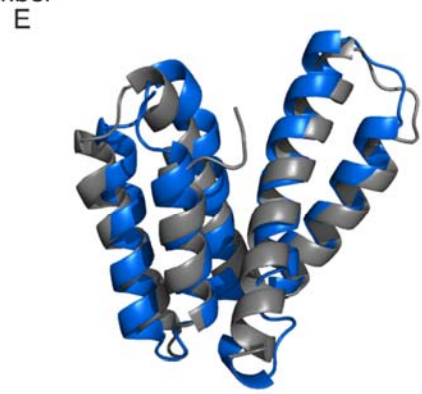
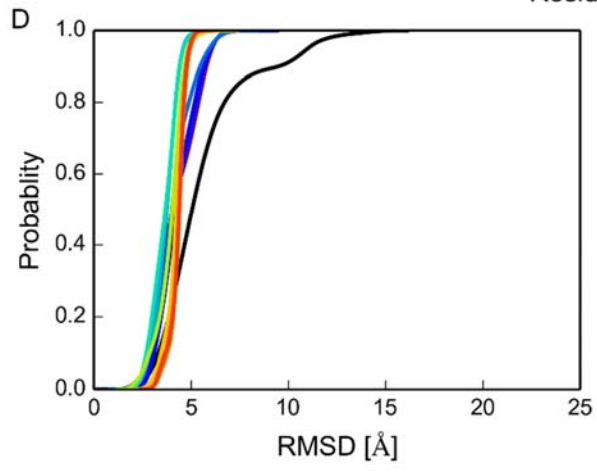
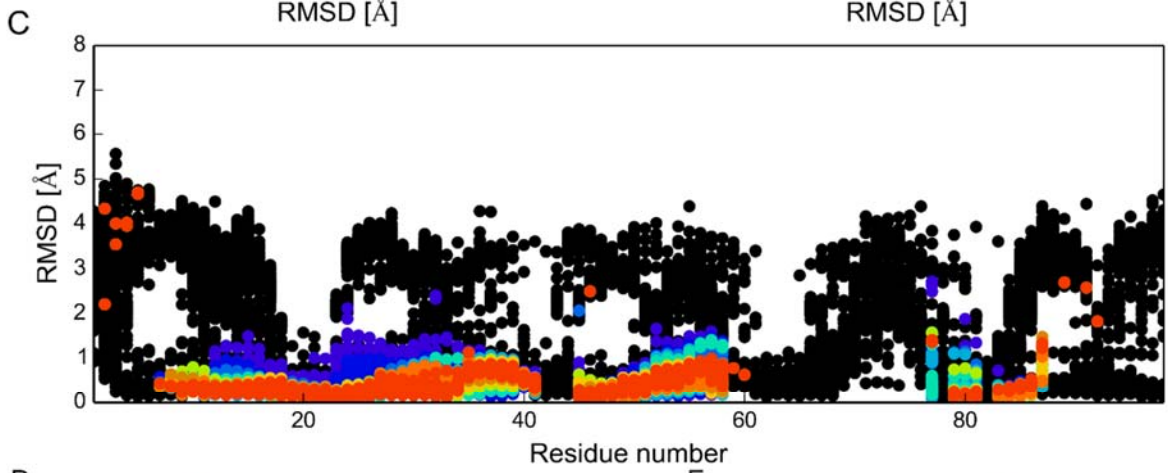
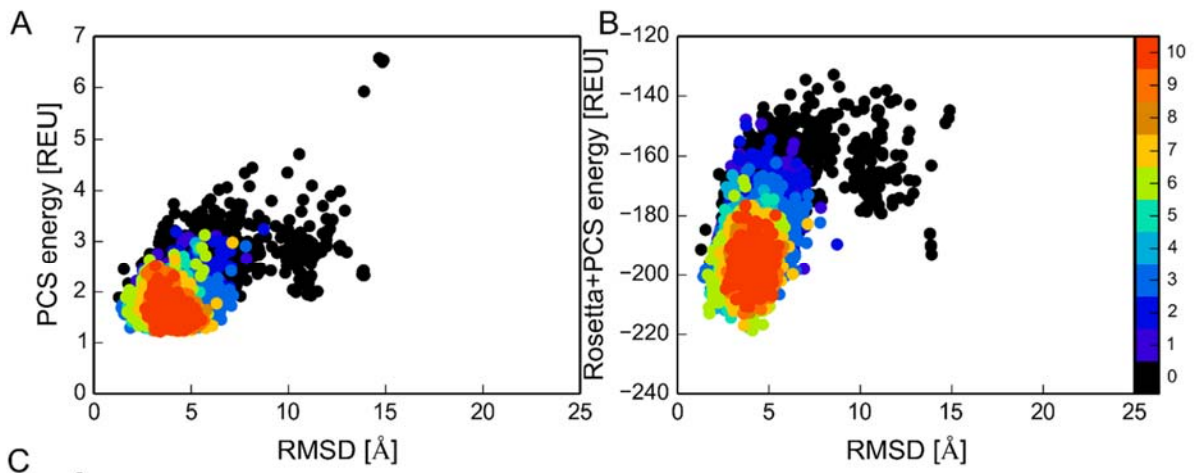


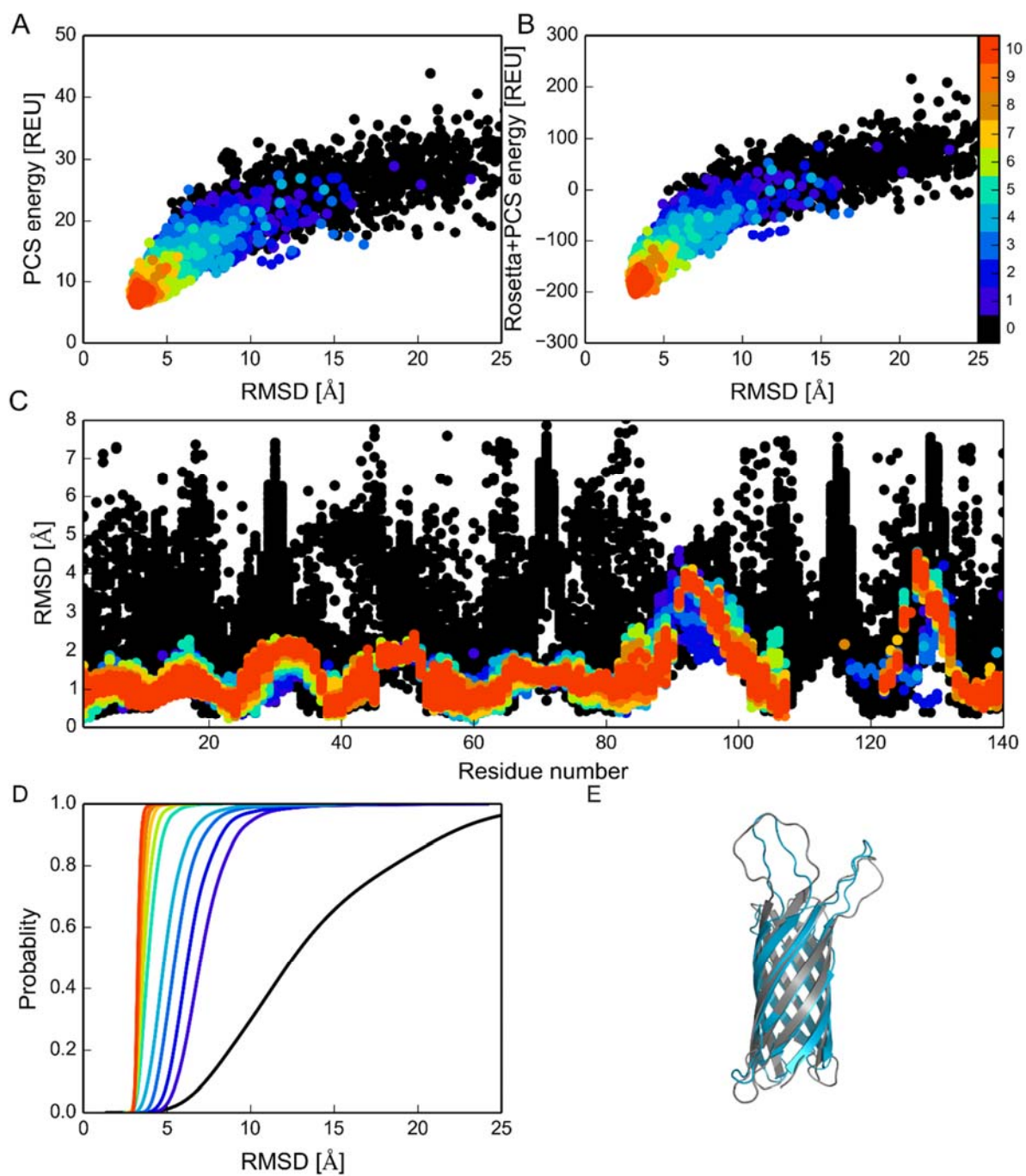
**Supplementary information**

Pseudocontact Shift-Driven Iterative Resampling for 3D Structure  
Determinations of Large Proteins

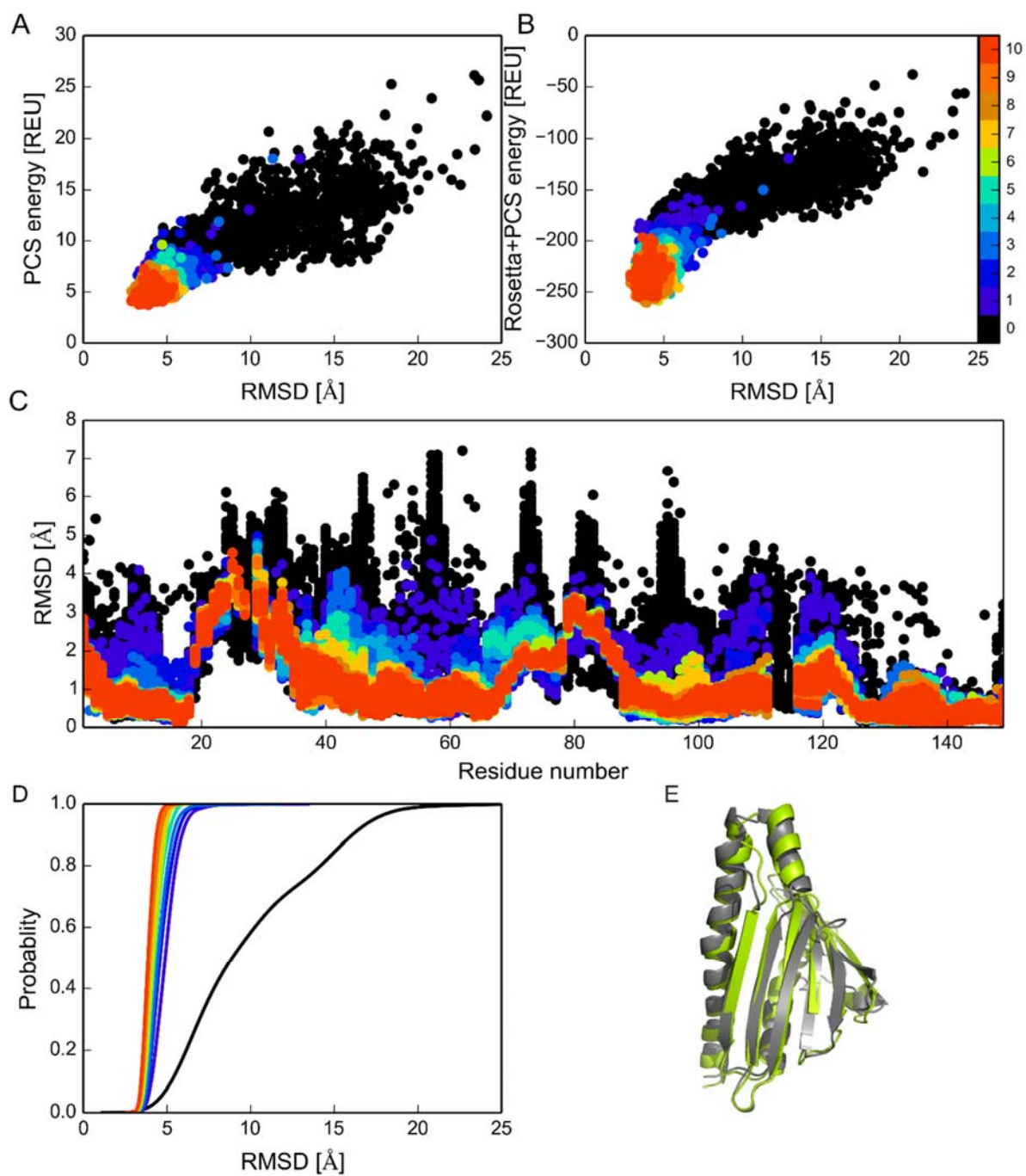
Kala Bharath Pilla, Gottfried Otting and Thomas Huber



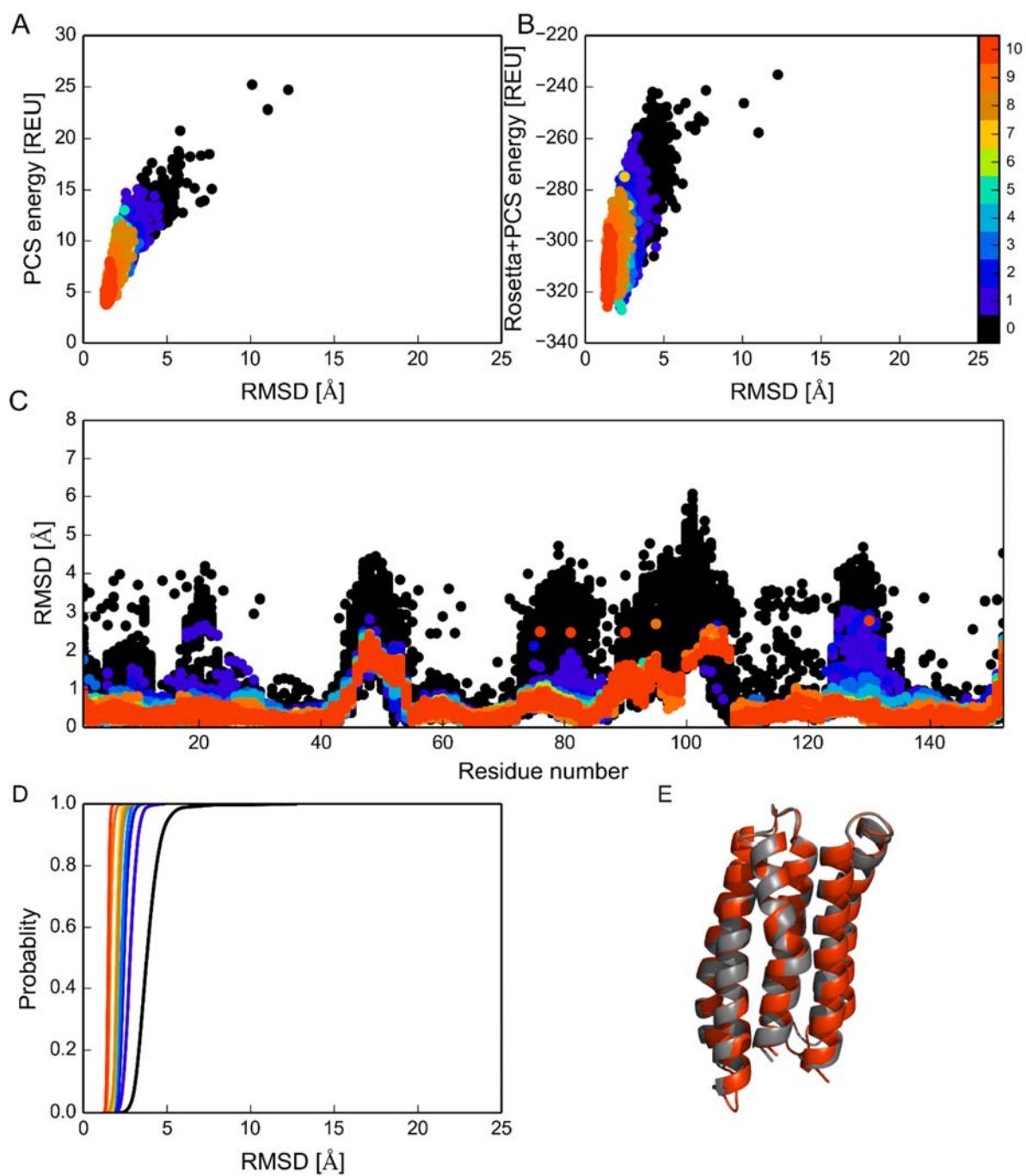
**Figure S1:** Results from PCS-driven iterative GPS-Rosetta applied to target B (ERp29-C). (A) Scatter plot of structures sampled by GPS-Rosetta. The PCS energy is plotted versus the C $\alpha$  RMSD of the NMR structure (PDB ID 2M66 [1]). The results from the different iterations are color-coded, with 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  $\Delta\chi$  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 standard fragment library of the Robetta server, while subsequent iterations took the PCSs into account. (D) Probability density plots illustrating how consecutive iterations shift the conformational sampling towards structures with lower C $\alpha$  RMSD to the crystal structure. (E) Superimposition of the structure with the lowest PCS energy (green) with the crystal structure (gray).



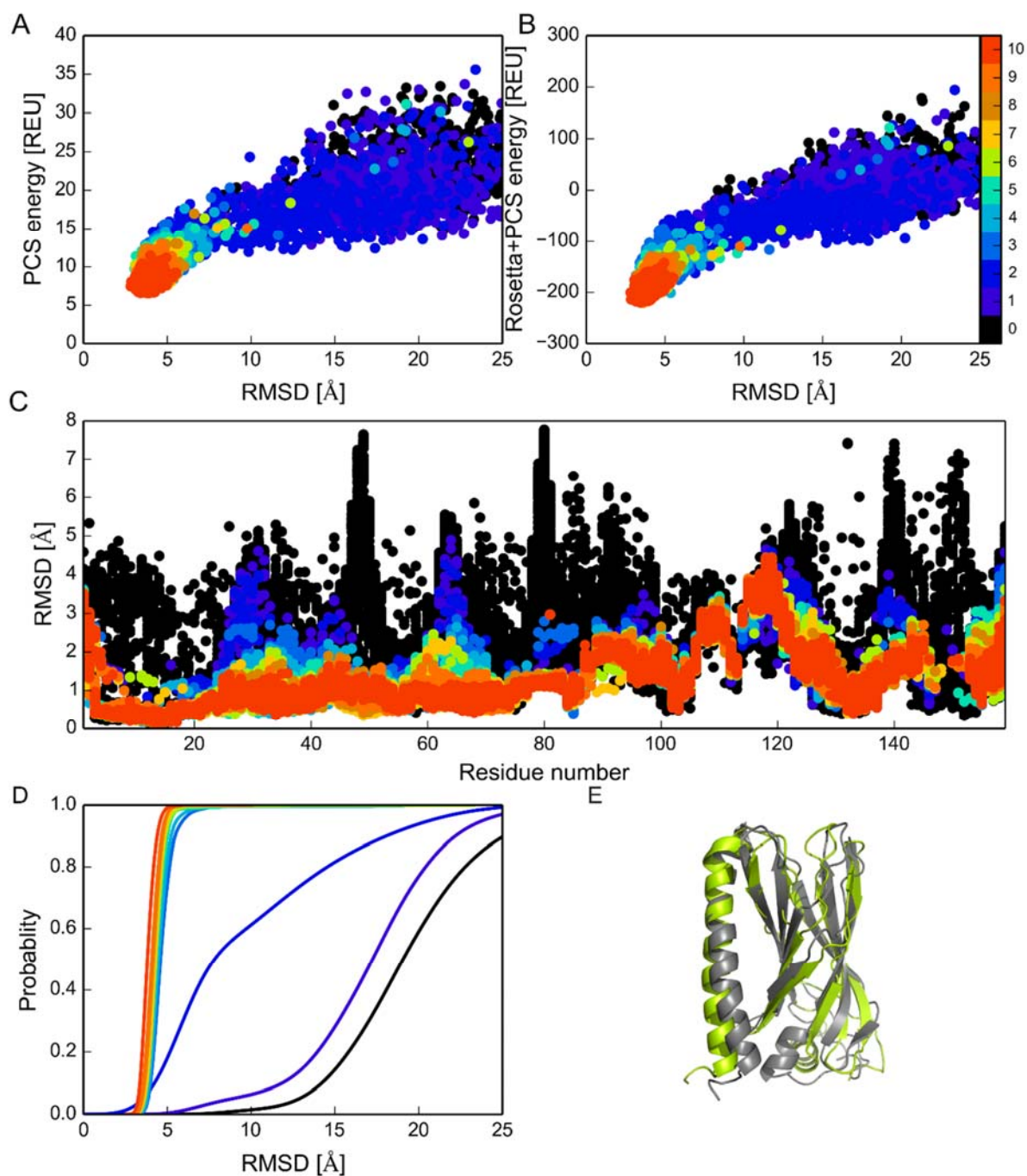
**Figure S2:** Results from PCS-driven iterative GPS-Rosetta applied to target C (OmpX). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:2M06 [2]



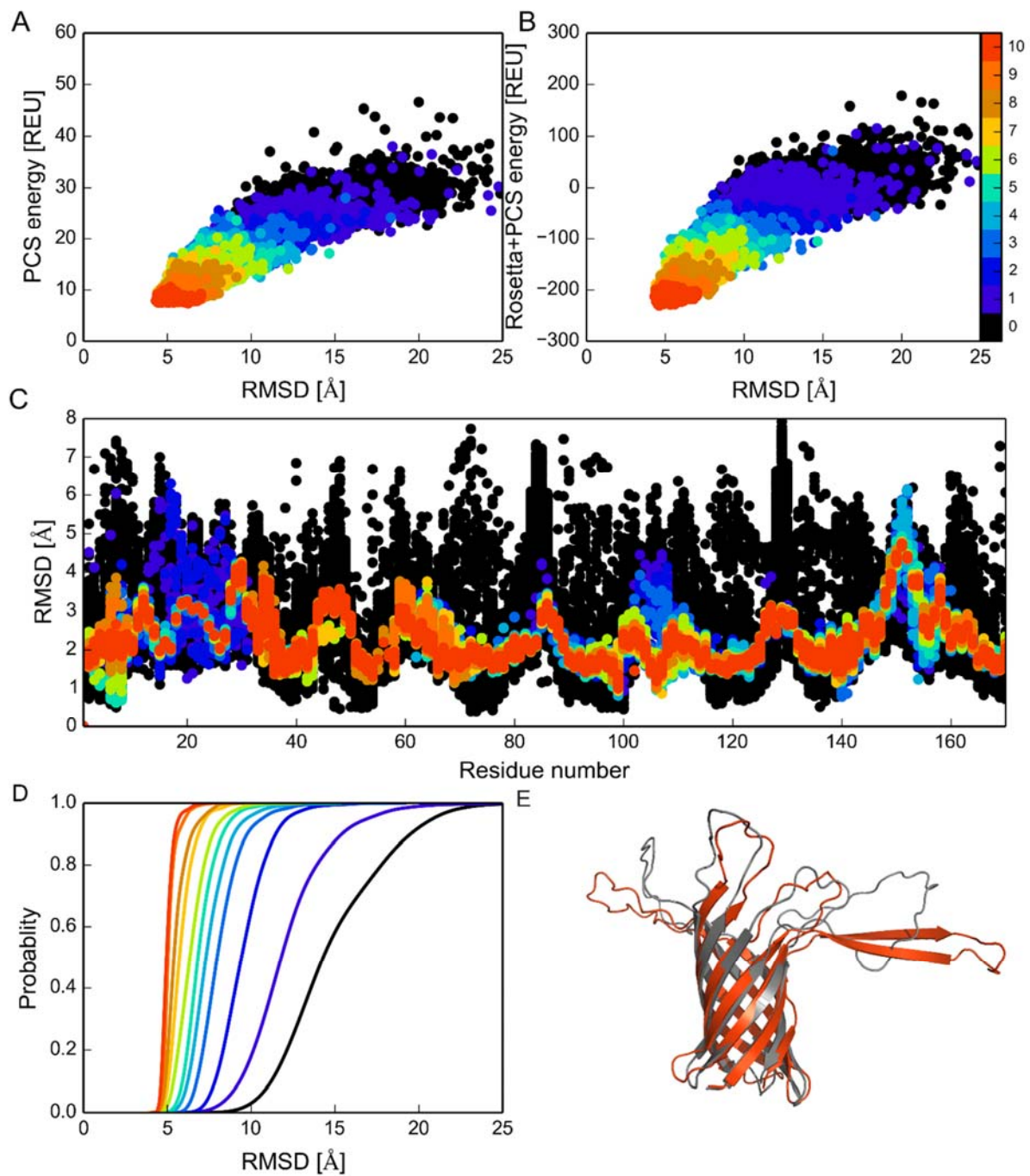
**Figure S3:** Results from PCS-driven iterative GPS-Rosetta applied to target D (Polyketide cyc-like protein). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:2M47



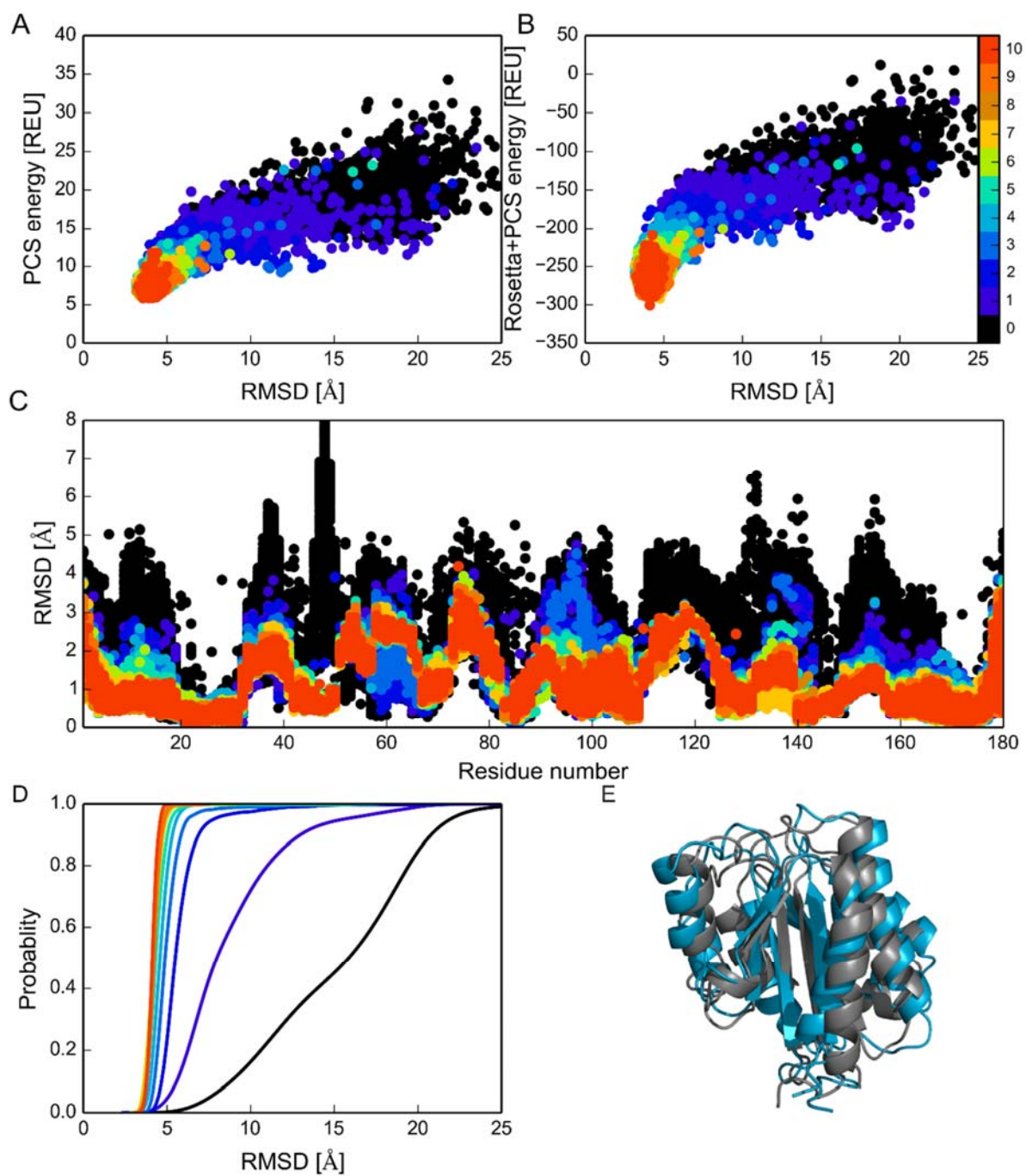
**Figure S4:** Results from PCS-driven iterative GPS-Rosetta applied to target E (CAP). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1S0P [3]



**Figure S5:** Results from PCS-driven iterative GPS-Rosetta applied to target F (LEA protein). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1YYC



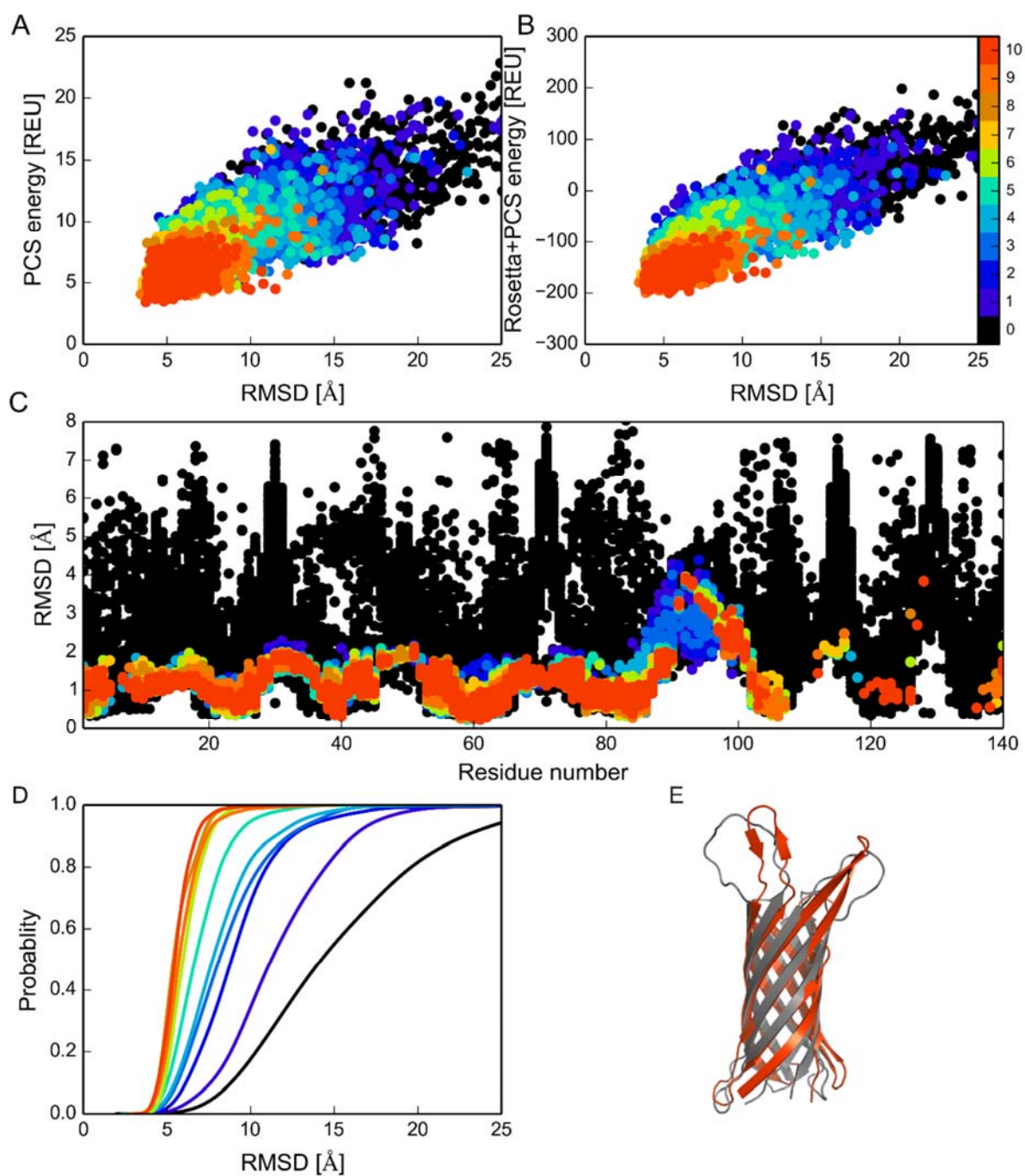
**Figure S6:** Results from PCS-driven iterative GPS-Rosetta applied to target G (OprH). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:2LHF [4]



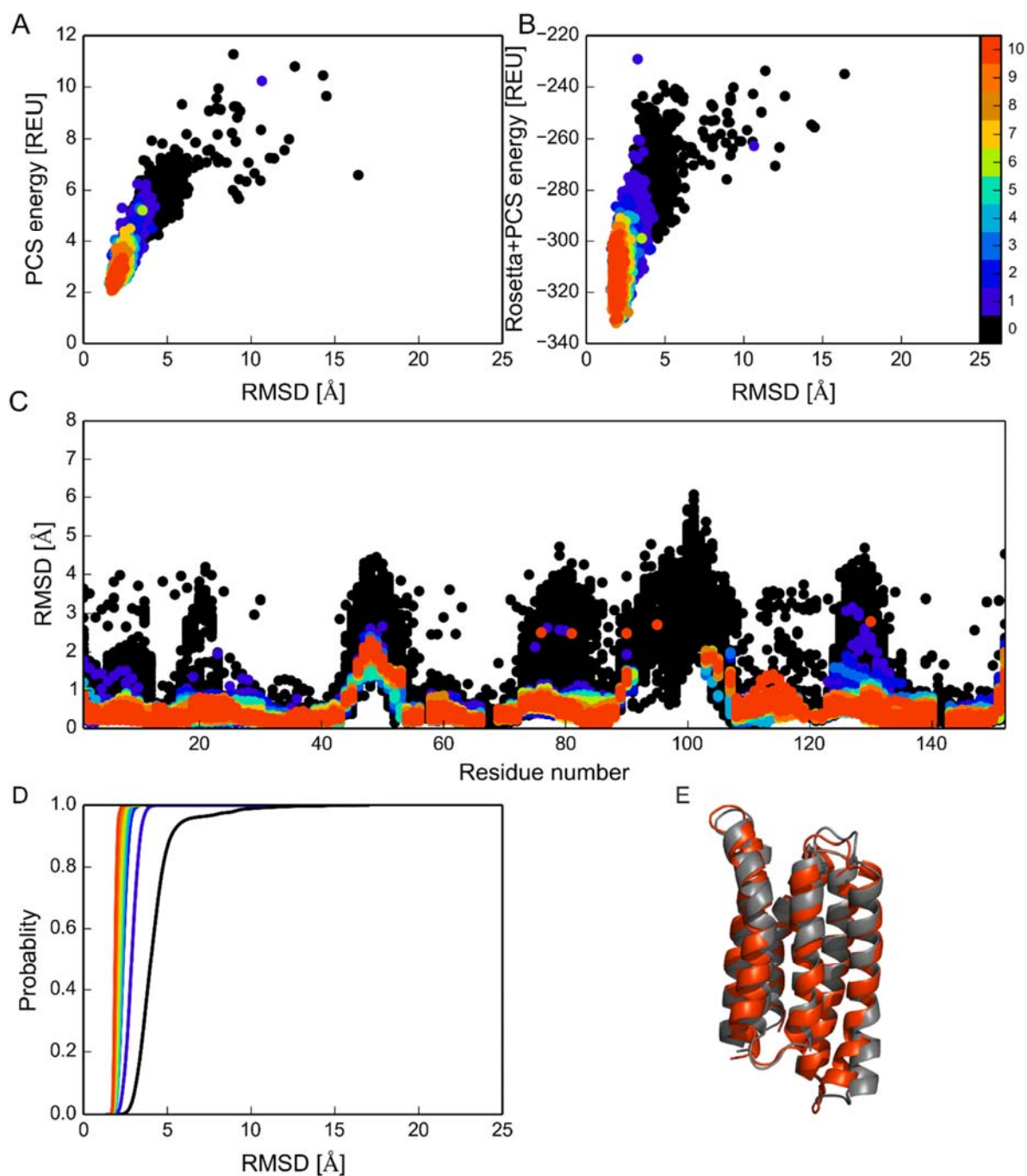
**Figure S7:** Results from PCS-driven iterative GPS-Rosetta applied to target H (Human leukocyte function associated antigen-1). The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1DGQ [5]

**Text S1: PCS data subsets:**

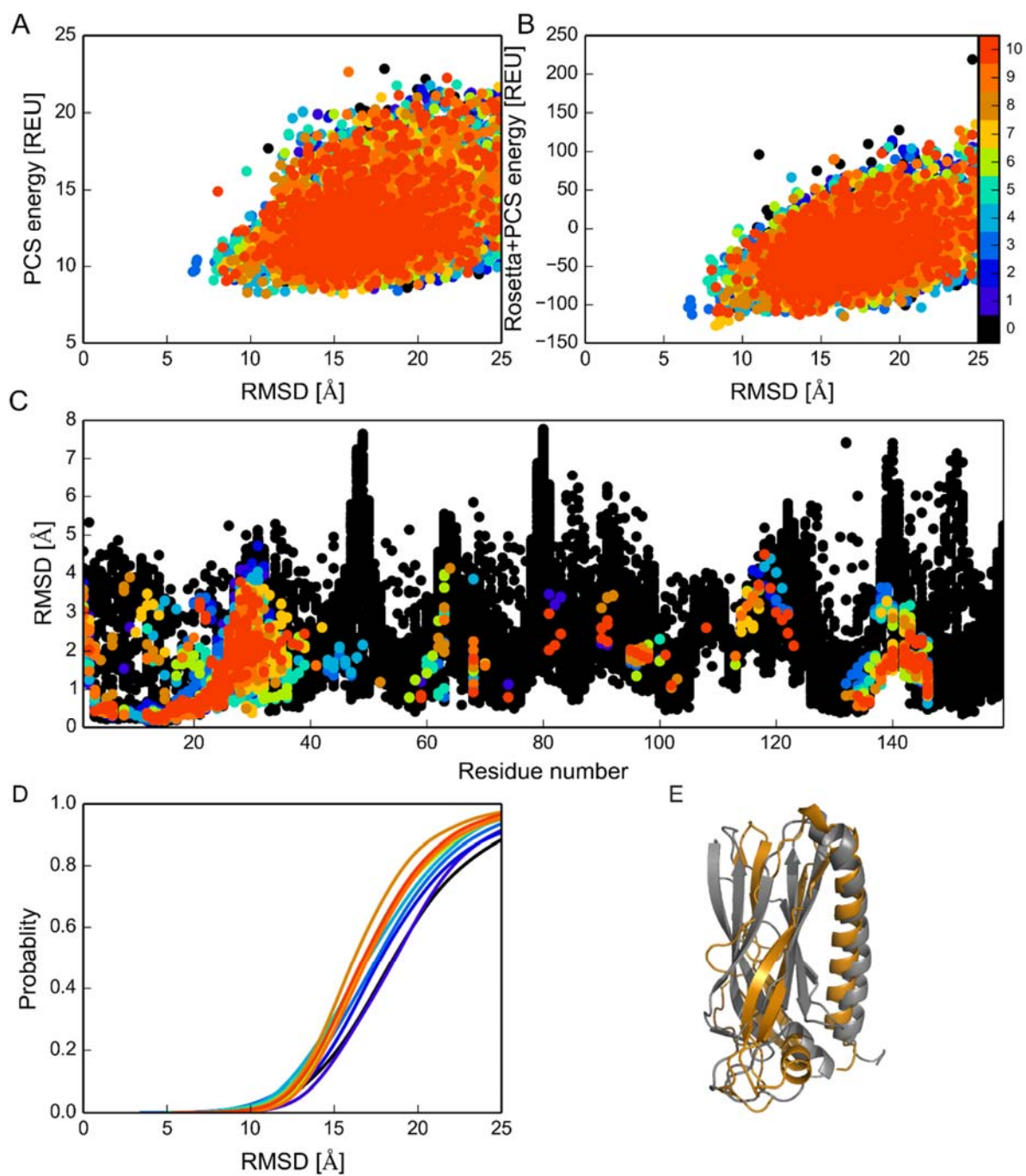
Iterative GPS-Rosetta protocol is tested with reduced PCS data for targets C, E and F. The PCS data contained for only for three metal centers and only two metals per center. The datasets are the same as used for the benchmark set except that one of the metal center is removed at random. The PCSs of the two metals that are included are for Thulium ( $\text{Tm}^{3+}$ ) and Terbium ( $\text{Tb}^{3+}$ ).



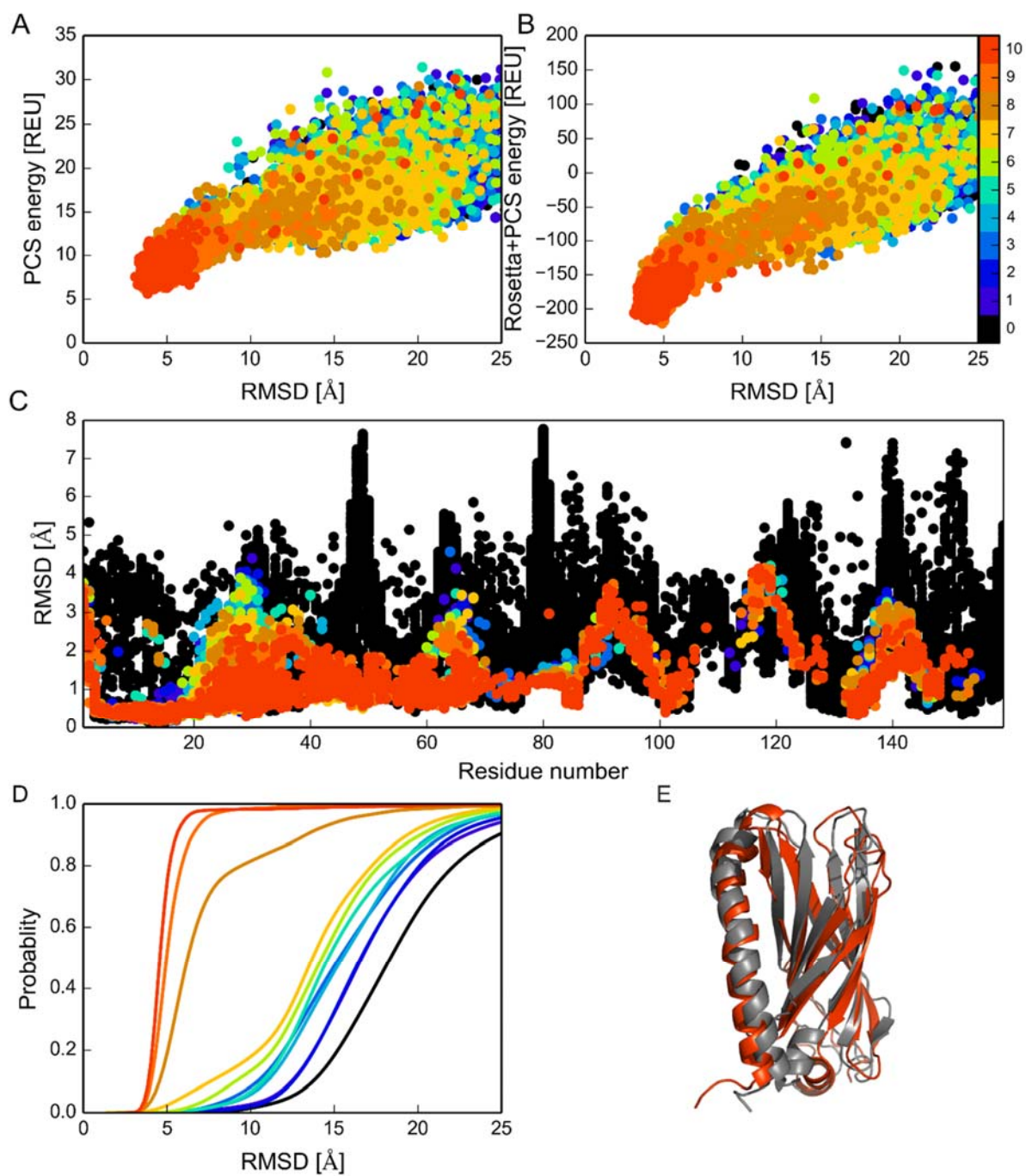
**Figure S8:** Results from PCS-driven iterative GPS-Rosetta applied to target C (OmpX) with PCS data from only three metal centers and with two metals per center. The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:2M06 [2]



**Figure S9:** Results from PCS-driven iterative GPS-Rosetta applied to target E (CAP) with PCS data from only three metal centers and with two metals per center. The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1S0P [3]



**Figure S10:** Results from PCS-driven iterative GPS-Rosetta applied to target F (LEA protein) with PCS data from only three metal centers and with two metals per center. The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1YYC



**Figure S11:** Results from PCS-driven iterative GPS-Rosetta applied to target F (LEA protein) with PCS data from only three metal centers and with four metals per center. The panels descriptions are the same as described in Figure S1 except the reference NMR structure used is PDBID:1YYC

**Table S1:** Comparison of axial and rhombic  $\Delta\chi$ -tensor components of the calculated structures and values reported previously. <sup>a</sup>The axial and rhombic components of the  $\Delta\chi$ -tensors are given in  $10^{-32} \text{ m}^3$ . <sup>b</sup>Same as previous two columns except that parameters are obtained from previous publications, Target A [6], Target B [1].

Target	Mutation site	$\text{Ln}^{3+}$	$\Delta\chi_{\text{ax}}^{\text{a}}$	$\Delta\chi_{\text{rh}}^{\text{a}}$	$\Delta\chi_{\text{ax}}^{\text{b}}$	$\Delta\chi_{\text{rh}}^{\text{b}}$
A (pSR11)	56	$\text{Dy}^{3+}$	-39.7	-24.0	$-36\pm 1$	$-4\pm 2$
		$\text{Tb}^{3+}$	-34.5	-20.8	$-31\pm 1$	$-5\pm 2$
		$\text{Tm}^{3+}$	28.5	12.3	$25\pm 1$	$6\pm 1$
		$\text{Yb}^{3+}$	10.4	5.0	$9\pm 1$	$2\pm 1$
	121	$\text{Dy}^{3+}$	-25.2	-10.3	$-25\pm 8$	$-10\pm 4$
		$\text{Tb}^{3+}$	-20.4	-8.7	$-19\pm 8$	$-6\pm 4$
		$\text{Tm}^{3+}$	18.8	9.0	$18\pm 2$	$3\pm 3$
		$\text{Yb}^{3+}$	-7.1	-4.4	$9\pm 8$	$4\pm 5$
	154	$\text{Dy}^{3+}$	-19.8	-3.9	$-21\pm 1$	$-7\pm 1$
		$\text{Tb}^{3+}$	-16.3	-1.9	$-18\pm 1$	$-7\pm 1$
		$\text{Tm}^{3+}$	11.2	2.2	$12\pm 1$	$6\pm 1$
		$\text{Yb}^{3+}$	4.3	1.0	$5\pm 1$	$3\pm 1$
	169	$\text{Dy}^{3+}$	-34.6	-4.3	$-35\pm 3$	$-4\pm 2$
		$\text{Tm}^{3+}$	29.0	3.3	$31\pm 1$	$4\pm 1$
		$\text{Yb}^{3+}$	10.5	1.3	$11\pm 1$	$2\pm 1$
B(ERp29-C)	C157S/S200C/K204D	$\text{Tb}^{3+}$	-10.4	-4.4	-10.6	-6.2
		$\text{Tm}^{3+}$	8.9	5.6	-9.6	-5.4
	C157S/A218C/A222D	$\text{Tb}^{3+}$	-9.6	-5.5	-12.7	-4.4
		$\text{Tm}^{3+}$	8.5	3.8	11.6	4.8

	C157S/Q241C/N245D	Tb <sup>3+</sup>	-18.1	-3.5	-17.5	-10.3
		Tm <sup>3+</sup>	12.9	3.7	11.5	2.2
	C157	Tb <sup>3+</sup>	-22.5	-5.9	-35.0	-10.2
		Tm <sup>3+</sup>	18.7	4.2	29.3	8.8



















91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108
Y	V	A	T	E	V	F	R	E	E	L	G	A	R	P	D	A	T
H	H	H	H	H	H	C	C	H	H	H	C	C	C	C	C	C	E
*	*	*	*	*		*		*			*				*	*	
*	*	*	*	*		*		*			*				*	*	
*	*	*	*	*		*		*		*	*				*	*	
*	*	*	*	*		*		*		*	*				*	*	
109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126
K	V	L	I	I	I	T	D	G	E	A	T	D	S	G	N	I	D
E	E	E	E	E	E	E	C	C	C	C	C	C	C	C	C	C	H
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144
A	A	K	D	I	I	R	Y	I	I	G	I	G	K	H	F	Q	T
H	H	C	C	C	E	E	E	E	E	E	E	C	C	C	C	C	C
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162
K	E	S	Q	E	T	L	H	K	F	A	S	K	P	A	S	E	F
H	H	H	H	H	H	H	H	H	H	C	C	C	C	H	H	H	H
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
V	K	I	L	D	T	F	E	K	L	K	D	L	F	T	E	L	Q
E	E	E	E	C	C	H	H	H	H	H	H	H	H	H	H	H	H
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
181	182	183	184	185	186	187	188										
K	K	I	Y	V	I	E	G										
H	H	H	H	H	H	H	C										
*	*	*	*	*	*	*	*										
*	*	*	*	*	*	*	*										
*	*	*	*	*	*	*	*										
*	*	*	*	*	*	*	*										

## References:

- [1] H. Yagi, K.B. Pilla, A. Maleckis, B. Graham, T. Huber, G. Otting, Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites., *Structure*. 21 (2013) 883–890. doi:10.1016/j.str.2013.04.001.
- [2] F. Hagn, M. Etzkorn, T. Raschle, G. Wagner, Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins., *J. Am. Chem. Soc.* 135 (2013) 1919–25.

doi:10.1021/ja310901f.

- [3] D. Ksiazek, H. Brandstetter, L. Israel, G.P. Bourenkov, G. Katchalova, K.-P. Janssen, et al., Structure of the n-terminal domain of the adenylyl cyclase-associated Protein (CAP) from *Dictyostelium discoideum*, *Structure*. 11 (2003) 1171–1178. doi:10.1016/S0969-2126(03)00180-1.
- [4] T.C. Edrington, E. Kintz, J.B. Goldberg, L.K. Tamm, Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from *Pseudomonas aeruginosa*., *J. Biol. Chem.* 286 (2011) 39211–23. doi:10.1074/jbc.M111.280933.
- [5] G.B. Legge, R.W. Kriwacki, J. Chung, U. Hommel, P. Ramage, D.A. Case, et al., NMR solution structure of the inserted domain of human leukocyte function associated antigen-1., *J. Mol. Biol.* 295 (2000) 1251–64. doi:10.1006/jmbi.1999.3409.
- [6] D.J. Crick, J.X. Wang, B. Graham, J.D. Swarbrick, H.R. Mott, D. Nietlispach, Integral membrane protein structure determination using pseudocontact shifts., *J. Biomol. NMR.* (2015) 1–9. doi:10.1007/s10858-015-9899-6.