Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Crowning Proteins: Modulating the Protein Surface Properties using Crown Ethers**

*Cheng-Chung Lee, Manuel Maestre-Reyna, Kai-Cheng Hsu, Hao-Ching Wang, Chia-I Liu, Wen-Yih Jeng, Li-Ling Lin, Richard Wood, Chia-Cheng Chou, Jinn-Moon Yang, and Andrew H.-J. Wang**

anie_201405664_sm_miscellaneous_information.pdf
anie_201405664_sm_VideoS1.mp4
anie_201405664_sm_VideoS2.mp4
anie_201405664_sm_VideoS3.mp4
anie_201405664_sm_VideoS4.mp4
anie_201405664_sm_VideoS5.mp4
anie_201405664_sm_VideoS6.mp4
anie_201405664_sm_VideoS7.mp4

Table of Contents

Supplementary materials and methods....2
Protein preparation....2
Crystallization....3
Data collection....3
Structure determination and refinement....4
Thermofluorassay....5
Sample preparation....5
Thermofluorassay....5
Molecular dynamics simulations....5
File preparation....5
Solvent parameterization and system preparation....6
Equilibration....6
Production MD....6
Data evaluation....7
Comparison between crown ether-binding lysines with their PDB crown ether-free counterparts....7
Supplementary Table 1. Summary of crystallization conditions used for each of the protein....8
Supplementary table 2. List of proteins used for crystallization trials , and crystallization improvements derived from the presence of 18-crown-6....9
Supplementary table 3. Data collection and refinement statistics for proteins containing crown ethers....10
Supplementary table 4. Data collection and refinement statistics for crystals not containing crown ethers....11
Supplementary table 5. Simulated B-factors highlighting the relative stabilities of Pin1R14A●CR complexes in-solution and in-crystal....12
Supplementary table 6. Detailed description of simulated systems....13
Supplementary table 7. Detailed description of simulated systems....14
Supplementary figure 1. Thermofluor scans of Pin1R14A, DMP19, and RbmA....15
Supplementary figure 2. Differential crystal growth in the presence and absence of 18-Crown-6 CR....16
Supplementary figure 3. CR dependent changes in crystallization condition composition in RbmA, lysozyme, and hemoglobin....17
Supplementary figure 4. Crown ether dependent crystallization of Pin1 and hemoglobin....19
Supplementary figure 5. Additional CR co-crystal structures....21
Supplementary figure 6. The complex C-crown binding network of the DMP19 dimer....22
Supplementary figure 7. Pin1 molecular dynamics experiments....25
Supplementary figure 8. Comparison of side-chain B-factors of crown ether interacting lysines....26
References....27

Supplementary materials and methods

Collection of protein-PEG complex structures

To explore the potential relevance of crown ethers in crystallization, we collected deposited structures containing low molecular weight polyethylene glycols (lmwPEGs, MW < 600 g/mol) from the PDB. 1,990 X-ray crystallographic protein structures in complex with lmwPEG were downloaded. These lmwPEGs have either linear or ring-shaped conformations, each set comprising 2,113 and 1,404 lmwPEGs, respectively. For each PDB structure, we generated its symmetry related molecules with Pymol[1]. The lmwPEGs were further grouped into intra-unit and inter-unit PEGs. A lmwPEG is regarded to be inter-unit if it contacts ≥2 molecules from different asymmetric units within a 4.5 Å distance; conversely, it is an intra-unit PEG if it only interacts with molecules within a single asymmetric unit. Ring-shaped lmwPEGs contain 969 (69.0%) inter- and 435 (31.0%) intra-unit PEGs, while linear lmwPEGs contain 1,214 (57.5%) inter- and 899 (42.5%) intra-unit PEGs.

Additionally, we grouped dual, inter-unit lmwPEG structures interacting across asymmetric units (within a 4.5 Å distance). The dual lmwPEGs were classified into three groups, including (1) ring-shaped to ring-shaped, (2) ring-shaped to linear, and (3) linear to linear groups, which contained 19, 35, and 38 pairs, respectively.

Protein preparation

Myoglobin, hemoglobin, trypsin and lysozyme freeze-dried samples were purchased from Sigma-Aldrich Co. Purified samples of DMP19, SARS-CoV $3CL^{pro}$, and RbmA were prepared as previously described[2–4].

Cloning, expression, purification and crystallization procedures for Pin1R14A were slightly modified from previously described protocols[5]. The Pin1-R14A mutant gene was cloned into the pET28a vector, resulting in an expression cassette incorporating an N-terminal His_{12}-tag and a thrombin cleavage site. Protein production was carried out in transformed *Escherichia coli* BL21(DE3) cells carrying the correct construct. The protein was purified by Ni-affinity chromatography, digested with thrombin, followed by a second Ni-affinity chromatography step. The purified protein was prepared at 10–20 mg ml^{-1} in 10 mmol/l HEPES buffer pH 7.5 containing 100 mmol/l NaCl and 1 mmol/l DTT for crystallization.

All other proteins were obtained as freeze-dried powders, which were dissolved in different buffers and at different final concentrations, according to previously published crystallization conditions (supplementary table 1).

Crystallization

18-Crown-6 (Sigma-Aldrich) was dissolved in the corresponding protein buffer at a final concentration of 100 mmol/l. Protein solutions were diluted with an equal volume of 18-Crown-6 solution, and incubated for 20 min at 25 °C. Crystallization was carried out by vapor diffusion in a sitting drop set-up (1:1 protein:liquor mixtures).

For sparse-matrix screening, 400 nl protein drops were pipetted into intelli-plates (Art Robbins Instruments) and mixed with 400 nl of mother liquor. A set of 384 common, commercially available conditions was chosen for all proteins, plus a variable number of additional screens, in which it was known that the corresponding protein would crystallize in the absence of 18-crown-6 (supplementary table 1).

Optimizations and crystal reproduction experiments were also performed in a sitting drop setup, in 24-well Cryschem M plates (Hampton), by mixing 1.5 µl protein-additive solution with 1.5 µl liquor. DMP19 crystals grew only in this setup, and in the same conditions as the previously reported CR-free crystals[2]. Surprisingly, crystals grown in this manner belonged to space-group $P4_12_12$, instead of the $P21$ reported in the literature.

Protein crystallization experiments were performed for three weeks, after which the crystal-yielding conditions were tallied (Fig. 2). Conditions in which the protein crystallized only in the absence or presence of 18-crown-6, but not in both, were defined as unique hits. Conversely, in common hits protein crystals grew both with and without 18-crown-6. In these cases, in our analysis crystal improvement was taken into account, as evidenced by a variety of factors, including crystal size, improved mono-crystaline growth, and resolution in the tested crystals.

Data collection

Diffraction data were collected in-house for the hemoglobin●crown co-crystals grown at 20 °C in 30% (w/v) PEG 3350, 0.2 mol/l Lithium sulphate, 0.1 mol/l Tris pH 8.0, using a Rigaku FR-E+ SuperBright rotating anode (λ=1.5418 Å) and an R-AXIS HTC image-plate detector.

Crystallization conditions for the Pin1R14A●crown-ether complex included 2.0-2.5 mol/l ammonium sulphate, 0.1 mol/l HEPES pH 7.5 and 1 mmol/l DTT, with crystals growing at low temperature (4 °C). Data collection took place at synchrotron beam line BL13B1 of the National Synchrotron Radiation Research Center (NSRRC) in Taiwan, using a Rayonix MX300HE CCD Area Detector (λ=0.97622 Å). All crystals were soaked in cryo-protectant (mother liquor plus 20% (v/v) glycerol) and flash-frozen in liquid nitrogen; the diffraction patterns were recorded at cryogenic temperatures. Diffraction data was processed and scaled using the program HKL2000[6].

Structure determination and refinement

All crystal structures were determined by molecular replacement using the program Molrep of the CCP4 program suite[7]. Crystal structures of human hemoglobin (PDB entry 1A3N) and human Pin1R14A (PDB entry 1PIN) were used as the search models for hemoglobin●crown ether and Pin1R14A●crown ether crystal structures determination.

For structure phasing, the DMP19 crystals with space group $P4_12_12$ were soaked with 1 mmol/l $AuCl_3$ (Sigma-Aldrich) for 2 days to obtain the gold derivative crystals. The single-wavelength anomalous dispersion (SAD) X-ray diffraction data was collected at a wavelength of 1.0400 Å on beam line BL12B2 at Spring8, Japan, with an ADSC Quantum 210 CCD detector. The phase of Au-SAD data was determined using the AutoSol module of the Phenix software package[8]. The resolution cut-off of input data for phasing was 3.6 Å with the anomalous difference > 3.0 sigma. Five Au sites were identified and the figure of merit was 0.262. The data resolution was then extended to 1.95 Å in order to calculate the density modification map and further model building was done using the AutoBuild module. The correlation coefficient between model and map was 0.86 and the R_{work}/R_{free} values were 0.22/0.24. This initial model was further refined using a high resolution native data set yielding R_{work}/R_{free} values of 0.17/0.20. The native data set was collected at NSRRC BL13C using ADSC Quantum 210 CCD detector with a wavelength of 1.000 Å.

All structures were refined using a combination of automatic refinement with Refmac5[9] (part of the CCP4 suite[7]) and manual refinement with Coot[10]. Difference Fourier (Fo-Fc) maps were calculated to locate the solvent molecules and ligands. This was performed by removing the ligand, submitting the resulting model to ten cycles of restrained refinement without prior phase information, and finally presenting the electron density around the ligands of the initial model at 1.5σ contour level. Crystallographic and refinement parameters for each structure are shown in supplementary tables 3 and 4.

All structure figures were produced with UCSF Chimera[1,11]. The stereochemistry of the models was validated with PROCHECK[12]. The atomic coordinates and structure factors of the all crystal structures in this work, including hemoglobin●crown ether and Pin1R14A●crown ether, have been deposited in the Protein Data Bank (accession codes on supplementary table 3).

Thermofluorassay

Sample preparation

Protein solutions were mixed with different amounts of crown ether, or PEG400 (final protein concentration was 0.5 mg/ml, final additive concentrations were 0, 10, 50 and 100 mM). The protein-additive solutions were then mixed with the SYPRO Orange Protein Gel Stain stock solution (Invitrogen) for a final 100 fold stain dilution. Each protein assay was performed in the proteins' storage buffer. PEG400 concentrations were estimated by considering an average molecular weight of 400 g/mol, and using an average density of 1.13 g/ml.

Thermofluorassay

Triplicate samples (25 µl) were loaded into 96 well RT-PCR plates, and loaded into a CFX Connect thermocycler (BioRad). Protein samples were heated from 15 ºC to 95 ºC, with a slope of 0.5 ºC/min. SYPRO Orange fluorescence was monitored at 575 nm.The in-built software was employed for data processing, and QtiPlot for the averaging of melting curves. Since higher baselines in thermofluorassays might be indicative of protein aggregation[13], we confirmed via analytical gel filtration that the oligomeric states of DMP19 and RbmA were not affected by CR (data not shown).

Molecular dynamics simulations

All molecular dynamics simulations were run using the Amber12 software package[14,15], with files prepared with a combination of AmberTools12 and Ambertools13. Data were wrapped and analyzed using VMD[16], PBCTools 2.6, and ptraj[17].

File preparation

PDB files for Pin1R14A crystal structures presenting either 18-Crown-6 ether (PDBID 3WH0) or PEG400 (PDBID 2ZR5) in space-group $P3_121$ were used for all simulations. First, protein models were completed by manually modeling the disordered loop spanning amino acids 33 to 44 using Coot. Subsequently, the structure was subjected to ten cycles of phase-free restrained refinement using Refmac5[9]. Further, since the full PEG molecule was not defined in the 2ZR5 crystal structure, nona-ethylene glycol and dodecaethylene glycol were modeled into it as roughly equivalent to the PEG400 and PEG500 average sizes. Next, files were prepared for simulation using the H++ server[18]. The resulting reasonable protonation predictions at pH 7.5 were manually added into the original files. For in-crystal simulations, the $P3_121$ unit cell was constructed using UCSF Chimera[11], including all crystallographic

waters and ligands (i.e. waters and ligands whose positions could be experimentally determined during structure solution and refinement). For the in-solution simulations, samples were prepared in the same way but with just one water-free protein molecule.

Solvent parameterization and system preparation

Prepared protein files were further manipulated using the AddToCell software[19,20] as shown in supplementary tables 6 and 7. Both simulations in-crystal, and in-solution were solvated in periodic SPC/E water boxes[21,22], which corresponded to the unit cell geometry in the former, while an orthorhombic box extending at least 15 Å in every direction from the protein surface was used for the latter. The boxes had been filled with ammonium and sulphate ions, as well as HEPES, DTT, and either PEG400 or 18-crown-6 molecules, in ratios which closely corresponded to the crystallization conditions. HEPES, DTT, PEG400 and 18-crown-6 were parametrized with antechamber[14], while SPCE waters, and ammonium parameters were obtained from previous studies[20]. Finally, sulphate was parameterized by a combination of previous information and RESP calculations[23,24]. Since high salt concentrations were being used, several water concentrations were tested in order to achieve volume consistency.

Equilibration

All systems were minimized for 5000 cycles with 500 kcal/mol − $Å^2$ harmonic restraints for the protein component, followed by restraint free steepest descent minimization for the entire system for a further 5000 cycles. The system was then equilibrated for 50 ps to 300 K using a randomly seeded Langevin thermostat (γ=5 ps^{-1}), and weak 10 kcal/mol − $Å^2$ restraints, with SHAKE-constrained[25] bond lengths for hydrogens. In the final step, constant pressure was introduced by slowly raising system pressure to one atmosphere for a further 50 ps at 300 K, with all other values unchanged.

System volume was then monitored by production dynamics of a maximum of 20 ns with an 8.0 Å cut-off. As soon as system volume deviated more than 0.5% from the theoretical cell volume, the simulation was halted, and the water concentration was readjusted, until the volume remained within the 0.5% error for at least the first 20 ns.

Production MD

Equilibrated production systems were run for 100 ns at constant pressure (1 atm) and temperature (300 K), maintaining the SHAKE bond length constrains[25], and Langevin thermostat[26]. Additionally, a 96x96x96 PME grid yielded high accuracy in unit cell electrostatic energy calculation.

Data evaluation

MD data was analyzed using a combination of VMD[16] and ptraj[17]. Wrapping, aligning and centering were performed in VMD, while calculation of per-residue-root-mean-square fluctuations, and deviations (RMSF and RMSD, respectively)were done using ptraj. The only exception were the crown ethers, which, due to a rotational motion along the ring's center, had artificially increased RMSD, and per-atom B-factors. Consequently, we calculated a dummy atom with coordinates corresponding to the center of mass (com) of each crown ether molecule, and then calculated their RMSF values. Accordingly, crown ether RMSDs presented in this work refer to crown ether com RMSDs, and were not calculated for each crown-ether atom independently. Finally, RMSFs were converted to B-factors, as previously described in literature[27].

Comparison between crown ether-binding lysines with their PDB crown ether-free counterparts

In order to evaluate the lysine stabilization capacity of crown ethers, we downloaded 30 Pin1R14A and 204 hemoglobin structures from the PDB, that is the entire high-resolution structural catalog for both proteins. B-factors for all atoms in K97 and K117 in Pin1R14A; K11 and K40 in α-hemoglobin; and K59 in β-hemoblogin were obtained from the PDB files. The per-residue B-factor was calculated by averaging the B-factors of the corresponding lysine side-chain atoms of a single structure, resulting in one data point in the B-factor distribution for each lysine. The average B-factor of the entire distribution was obtained by averaging all per-residue B-factors. B-factor distributions were plotted as a box-and-whisker plot using Graphpad Prism, in which the average B-factor of the distribution was compared to the corresponding per-residue value of our crystal structures.

Supplementary Table 1. Summary of crystallization conditions used for each of the protein.

[1]: Molecular Dimensions [2]: Expanded JCSG core screen was mixed in-house based on[28] [3]: Emerald Biosystems [4]:Hampton [5]: Jena Bioscience [6]: Qiagen [*]: common five screens, used for all proteins in the test. They are described once, and thereafter abbreviated as "common five screens"

Protein	Special characteristics	crystallization screen	Temperature (C)	Total number of conditions	Citation	Storage buffer	concentration
PIN1R14A	LmwPEG is required for crystallization	*Structure screen combination*[1*] *Stura Footprint*[1*] *3D structure*[1*] *Clear Strategy combination*[1*] *Expanded JCSG core screen*[2*] Cryo Screen I & II[3]	20/4	960	[5,29]	10 mM HEPES pH 7.5 100 mM NaCl 1 mM DTT	10-20 mg/ml
Hemoglobin	Commercially available	Common five screens Macrosol[1] Grid Screen Ammonium Sulfate[4] Grid screen Sodium Malonate[4] Wizard screen I&II[3] Cryo Screen I & II[3]	20	672		20 mM Tris/HCl pH 8	20 mg/ml
DMP19	Recently solved, no information on CR/lmwPEG requirements	Common five screens	20	384	[2]	20 mM Tris/HCl pH 8	20 mg/ml
SARS-CoV $3CL^{pro}$	Easily available	Common five screens Cryo Screen I & II[3] Grid screen PEG6K[4] PEG screen[5]	20	576	[3]	50 mM Tris/HCl pH 8	10 mg/ml
RbmA	Known CR-dependent crystallization improvement.	Common five screens Wizard screen I&II[3] Wizard screen III & IV[3] Cryo Screen I & II[3] Screen I & II[4] Index[4] JCSG core I suite[6]	20	960	[4]	20 mM Tris/HCl pH 8 100 mM NaCl	16-24 mg/ml
Lysozyme	Commercially available	Common five screens Cryo Screen I & II[3]	20	576		20 mM Tris/HCl pH 8	40 mg/ml
Myoglobin	Commercially available	Common five screens Cryo Screen I & II[3]	20	480		20 mM Tris/HCl pH 8	20 mg/ml
Trypsin	Commercially available	Common five screens Cryo Screen I & II[3]	20	480		20 mM Tris/HCl pH 8	20 mg/ml

Supplementary table 2. List of proteins used for crystallization trials , and crystallization improvements derived from the presence of 18-crown-6.

V: confirmed positive effect. X: confirmed negative effect/no effect.

Protein	Optimal 18-Crown-6 concentration (mM)	New crystallization condition	New Space group	Crystal quality improvement	Conformational change	Crown ether involved in crystal contacts
DMP19	25	V	V	V	V	V
PIN1R14A	25	V	X	V	X	V
RbmA	50	V	V	V	V	X
Hemoglobin	25	V	X	X	X	V
SARS-CoV $3CL^{pro}$	25	V	X	X	X	X
Lysozyme	25	V	X	X	X	X
Myoglobin	25	X	X	X	X	X
Trypsin	X	X	X	X	X	X

Supplementary table 3. Data collection and refinement statistics for proteins containing crown ethers.

Values corresponding to the highest resolution shell are shown in parentheses.

	Pin1R14A/CR (3WH0)	hemoglobin/CR (3WHM)	DMP19/CR (3WUR)	DMP19/Au/CR
Data collection				
Wavelength (Å)	1.0	1.5418	1.0	1.04
Space group	$P3_121$	$P2_1$	$P4_12_12$	$P4_12_12$
Cell dimensions				
a, *b*, *c* (Å)	68.47, 68.47, 80.22	54.16, 104.64, 59.96	67.10, 67.10,176.91	67.32, 67.32, 176.69
α, β, γ (°)	90, 90, 120	90, 108.69 , 90	90, 90, 90	90, 90, 90
Resolution (Å)	30-1.6 (1.66-1.60)	25-1.85 (1.92-1.85)	30-1.45 (1.48-1.45)	30-1.95 (2.02-1.95)
Observed reflections	316,092	106,505	694,135	240,339
Unique reflections	29,138	50,382	71,243	30,264
R_{merge}	3.7 (52.8)	3.2 (30.1)	4.1 (37.3)	6.4 (41.4)
$I/\sigma(I)$	55.3 (4.7)	23.0 (2.1)	58.4 (5.6)	36.7 (7.0)
Completeness	99.8 (100.0)	93.4 (94.0)	98.0 (97.9)	98.8 (99.9)
Multiplicity	10.8 (10.9)	2.1 (2.2)	9.7 (6.3)	7.9 (8.0)
Refinement				
Resolution (Å)	30-1.6	25-1.85	30-1.45	
No. of reflections	25,185/1,348	41,996/2,256	67,437/3,570	
R_{work}/R_{free}	21.0/24.7	20.4/25.4	16.9/20.4	
No. of atoms/ Avg B factor (Å^2)				
Protein	1,156/16.3	4,300/21.3	2,594/18.5	
Crown ether	36/32.1	90/32.4	108/24.4	
Water	192/31.4	441/30.1	509/35.0	
Other	13/42.3	178/19.2	10/22.4 (SO_4) 8/20.1 (Ethylene glycol) 10/24.0 (Tartrate)	
RMSD				
Bond lengths (Å)	0.029	0.008	0.033	
Bond angles (°)	2.77	1.30	3.02	
Ramachandran statistics (%)				
Most favoured	92.7	93.3	95.5	
Additionally allowed	7.3	6.7	4.5	
Generously allowed	0.0	0.0	0.0	
Disallowed	0.0	0.0	0.0	

Supplementary table 4. Data collection and refinement statistics for crystals not containing crown ethers.

	SARS-CoV $3CL^{pro}$	Lysozyme
Data collection		
Wavelength (Å)	1.0	1.5418
Space group	C2	$P4_32_12$
a, *b*, *c* (Å)	108.51, 81.62, 53.35	77.08, 77.08, 38.32
α, β, γ (°)	90.0, 105.17, 90.0	90.0, 90.0 , 90.0
Resolution (Å)	30-2.2 (2.28-2.20)	30-1.5 (1.55-1.50)
Observed reflections	80,792	140,935
Unique reflections	22,488	17,676
R_{merge}	3.6 (27.0)	5.1 (35.2)
$I/\sigma(I)$	31.4 (4.0)	25.2 (6.0)
Completeness	98.7 (94.5)	92.6 (97.0)
Multiplicity	3.6 (3.3)	8.0 (4.9)
Refinement		
Resolution (Å)	30-2.2	30-1.5
No. of reflections	19,146/1,042	16,705/899
R_{work}/R_{free}	22.5/26.5	28.4/32.4
MR model (PDB code)	2ZU5	1LKR
Z	1	1

After initial refinement for 15 cycles, and verification that no crown-ethers were present in the structure, no further processing was undertaken.

Supplementary table 5. Simulated B-factors highlighting the relative stabilities of Pin1R14A●CR complexes in-solution and in-crystal

	Crystal simulation with CR	Crystal simulation without CR	In-solution simulation with CR	In-solution simulation without CR
Pin1R14A				
C-crown	17.9±7.8 Å^2	NA	198.5 Å^2	NA
C'-crown	17.9±7.8 Å^2	NA	49820.3 Å^2	NA
K-crown	21.06±15.6 Å^2	NA	78.4 Å^2	NA
K'-crown	21.06±15.6 Å^2	NA	13238.4 Å^2	NA
K97	24.2 Å^2	43.3 Å^2	113.7	124.9
K117	19.5 Å^2	49.9 Å^2	51.9	78.2

Supplementary table 6. Detailed description of simulated systems.

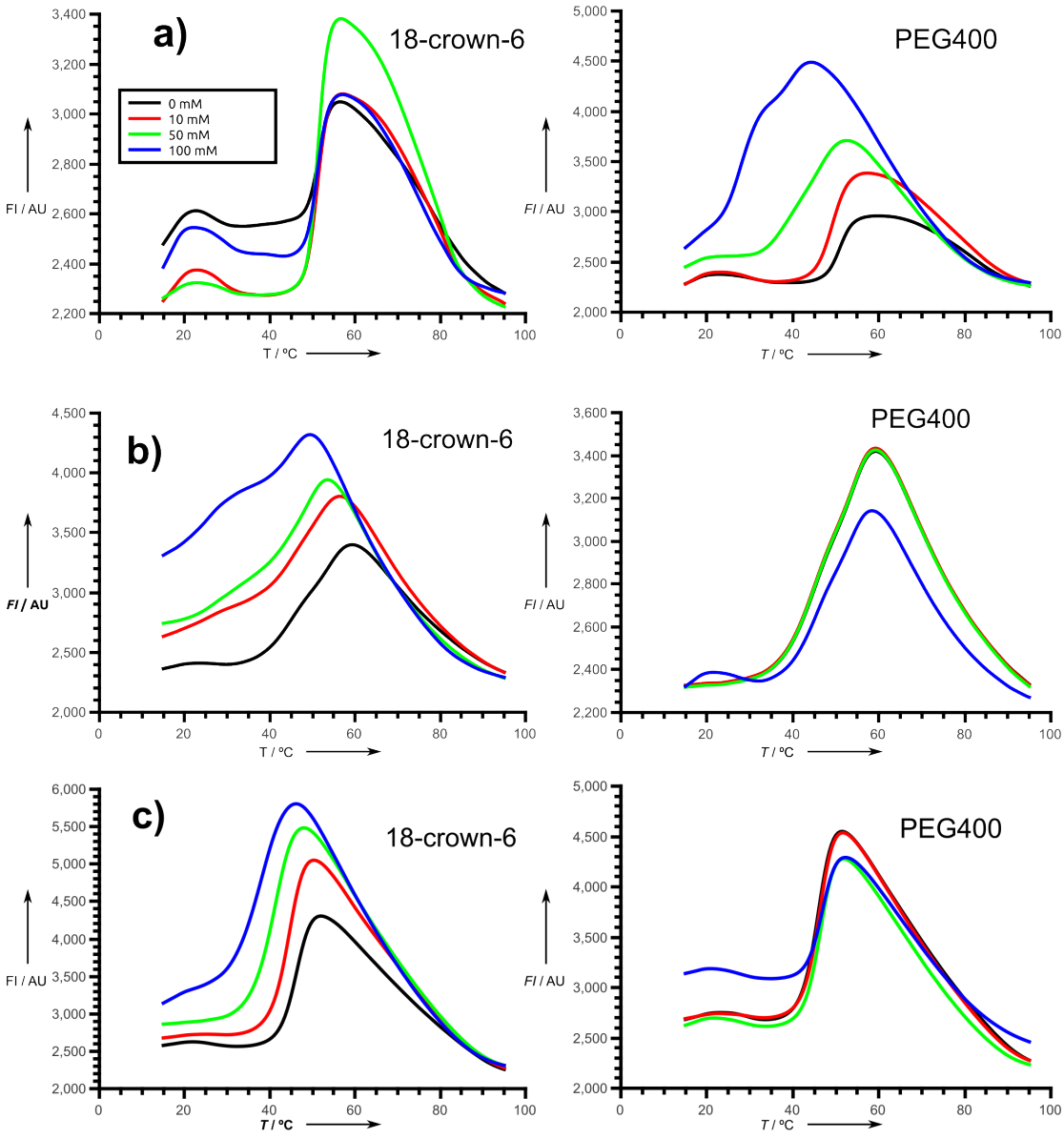
Crystallographic components refer to molecules or atoms which were already present within the deposited PDB files; simulated components were generated in the process of setting up the molecular dynamics system. * and **: solute concentration, in mol/l, are corrected by the solvent content percentage, which is ~58% for in-crystal simulations, and ~96% for in-solution simulations.

		3WH0							
		Crystal				Solution			
		Crown		No crown		Crown		No crown	
		Number	Concentration (M)*	Number	Concentration*	Number	Concentration**	Number	Concentration (M)**
Crystallographic components	Protein	6	0.03	6	0.03	1	0.03	1	0.03
	CR	12	0.1	0	0	2	0.096	0	0
	PEG 400	0	0	0	0	0	0	0	0
	Water	1146	55.68	1146	57.15	0	50.77	0	52.18
	Sulphate	6	0.94	6	0.94	1	1.03	1	1.03
Simulated components	CR	0	0.1	0	0	28	0.096	0	0
	PEG 400	0		0		0		0	
	DTT	1	0.009	1	0.009	1	0.003	1	0.003
	HEPES	6	0.05	6	0.05	6	0.019	6	0.019
	Ammonium	203	1.78		1.78	641	2.05	641	2.05
	Sulphate	106	0.94		0.94	320	1.03	320	1.03
	Water	5190	55.68	5358	57.15	15848	50.77	16266	52.18
Theoretical volume ($Å^3$)			325714.89		325714.89		537373.99		537373.99

Supplementary table 7. Detailed description of simulated systems.

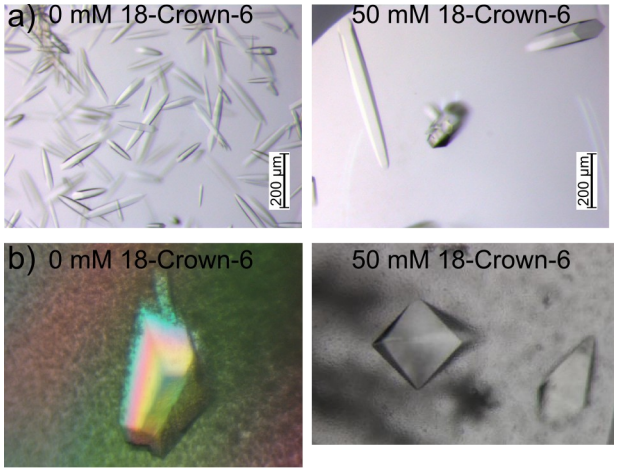
Crystallographic components refer to molecules or atoms which were already present within the deposited PDB files; simulated components were generated in the process of setting up the molecular dynamics system. * and **: solute concentration, in mol/l, are corrected by the solvent content percentage, which is ~58% for in-crystal simulations, and ~96% for in-solution simulations.

		2ZR5							
		Crystal				Solution			
		PEG 500		PEG400		PEG500		PEG400	
		Number	Concentration (M)*	Number	Concentration*	Number	Concentration**	Number	Concentration (M)**
Crystallographic components	Protein	6	0.03	6	0.03	1	0.028	1	0.028
	CR	0	0	0	0	0	0	0	0
	PEG	6	0.097	6	0.097	1	0.059	1	0.059
	Water	654	52.9	654	53.5	0	50.34	0	50.35
	Sulphate	6	0.94	6	0.94	1	1.30	1	1.30
Simulated components	CR	0	0	0	0	0	0	0	0
	PEG	5	0.097	5	0.097	19	0.059	19	0.059
	DTT	1	0.009	1	0.009	1	0.003	1	0.003
	HEPES	6	0.05	6	0.05	36	0.106	36	0.106
	Ammonium	259	2.29	259	2.29	881	2.59	881	2.59
	Sulphate	137	1.22	137	1.22	441	1.30	441	1.30
	Water	5343	52.9	5402	53.5	17103	50.34	17107	50.35
Theoretical volume ($Å^3$)			324076.28		324076.28		587485.19		587485.19



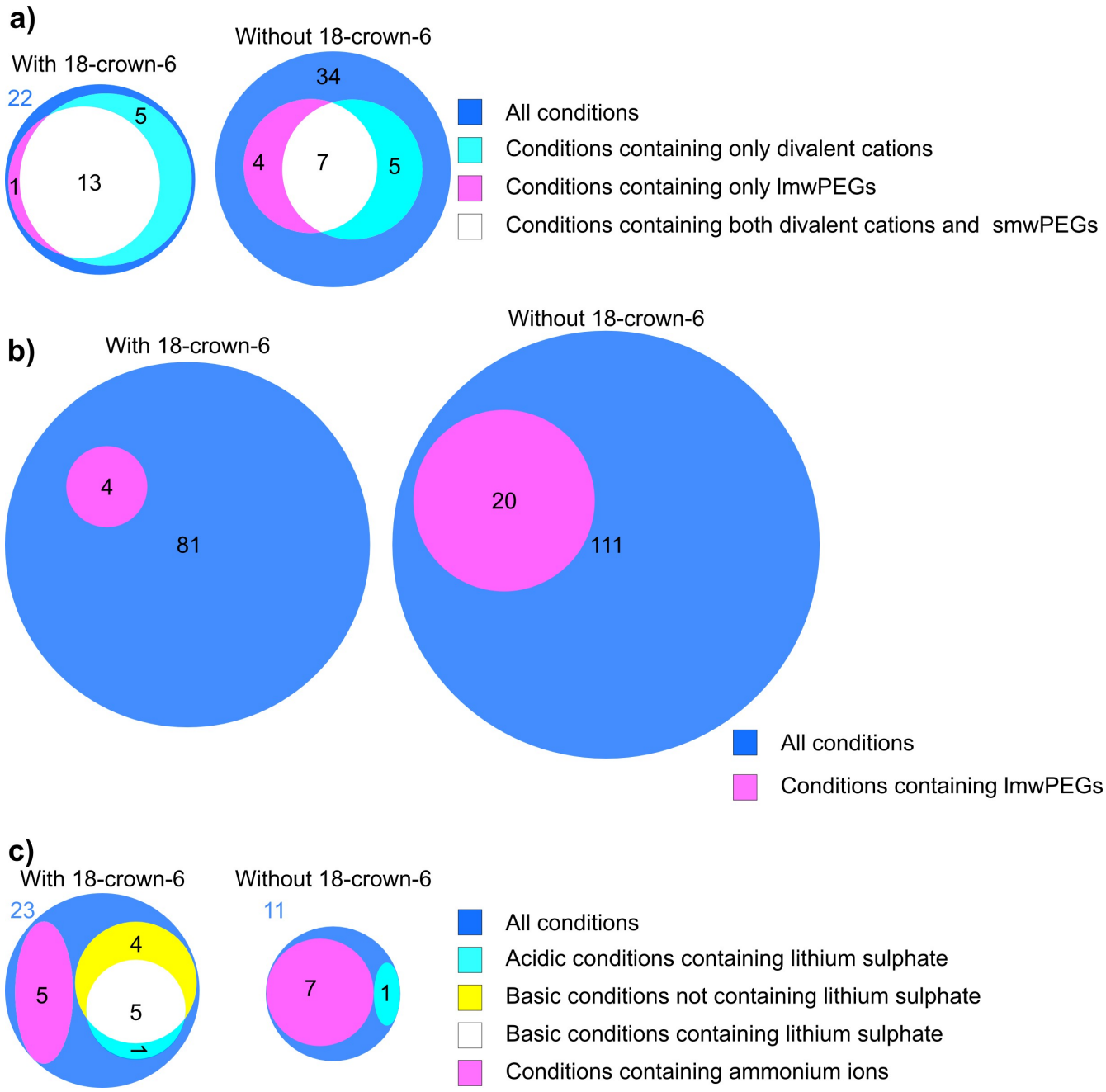
Supplementary figure 1. Thermofluor scans of Pin1R14A, DMP19, and RbmA.

Protein-additive mixtures were analyzed via thermofluorassay with variable concentrations of additives. Fluorescence intensities (*FI*) were measured at increasing temperatures (*T*). a) Pin1R14A melting curves in the presence of varying concentrations of 18-crown-6 (left), and PEG400 (right). b) DMP19 melting curves under the same conditions as a). c) RbmA thermofluorassay control experiment.



Supplementary figure 2. Differential crystal growth in the presence and absence of 18-Crown-6 CR.

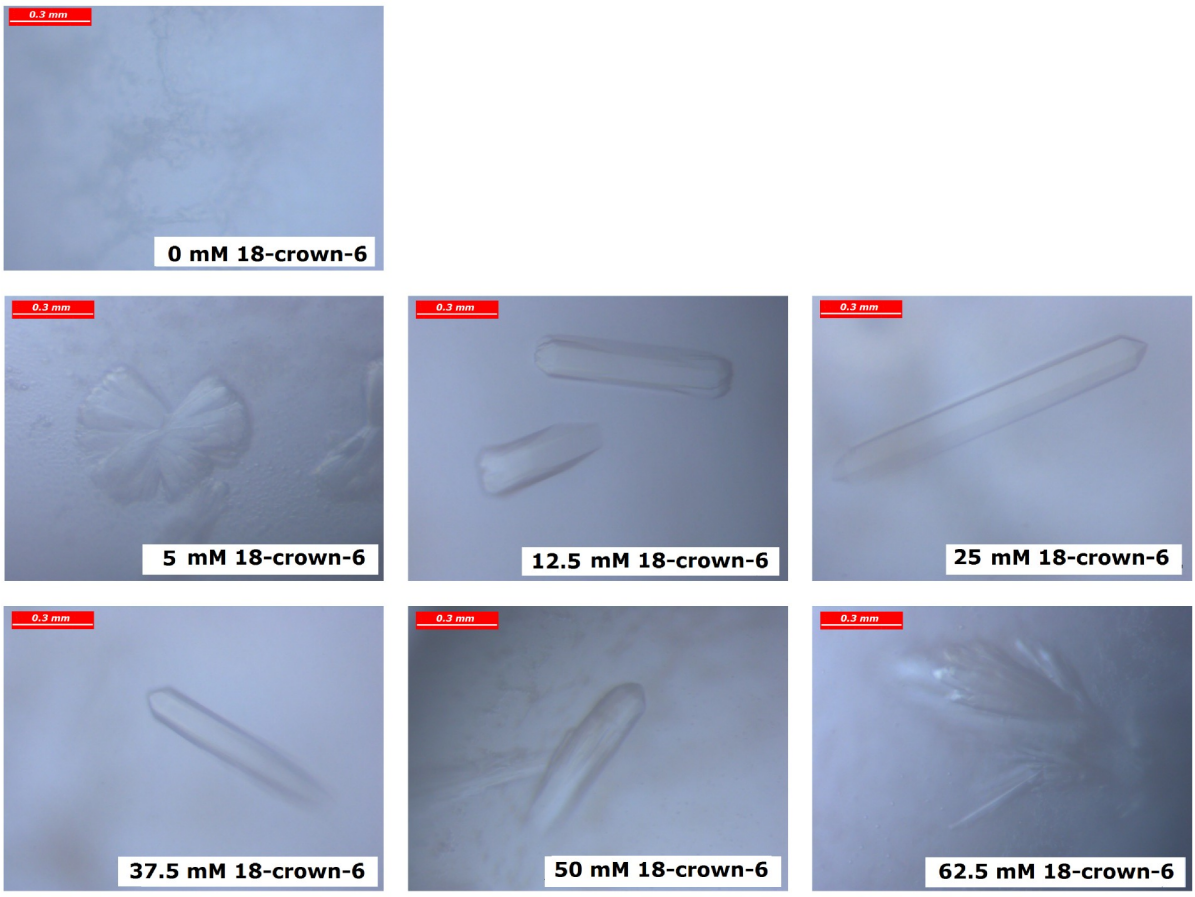
a) RbmA crystal growth without CR (left) and with CR (right). CR addition resulted in large diffracting crystals, which allowed for space-group determination and structure solution.Condition composition: 30% (w/v) PEG400, 0.1M Tris pH 8.5, 0.2M $MgCl_2$. b) DMP19 growth without CR (left) and with CR (right).



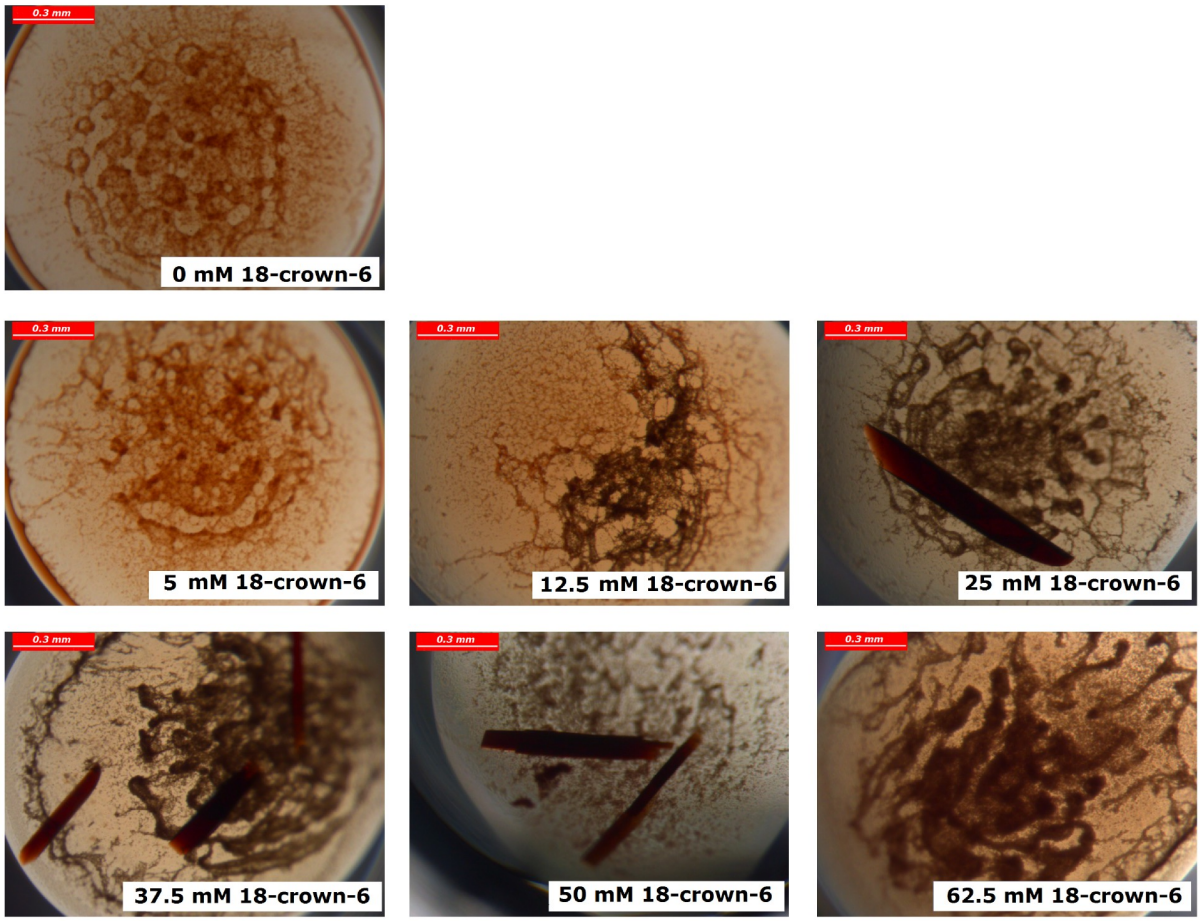
Supplementary figure 3. CR dependent changes in crystallization condition composition in RbmA, lysozyme, and hemoglobin.

a) Results for RbmA. b) Results for lysozyme c) Results for hemoglobin. In each of the venn diagrams presented in this figure, the size of the blue circle represents the total number of hits for each of the proteins, while the different colored circles represent the number of hits of a certain chemical species, or pH range. Intersections between different circles are white. Each panel contains a figure legend with a key for the corresponding diagram.

a)


0.3 mm
0 mM 18-crown-6
5 mM 18-crown-6
12.5 mM 18-crown-6
25 mM 18-crown-6
37.5 mM 18-crown-6
50 mM 18-crown-6
62.5 mM 18-crown-6

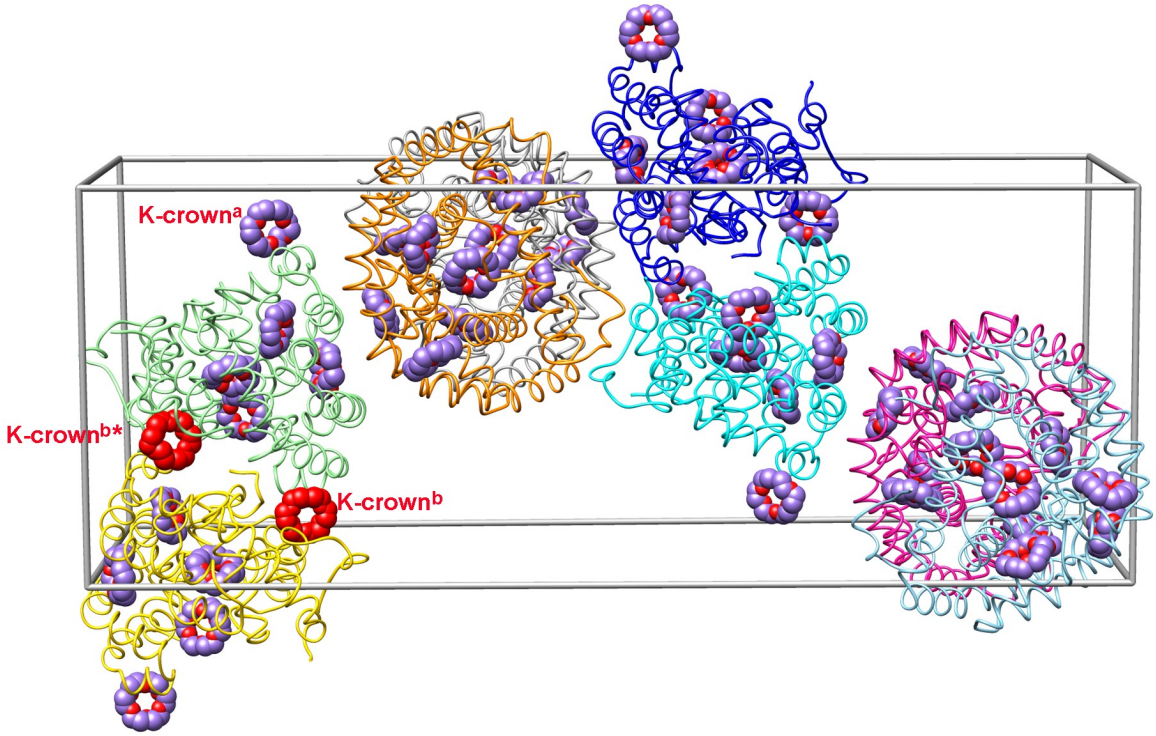
b)


0.3 mm
0 mM 18-crown-6
5 mM 18-crown-6
12.5 mM 18-crown-6
25 mM 18-crown-6
37.5 mM 18-crown-6
50 mM 18-crown-6
62.5 mM 18-crown-6

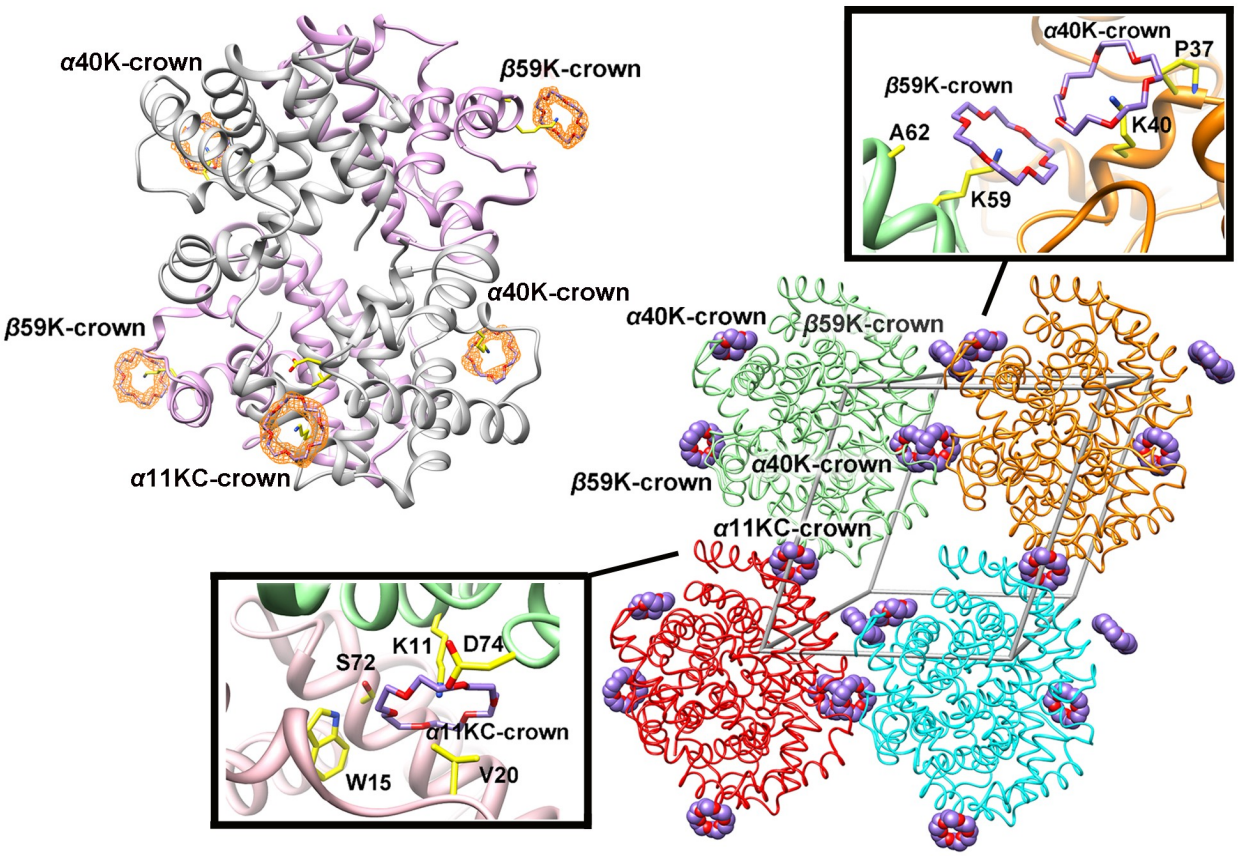
Supplementary figure 4. Crown ether dependent crystallization of Pin1 and hemoglobin.

Pin1 (a) and hemoglobin (b) were subjected to crystallization in a hanging drop set-up, (conditions for Pin1: 2.0-2.5 M ammonium sulphate, 0.1 M HEPES pH 7.5 and 1 mM DTT, 4ºC. Conditions for hemoglobin: 30% (w/v) PEG 3350, 0.2 M Lithium sulphate, 0.1 M Tris pH 8.0, 20ºC). Different amounts of crown ether were mixed with the protein, maintaining protein concentration constant. In both cases optimal crown ether concentration was ~25 mM. Lower crown ether concentrations did not yield high quality crystals, and higher concentrations resulted in protein precipitation or twinning.

a)

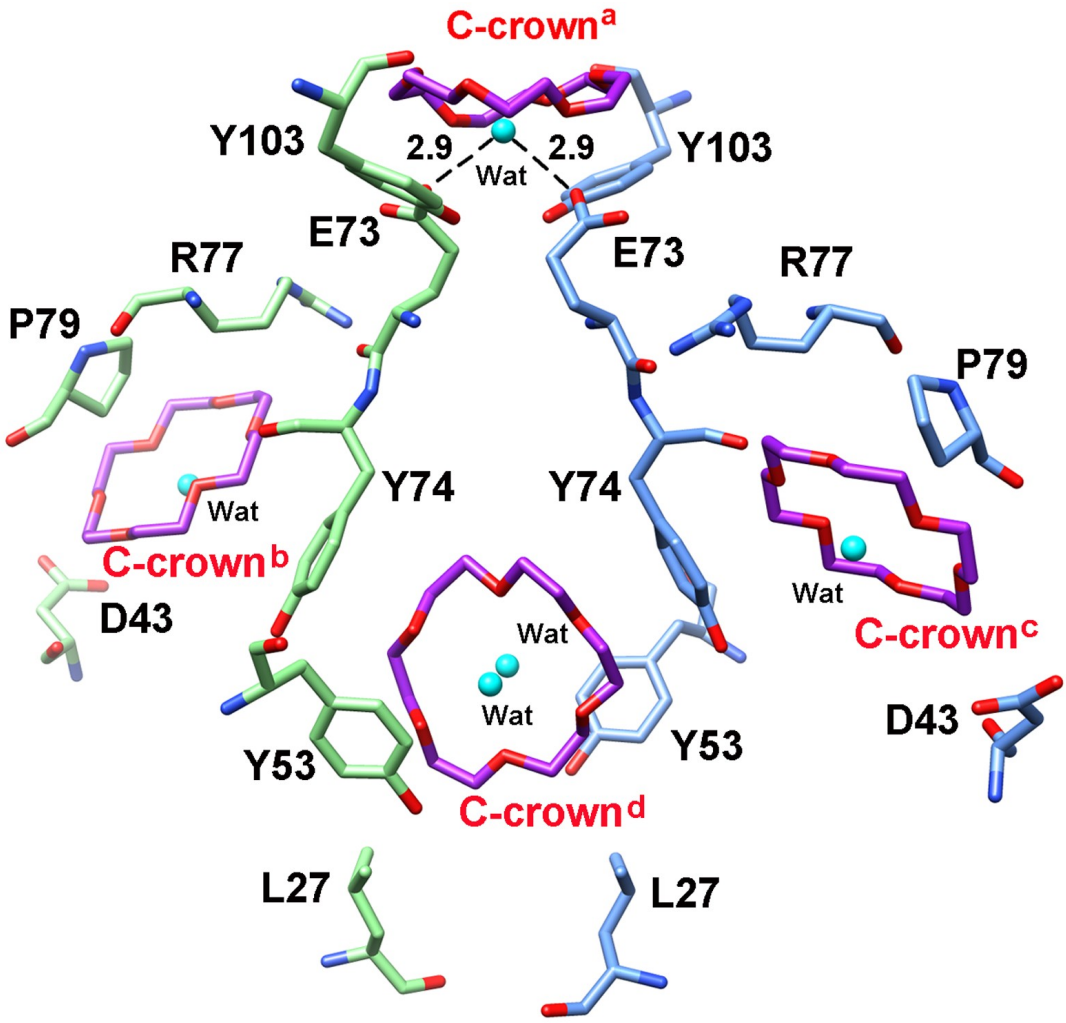

K-crown^a
K-crown^b*
K-crown^b

b)


α40K-crown
β59K-crown
α40K-crown
β59K-crown
α11KC-crown
α40K-crown
P37
β59K-crown
A62
K40
K59
α40K-crown
β59K-crown
α40K-crown
β59K-crown
α11KC-crown
K11
D74
S72
α11KC-crown
W15
V20

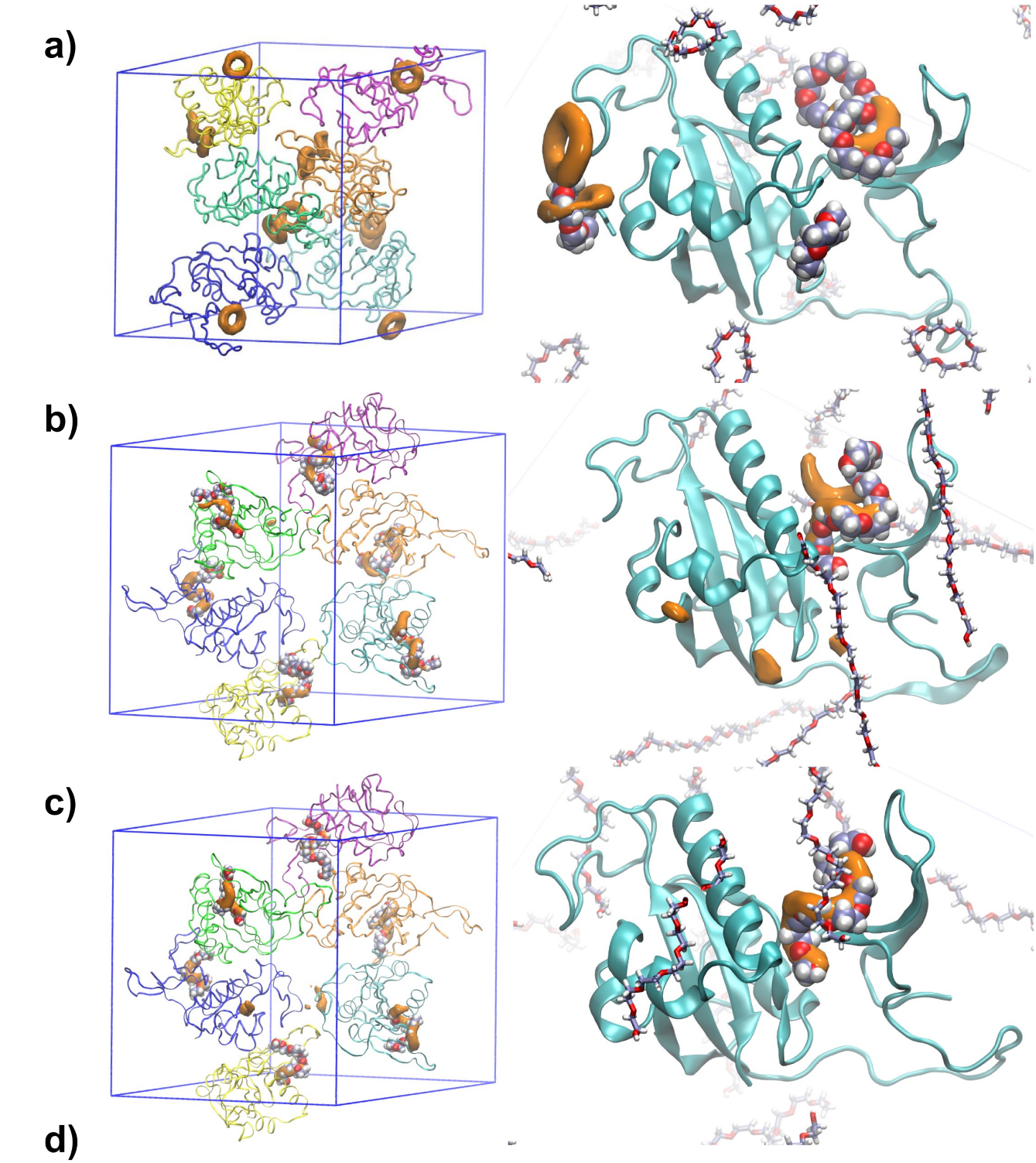
Supplementary figure 5. Additional CR co-crystal structures.

a) the unit cell of the DMP19 CR co-crystal. Each DMP19 dimer is presented in a different color; CRs are drawn in purple with red oxygen atoms. The two CR molecules interacting across different unit cells, K-crown^{b} and K-crown^{b}*, are shown in red. b) The co-crystal structure of the Hemoglobin-CR complex. Left: Five CRs bind to one hemoglobin molecule. The α and β subunits of hemoglobin are shown in gray and pink, respectively. Right: Crystal packing of hemoglobin and CRs. Four hemoglobin molecules are shown in different colors.



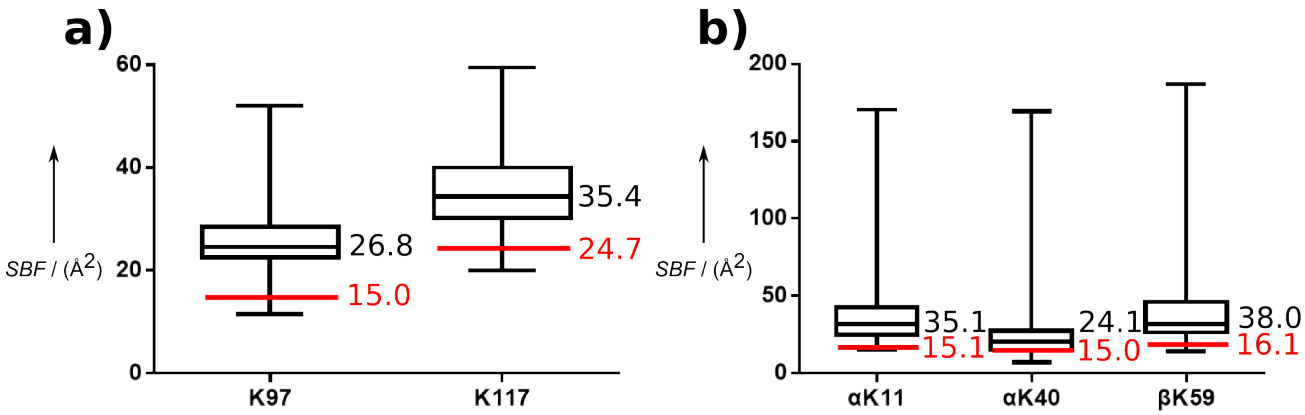
Supplementary figure 6. The complex C-crown binding network of the DMP19 dimer.

CRs are shown in purple, while elements of each of the DMP19 monomers are shown in blue and green. C-crown[a] and C-crown[d] bind to both DMP19 monomers within the dimer, resulting in a pseudo two-fold axis On the other hand, C-crown[b] and C-crown[d] interact each with a single DMP19 monomer in exactly the same fashion. C-crowns[b-c] are bound within a long, hydrophobic channel.


a)
b)
c)
d)
RMSD / Å
T / (ns)
Apo-Pin1R14A
Crown-ether
PEG400
PEG500
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
0
10
20
30
40
50
60
70
80
90
100

Supplementary figure 7. Pin1 molecular dynamics experiments.

Pin1 crystal structures were submitted to temperature and pressure controlled 100 ns MD simulations, either as crystalline unit cells (left), or in solution (right). Additive concentrations (lmwPEG or 18-crown-6 ether) were adjusted to correspond to the experimental concentrations. Crystallographic additive molecules are shown as spheres, while simulated ones appear as Dreiding models. Occupancies for all additive molecules were calculated (orange blobs). These are a direct measure for rigidity and preferred conformations in the simulation. a) Simulations in the presence of 18-crown-6 ethers (correspond to videos S1 and S2). b) Simulations in the presence of PEG500 (correspond to videos S3 and S4). c) Simulations in the presence of PEG400 (correspond to videos S5 and S6). d) RMSD over time for each of the crystal simulations, including an additive-free, control simulation (orange, video S7), with a crown ether (blue), a PEG400 complex (green), and a PEG500 complex (yellow). Crown ethers appear to result in a slightly more stable trajectory.



Supplementary figure 8. Comparison of side-chain B-factors of crown ether interacting lysines.

The side-chain B-factors (SBF) of lysine interacting with 18-crown-6 ethers in a) Pin1R14A and b) hemoglobin, in red, against the average SBF of those same lysines from all corresponding structures deposited at the PDB (black line). B-factor distributions are represented as box-and-whisker plots. The bottom and top of the box are the first and third quartiles, respectively. The band inside the box represents the distribution average. Bottom and top ends of the whiskers are the minimum and maximum, respectively.

References

[1] W. Delano, PyMOL, Molecular Graphics System, Schroedinger, Inc., Portland (USA) **2008**.

[2] H.-C. Wang, T.-P. Ko, M.-L. Wu, S.-C. Ku, H.-J. Wu, A. H.-J. Wang, *Nucleic Acids Res.* **2012**, *40*, 5718–30.

[3] C.-C. Lee, C.-J. Kuo, T.-P. Ko, M.-F. Hsu, Y.-C. Tsui, S.-C. Chang, S. Yang, S.-J. Chen, H.-C. Chen, M.-C. Hsu, et al., *J. Biol. Chem.* **2009**, *284*, 7646–55.

[4] M. Maestre-Reyna, W.-J. Wu, A. H.-J. Wang, *PLoS One* **2013**, *8*, e82458.

[5] Y. Zhang, S. Daum, D. Wildemann, X. Z. Zhou, M. A. Verdecia, M. E. Bowman, C. Lücke, T. Hunter, K.-P. Lu, G. Fischer, et al., *ACS Chem. Biol.* **2007**, *2*, 320–8.

[6] Z. Otwinowski, W. Minor, *Methods Enzym.* **1997**, *276*, pp 307–326.

[7] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, et al., *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 235–42.

[8] P. D. Adams, P. V Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 213–21.

[9] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D. Biol. Crystallogr.* **2011**, *67*, 355–67.

[10] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 486–501.

[11] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–12.

[12] R. a. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **1993**, *26*, 283–291.

[13] S. Boivin, S. Kozak, R. Meijers, *Protein Expr. Purif.* **2013**, *91*, 192–206.

[14] D. A. Case, T. A. Darden, I. T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, et al., **2012**.

[15] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, R. C. Walker, *J Chem Theory Com* **2012**, *8*, 1542–1555.

[16] W. Humphrey, A. Dalke, K. Schulten, *J Mol Graph* **1996**, *14*, 33–38.

[17] D. R. Roe, T. E. Cheatham III, *J Chem Theory Com* **2013**, *9*, 3084–3095.

[18] R. Anandakrishnan, B. Aguilar, A. V Onufriev, *Nucleic Acids Res.* **2012**, *40*, W537–41.

[19] D. S. Cerutti, I. Le Trong, R. E. Stenkamp, T. P. Lybrand, *Biochemistry* **2008**, *47*, 12065–77.

[20] D. S. Cerutti, P. L. Freddolino, R. E. Duke, D. A. Case, *J. Phys. Chem. B* **2010**, *114*, 12811–24.

[21] S. Chatterjee, P. G. Debenedetti, F. H. Stillinger, R. M. Lynden-Bell, *J. Chem. Phys.* **2008**, *128*, 124511.

[22] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.-US* **1987**, *91*, 6269–6271.

[23] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F.-Y. Dupradeau, *Nucleic Acids Res.* **2011**, *39*, W511–7.

[24] C. J. M. Huige, C. Altona, *J Comput Chem* **1995**, *16*, 56–79.

[25] J.-P. Ryckaert, G. Ciccotti, H. J. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.

[26] R. J. Loncharich, B. R. Brooks, R. W. Pastor, *Biopolymers* **1992**, *32*, 523–35.

[27] N. J. Christensen, K. P. Kepp, *PLoS One* **2013**, *8*, e61985.

[28] R. Page, R. C. Stevens, *Methods* **2004**, *34*, 373–89.

[29] G. G. Xu, Y. Zhang, A. Y. Mercedes-Camacho, F. A. Etzkorn, *Biochemistry* **2011**, *50*, 9545–50.