–3642. [PubMed: 22121340]

57. Fu B, Sahakyan AB, Camilloni C, Tartaglia GG, Paci E, Caflisch A, Vendruscolo M, Cavalli A. ALMOST: An All Atom Molecular Simulation Toolkit for Protein Structure Determination. J Comput Chem. 2014; 35:1101–1105. [PubMed: 24676684]

58. Tribello GA, Bonomi M, Branduardi D, Camilloni C, Bussi G. PLUMED 2: New Feathers for an Old Bird. Comput Phys Commun. 2014; 185:604–613.

59. Robustelli P, Kohlhoff K, Cavalli A, Vendruscolo M. Using NMR Chemical Shifts as Structural Restraints in Molecular Dynamics Simulations of Proteins. Structure. 18:923–933.

60. Camilloni C, Cavalli A, Vendruscolo M. Assessment of the Use of NMR Chemical Shifts as Replica-Averaged Structural Restraints in Molecular Dynamics Simulations to Characterize the Dynamics of Proteins. J Phys Chem B. 2013; 117:1838–1843. [PubMed: 23327201]

61. Worthylake DK, Wang H, Yoo S, Sundquist WI, Hill CP. Structures of the HIV-1 Capsid Protein Dimerization Domain at 2.6 Å Resolution. Acta Crystallogr, Sect D: Biol Crystallogr. 1999; 55:85–92. [PubMed: 10089398]

62. Zhang P, Meng X, Zhao G. Tubular Crystals and Helical Arrays: Structural Determination of HIV-1 Capsid Assemblies Using Iterative Helical Real-Space Reconstruction. Methods Mol Biol. 2013; 955:381–399. [PubMed: 23132072]

63. Meng X, Zhao G, Zhang P. Structure of HIV-1 Capsid Assemblies by Cryo-Electron Microscopy and Iterative Helical Real-Space Reconstruction. J Vis Exp. 2011:3041. [PubMed: 21860371]

64. Liu C, Perilla JR, Ning J, Lu M, Hou G, Ramalho R, Himes BA, Zhao G, Bedwell GJ, Byeon IJ, et al. Cyclophilin A Stabilizes the HIV-1 Capsid Through a Novel non-Canonical Binding Site. Nat Commun. 2016; 7:10714. [PubMed: 26940118]

65. Mattei S, Glass B, Hagen WJH, Kräusslich HG, Briggs JAG. The Structure and Flexibility of Conical HIV-1 Capsids Determined within Intact Virions. Science. 2016; 354:1434–1437. [PubMed: 27980210]

66. Hong M, Mishanina TV, Cady SD. Accurate Measurement of Methyl 13C Chemical Shifts by Solid-State NMR for the Determination of Protein Side Chain Conformation: the Influenza a M2 Transmembrane Peptide as an Example. J Am Chem Soc. 2009; 131:7806–7816. [PubMed: 19441789]

67. Li DW, Brüschweiler R. PPM: a Side-Chain and Backbone Chemical Shift Predictor for the Assessment of Protein Conformational Ensembles. J Biomol NMR. 2012; 54:257–265. [PubMed: 22972619]

68. Shen Y, Bax A. Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts using Artificial Neural Networks. J Biomol NMR. 2013; 56:227–241. [PubMed: 23728592]
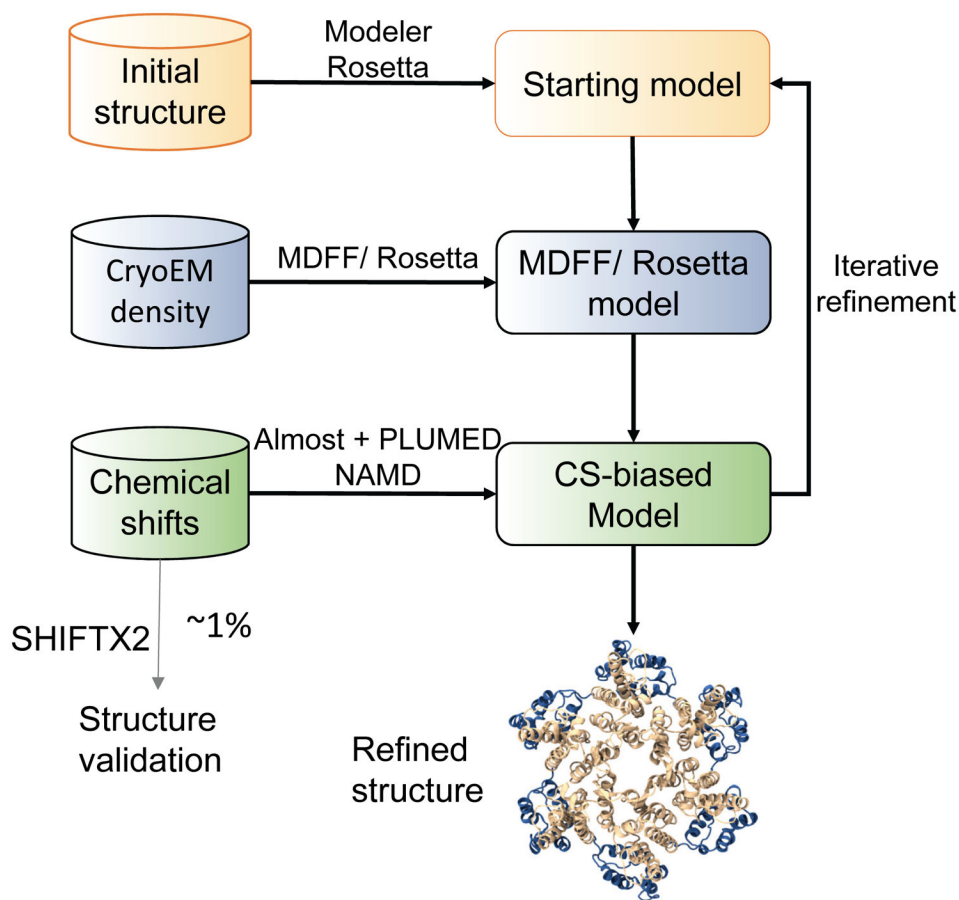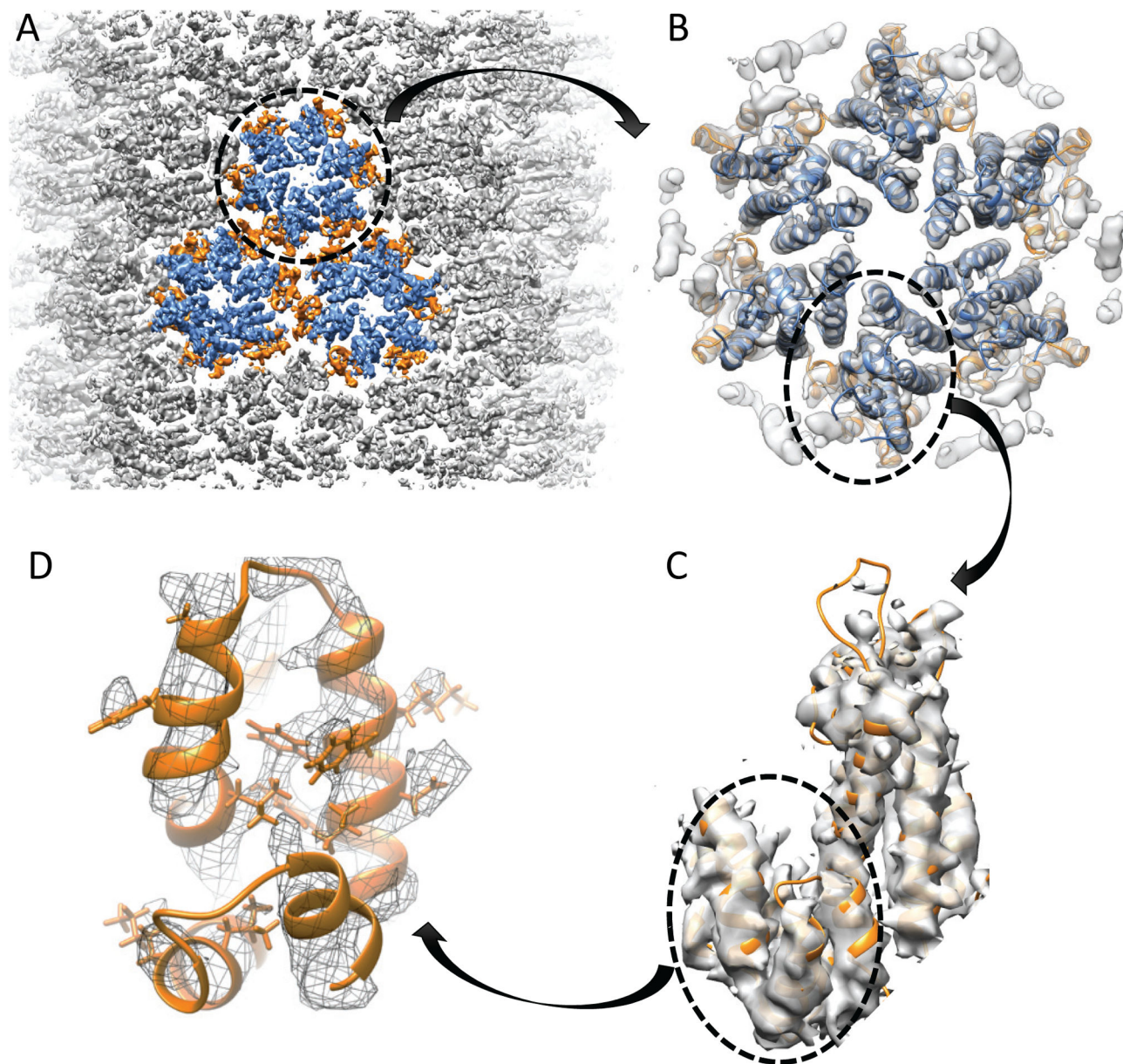
Figure 1.
Schematic diagram illustrating NMR CS-biased cryoEM structure refinement. A starting model is fit into the experimental cryoEM density map (4–6 Å resolution) using molecular dynamics flexible fitting (MDFF). The predicted chemical shifts (CSs) for the resulting model are compared with experimental NMR CSs, and the differences between these are incorporated as a biasing potential in MD, such as NAMD. The biasing forces are calculated using Almost and send to NAMD using PLUMED. The resulting CS-biased model is then iteratively refined. In parallel, CSs calculated from the resulting model using SHIFTX2 are compared to a small subset of experimental NMR CSs that have not been used during the CS-biased model refinement procedure for model validation.
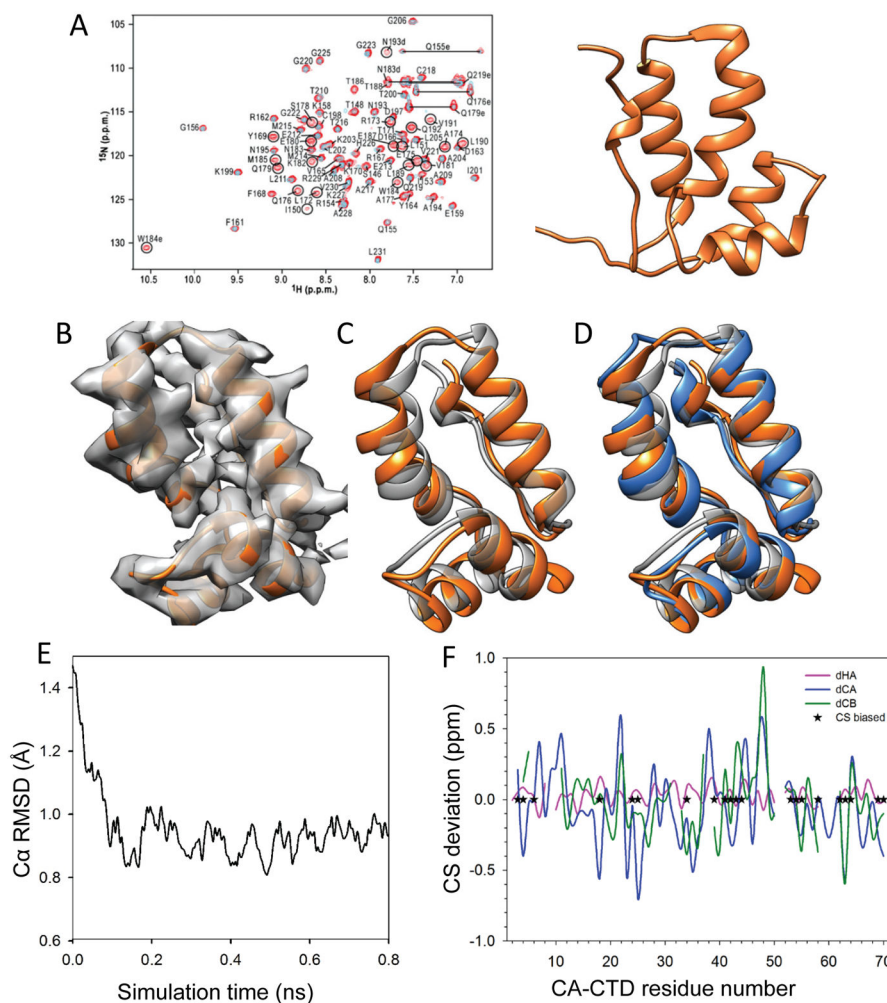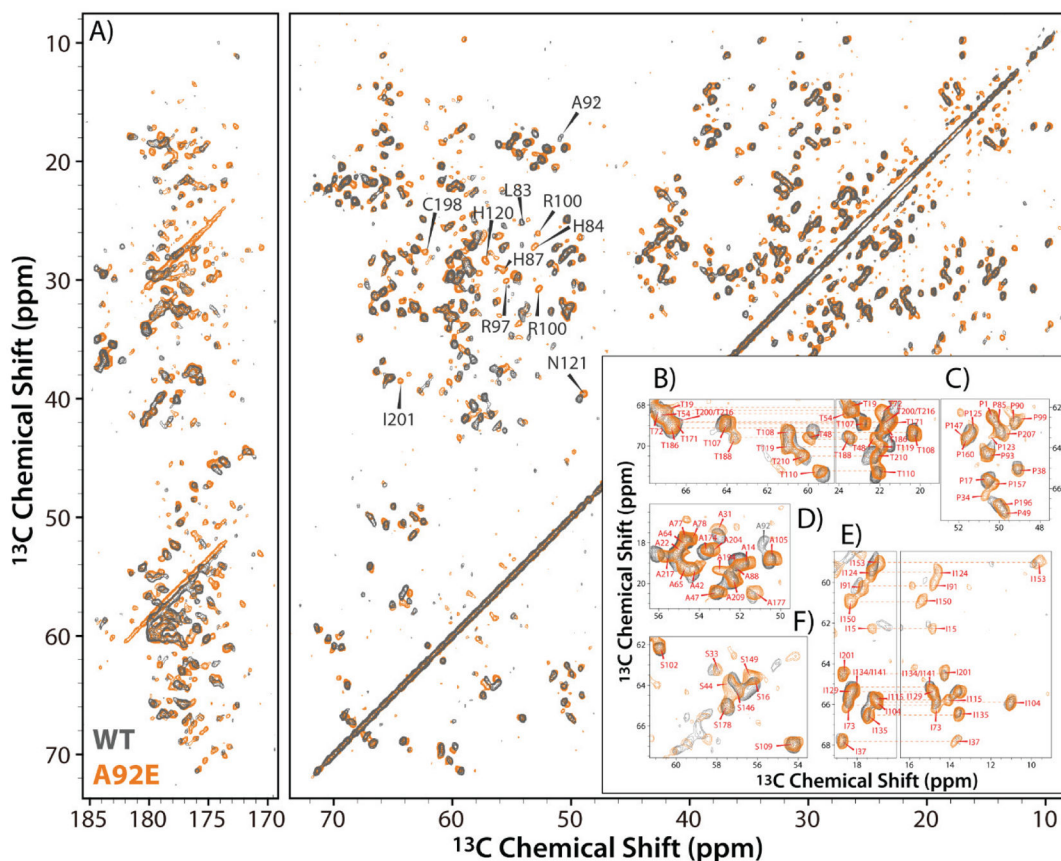
Figure 2.
CS-biased structure refinement using the CA-CTD as a model system. A) Solution NMR spectrum (left) of the CA-CTD domain and the solution NMR structure (PDB: 2kod, right) [4]. B) Simulated electron density map at 5 Å resolution based on the NMR structure, overlaid onto the NMR atomic model (PDB: 2kod, orange). C) MDFF fitting of the X-ray crystal structure of CA-CTD (PDB:1A43) into the simulated density map (grey model), superimposed onto the NMR structure (orange). D) Superposition of the CS-biased refined structure (blue), the MDFF model (grey) and the ground truth NMR structure (orange). E) Cα RMSD between the CS-biased refined model and the ground truth along the simulation trajectory. F) Differences between the predicted CSs of the refined model and the experimental solution chemical shifts: Cα; blue, Cβ; green and H; pink. Residues included in the biasing calculation are labeled with "*". (Panel A is from Byeon et al. [4], with permission).

Figure 3.
CryoEM density map of the CA tubular assembly at 5 Å resolution, reconstructed from CA tubes with a helical symmetry of (−12, 11). A) Surface rendering of the density map of the CA tube, countered at 1.5σ. Three connecting CA hexamers are colored: CA-NTD; blue, CA-CTD; orange. B) A single CA hexamer is cut out from the lattice array, and MDFF fit using the crystal structure (PDB: 4XFX). C) Surface rendering of the CA monomer, cut out from the CA hexamer and countered at 1.2σ, displaying the helical turns for the alpha helices and some densities for the side chains. D) Surface mesh rendering of the CA-CTD, countered at 1.2σ, displaying the side chain densities of some bulky residues.
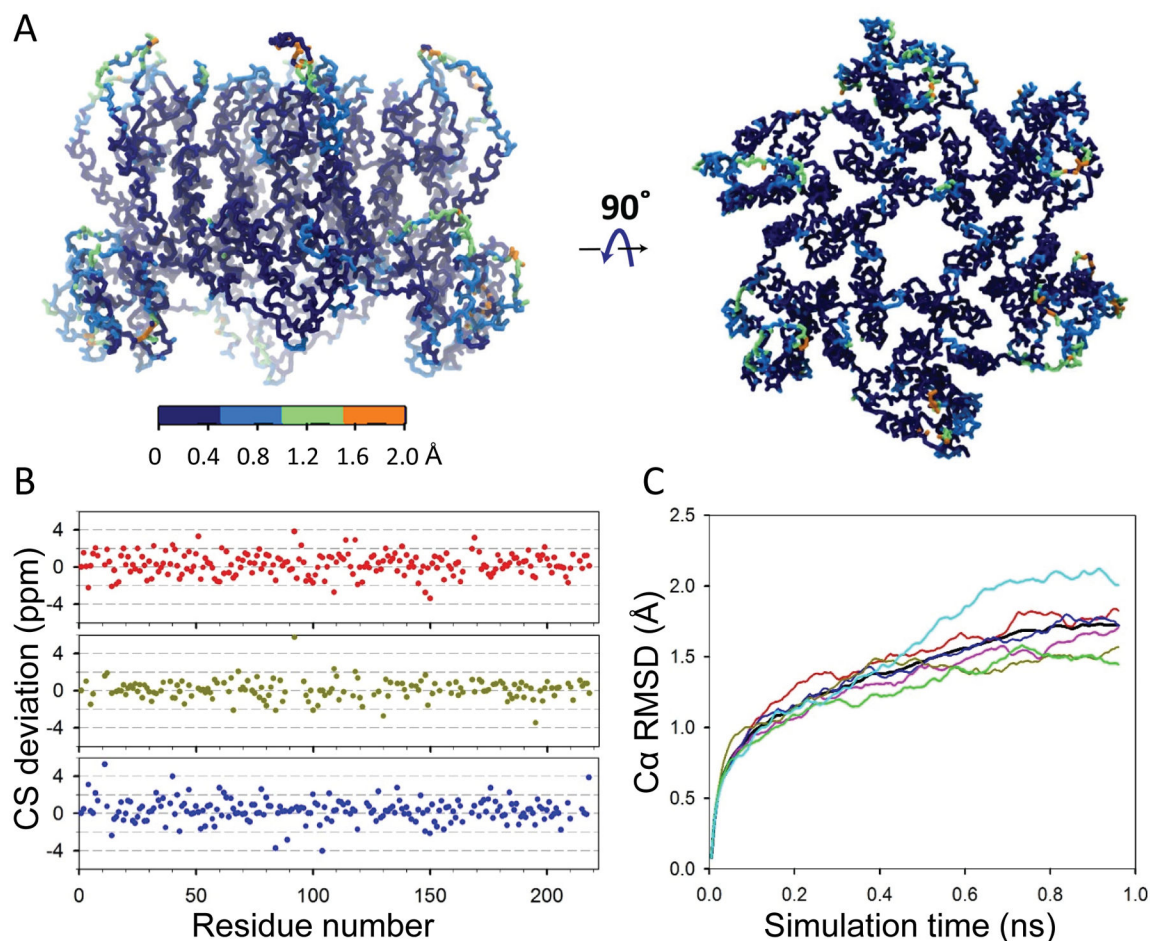
Figure 4.
A) 2D ^{13}C-^{13}C DARR spectrum of tubular assemblies of U-^{13}C,^{15}N-CA A92E (orange) superimposed on the 2D ^{13}C-^{13}C CORD spectrum of U-^{13}C,^{15}N-CA WT (gray), acquired at 19.9 T (850 MHz), using a mixing time of 50 ms. B)–F) Expansions of the DARR spectra for the aliphatic region, comprising Ala, Ile, Ser, and Thr resonances. Note, the excellent spectral resolution: B) C^α-C^β (left) and C^β-$C^{\gamma2}$ (right) resonances of Thr residues; C) C^α-C^δ resonances of Pro residues; D) C^α-C^β resonances of Ala residues; D) C^α-$C^{\gamma2}$ (left) and C^α-$C^{\delta1}$ (right) resonances of Ile residues; E) C^α-C^β resonances of Ser residues.

Figure 5.
Top: Experimental ^{13}C chemical shifts of sidechain carbons for individual Ile residues in tubular assemblies of CA A92E: Cγ1 (left), Cγ2 (middle), and Cδ1 (right) and those predicted using TALOS-N[68]. Bottom panels: ^{13}C chemical shifts for Ile side chain carbon atoms plotted vs. χ1 and χ2 torsion angles. Shifts predicted by SHIFTX2 for each CA monomer in the MDFF models are color coded by rotamer type: magenta (g+t), yellow (g−g−), grey (g−t), and blue (tt). For for the plots of shift vs. χ2, the experimental shifts are depicted next to each panel.

Figure 6.
MAS NMR CS-biased (backbone only) model of a CA hexamer in the CA tubular assembly. A) Root-mean-square fluctuations (RMSFs) resulting from the CS-biased simulation. Projection of the RMSFs on the structure of CA hexamer, ranging from 0 (dark blue) to orange (2.0 Å). B) Differences between the predicted CSs of the refined model and the experimental MAS NMR shifts for all CA residues, Cα in red, Cβ in dark green and C in blue. C) Time-evolution of the Cα RMSD between the CS-biased refined model and the starting MDFF model. All 6 CA chains are plotted in color, with the CA hexamer in black.