Structure, Volume *25*

Supplemental Information

Protein Structure Determination by Assembling Super-Secondary Structure Motifs Using Pseudocontact Shifts

Kala Bharath Pilla, Gottfried Otting, and Thomas Huber

Supplementary information

Protein Structure Determination by Assembling Super-Secondary Structure Motifs Using Pseudocontact Shifts

Kala Bharath Pilla, Gottfried Otting and Thomas Huber

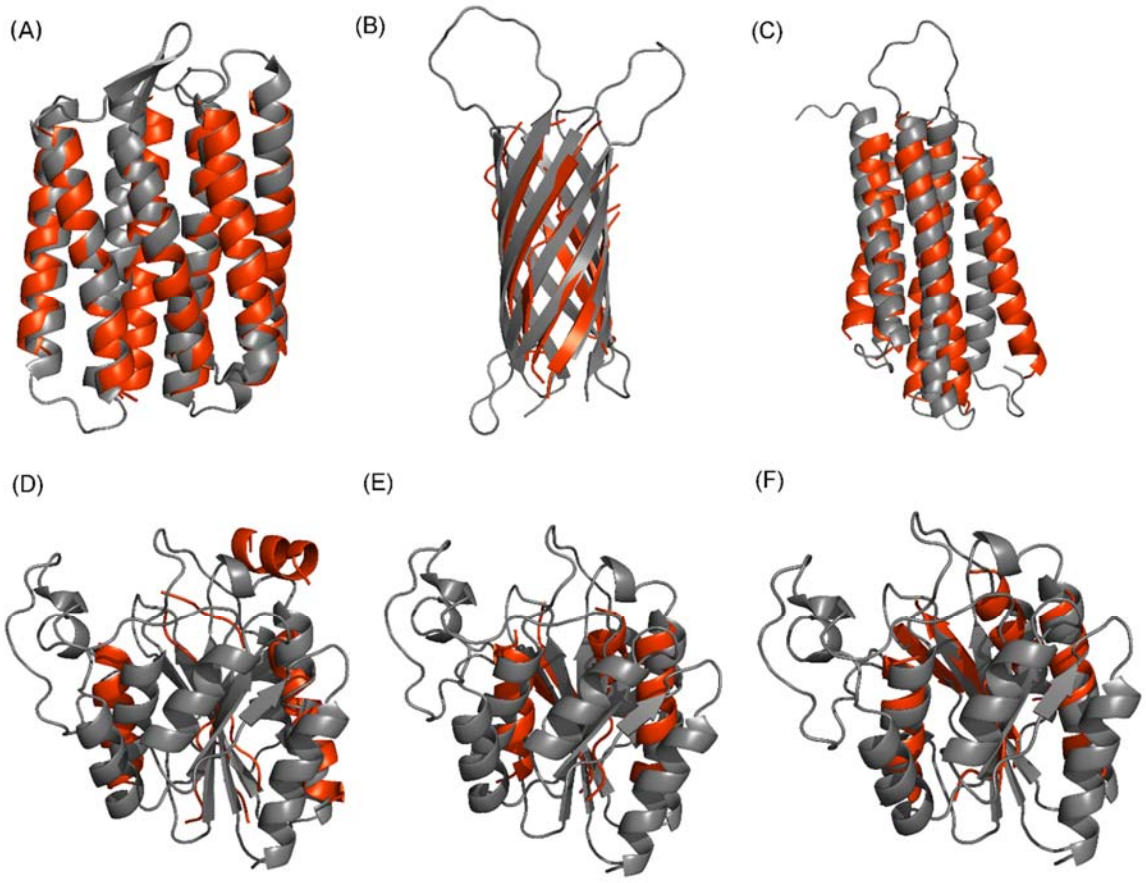

Figure S1, related to Figure 2: Superposition representation of the backbones of the best PCS-fitted Smotif assemblies calculated with DINGO-PCS (red) onto corresponding reference structures (gray). PCS data are reduced to three metal centers and two metals per center. (A) Target-A, 2.4 Å RMSD to the reference crystal structure [PDID 1H68]. (B) Target-C, 2.0 Å RMSD to the reference NMR structure [PDBID: 2M06]. (C) Target-G, 4.1 Å RMSD to the reference NMR structure [PDBID: 2JSW]. (D) Target-H, 8.1 Å RMSD to the reference NMR structure [PDBID: 2IUE]. (E) Target-H, 3.5 Å RMSD to the reference NMR structure [PDBID: 2IUE] but using PCS data reduced to three metal centers and four metals per center. (F) Target-H, 3.2 Å RMSD to the reference NMR structure [PDBID: 2IUE] but using PCS data reduced to four metal centers and two metals per center. See also Table S1.

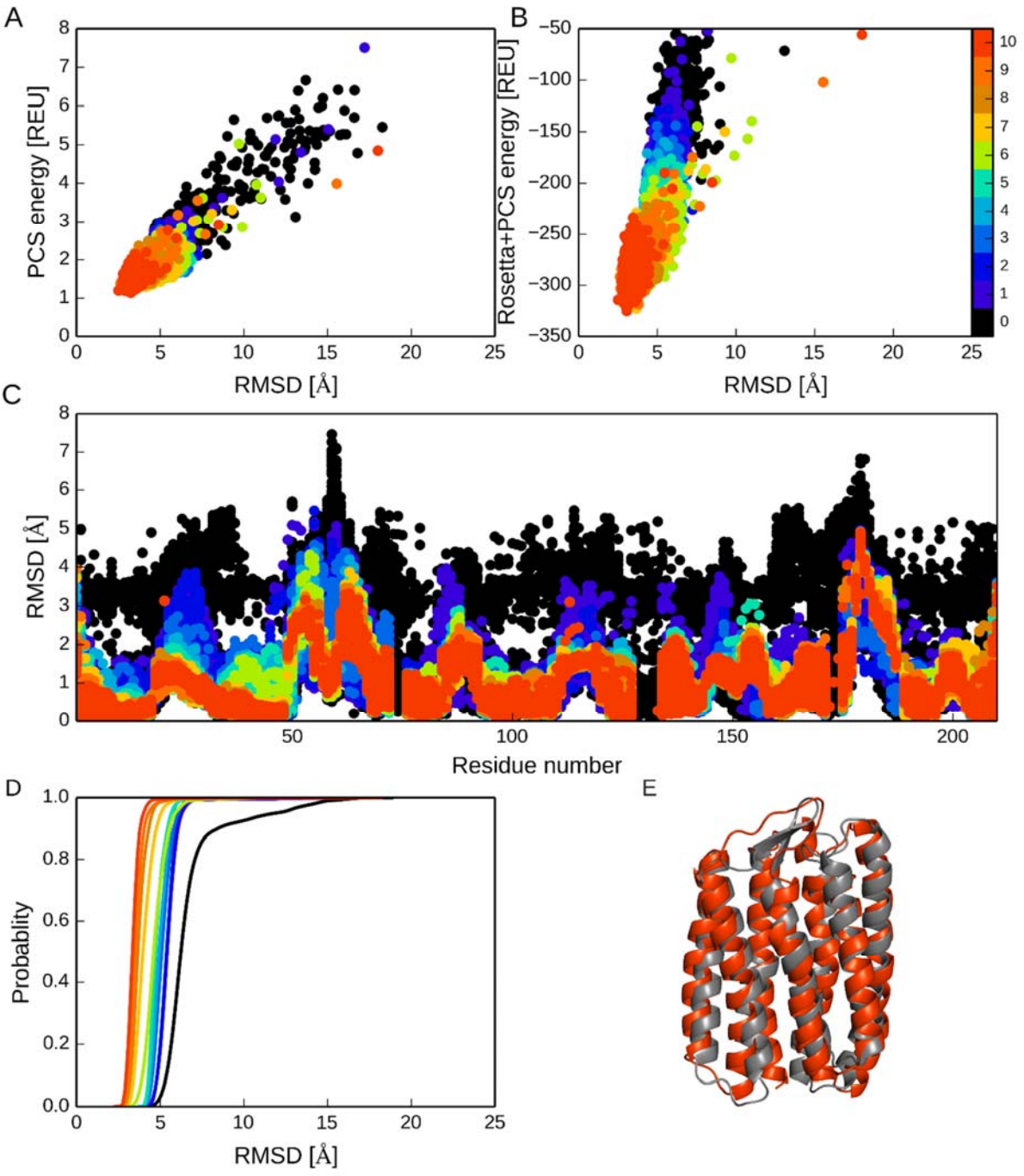

Figure S2, related to Figure 3: Results from PCS-driven iterative GPS-Rosetta applied to target-A* (pSRII). (A) Scatter plot of structures sampled by GPS-Rosetta. The PCS energy is plotted versus the Cα RMSD of the crystal structure [PDB ID: 1H68 (Royant et al., 2001)]. The results from the different iterations are color-coded, with the zeroth iteration in black and the next ten iterations in blue to red as shown in the color bar on the right. (B) Same as (A), but plotted against combined all-atom Rosetta energy and PCS energy. (C) Improvement in the quality of fragments

identified by overlapping Δχ tensors in the PCS-driven iterative scheme. The plot shows the RMSD calculated between each nine-residue fragment and its corresponding native fragment in the crystal structure. The zeroth iteration (black) used the Smotif-enhanced fragment library, while in subsequent iterations fragment libraries were recomputed from sampled structures considering the PCSs. (D) Probability density plots illustrating how consecutive iterations shift the conformational sampling towards structures with lower Cα RMSD to the crystal structure. (E) Superimposition of the structure with the lowest PCS energy (red) with the crystal structure (gray).

Table S1 related to Figure 2 and Figure S1

Performance comparison of the DINGO-PCS algorithm with reduced PCS-datasets

Target	Tags/(metals per tag) = total datasets	Total number of PCSs[a]	Cα RMSD[b]	Total Smotifs /Assembled Smotifs	PDB ID
A (pSRII)	3 tags/4 metals + 1 tag/3metals = 15 datasets	737	1.9 Å	6/6	1H68
	3 tags/2 metals = 6 datasets	303	2.4 Å	6/6	
C (OmpX)	4 tags/4 metals = 16 datasets	1421	2.1 Å	7/7	2M06
	3 tags/2 metals = 6 datasets	527	2.0 Å	7/7	
G (Talin, C-terminal actin binding site)	4 tags/4 metals = 16 datasets	1809	4.5 Å	4/4	2JSW
	3 tags/2 metals = 6 datasets	707	4.1 Å	4/4	
H (Pactolus domain-1)	4 tags/4 metals = 16 datasets	1961	5.1 Å	10/10	2IUE
	3 tags/2 metals = 6 datasets	773	8.1 Å	6/10	
	3 tags/4 metals = 12 datasets	1449	3.5 Å	6/10	
	4 tags/2 metals = 8 datasets	1051	3.2 Å	6/10	

[a] Total number of PCSs calculated for all amino acid residues (Smotifs + loops) in the target.

[b] The C^{α} RMSD was calculated between the best Smotif assembly calculated by DINGO-PCS, which was identified as the structure best fulfilling the PCS data, and the residues covered by Smotifs in the corresponding reference structure.