Article

Facile Fabrication of Protein−Macrocycle Frameworks

Kiefer O. Ramberg, Sylvain Engilberge, Tomasz Skorek, and Peter B. Crowley*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Precisely defined protein aggregates, as exemplified by crystals, have applications in functional materials. Consequently, engineered protein assembly is a rapidly growing field. Anionic calix[n]arenes are useful scaffolds that can mold to cationic proteins and induce oligomerization and assembly. Here, we describe protein-calixarene composites obtained via cocrystallization of commercially available sulfonato-calix[8]arene (**sclx$_8$**) with the symmetric and "neutral" protein RSL. Cocrystallization occurred across a wide range of conditions and protein charge states, from pH 2.2− 9.5, resulting in three crystal forms. Cationization of the protein surface at pH $\sim$ 4 drives calixarene complexation and yielded two types of porous frameworks with pore diameters >3 nm. Both types of framework provide evidence of protein encapsulation by the calixarene. Calixarene-masked proteins act as nodes within the frameworks, displaying octahedral-type coordination in one case. The other framework formed millimeter-scale crystals within hours, without the need for precipitants or specialized equipment. NMR experiments revealed macrocycle-modulated side chain pK_a values and suggested a mechanism for pH-triggered assembly. The same low pH framework was generated at high pH with a permanently cationic arginine-enriched RSL variant. Finally, in addition to protein framework fabrication, **sclx$_8$** enables *de novo* structure determination.
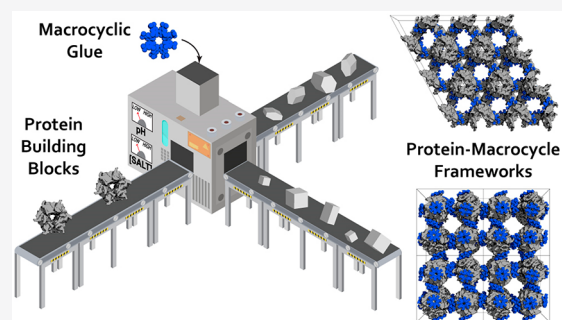
INTRODUCTION

Protein-based materials have great potential to serve society.[1−4] With their periodic arrangement of functional building blocks, crystals have applications in catalytic devices.[3−7] Porous crystals are of particular interest considering their capacity to capture (store) and transform biomolecules.[6,8−14] While great advances have been achieved with metal organic frameworks (MOFs)[12] and covalent organic frameworks (COFs),[13] protein-based frameworks have proved more challenging.[11,14−17] Yet, biocompatible and biodegradable frameworks are highly desirable given the demands for new therapeutics and biomaterials as well as sustainable manufacturing processes.[2] This paper describes the facile fabrication of millimeter-scale, high-porosity, solid-state composites of precisely arrayed protein and synthetic components.

Designed protein oligomerization and protein crystal engineering are progressing toward the goal of protein-based devices.[9−11,14−28] The application of Coulombic forces for guided assembly continues to deliver satisfactory results such as the cocrystallization of binary mixtures of oppositely charged homologues.[25−27] Multivalent ligands and "molecular glues" offer alternative approaches to controlled assembly without the requirement for engineered surface features in the target protein.[3,18−23] For example, anionic calix[n]arenes that host arginine or lysine side chains[29−32] can direct the assembly of cationic proteins.[20,33−37] The commercially available sulfonato-calix[8]arene (**sclx$_8$**), a 1.5 kDa flexible phenolic macrocycle with variable cavity dimensions,[29] a solvent accessible surface area (ASA) of $\sim$1600 Å^2, and a formal net charge ranging from −8 to −12,[38] shows particular promise.[35−37] We have demonstrated autoregulated oligomerization of the lysine-rich cytochrome *c* (cyt*c*, isoelectric point, p*I* $\sim$ 9.5), with 1 equiv **sclx$_8$** forming a tetramer, and 3 equiv yielding a calixarene-coated (encapsulated) protein.[35] The ligand:protein ratio influences also the formation of cyt*c*-**sclx$_8$** crystalline frameworks with varying porosities, ranging from 65 to 85% solvent content.[35,37] The most porous framework is mediated exclusively by protein-calixarene contacts and requires at least 3 equiv **sclx$_8$** with respect to the cationic protein. Crystal engineering and the use of effector ligands have provided access to different architectures including a duplicated framework.[37] Here, we describe the preparation of crystalline frameworks comprising a symmetric "neutral" protein and **sclx$_8$**.

The 6-bladed β-propeller *Ralstonia solanacearum* lectin (RSL, p*I* $\sim$ 6.5) was selected as the model protein.[39−43] RSL is a rigid, C_3-symmetric spheroid with high thermal stability making it an interesting candidate for protein-based frame-

Table 1. Cocrystallization Conditions and Structure Properties

Form	Protein	equiv sclx$_8$	(NH$_4$)$_2$SO$_4$ (M)	Buffer	pH	Additive (0.2 M)	Space Group	a x b x c (Å)	Res (Å)	PDB id
I	RSL	80	1.6	CAPS	9.5	Li$_2$SO$_4$	P2$_1$3	64^3	1.2	6Z60
	RSL	80	1.6	Tris-HCl	8.5	Li$_2$SO$_4$			1.3	6Z62
	RSL-R$_6$	50	1.6	Tris-HCl	8.5	Li$_2$SO$_4$		^{3}	1.1	6Z5Z
	RSL	50	2.4	MES	6.8				1.1	6Z5W
	RSL	80	2.0	acetate	4.8				1.1	6Z5X
II	RSL	50	0.8	citrate	4.0		I23	104^3	1.3	6Z5G
	RSL	10	0.8	Gly-HCl	2.2	MgCl$_2$			1.6	6Z5M
III	RSL	10		acetate	4.0		P3	60 × 60 × 64	1.3	6Z5Q
	RSL*	10		acetate	4.0			60 × 60 × 64	1.3	7ALF
	RSLex	15		acetate	4.0			60 × 60 × 64	1.5	7ALG
	RSL-R$_8$	50	1.3	Tris-HCl	8.5	Li$_2$SO$_4$		60^3	1.4	6Z5P

works.[43] Trimeric RSL possesses pseudo C_6-symmetry due to the ∼40% sequence identity between its N- and C-terminal halves. Such high symmetry is advantageous for framework fabrication as evidenced with other lectins,[19,23,44] related β-propellers,[45] ferritin,[15,16,24,25] viral capsids,[8,46] and engineered cages.[6,28,47−49] sclx$_8$-mediated assembly was tested with RSL and several variants including RSLex[43] and MK-RSL (each containing one extra lysine), and two arginine-enriched mutants, RSL-R$_6$ and RSL-R$_8$ (Figures S1 and S2). In RSL-R$_6$, the three lysines of native RSL are replaced by arginine.[42] RSL-R$_8$ includes these mutations as well as two acidic residues replaced by arginine. The chemically modified variants, methylated RSL (RSL*) and acetylated RSL (RSL-Ac), were tested also.

We present three types of RSL-sclx$_8$ frameworks dependent on the protein charge characteristics (pH trigger) and the cocrystallization conditions. Two of the frameworks require acidic conditions and are porous with >55% solvent content and pore diameters >3 nm. These frameworks are consistent with protein encapsulation[50] by calix[8]arene in solution[35] and suggest a molecular basis for reentrant condensation.[51,52] The low pH framework was recapitulated with the highly cationic RSL-R$_8$ variant. NMR experiments provide further evidence of a pH trigger, arising from protonation of acidic side chains in RSL. One of the fabrication processes is rapid (hours), yields millimeter-scale crystals, and requires neither precipitants nor specialized equipment. Thus, we demonstrate the general utility of sclx$_8$ for protein framework assembly as well as X-ray structure determination by anomalous methods.

■ EXPERIMENTAL SECTION

Materials. Stock solutions of sclx$_8$ (Tokyo Chemical Industry) were prepared in water, and the pH was adjusted to 8.0.

Protein Production. Unlabeled and ^{15}N-labeled RSL samples were produced in *E. coli* BL21 transformed with the pET25rsl plasmid. The modified pET25rsl vectors that encode RSL-R$_6$ and RSLex were reported previously.[42,43] The vectors encoding RSL-R$_8$ (K25R/K34R/E43R/D46R/K83R) and MK-RSL were produced by Genscript. All proteins, except RSL-R$_8$, were purified by mannose-affinity chromatography.[39] Attempts to purify RSL-R$_8$ by affinity chromatography failed due to the coelution of *E. coli* proteins, likely as a consequence of arginine "stickiness".[53] Consequently, RSL-R$_8$ was purified on carboxymethyl resin equilibrated with 0.02 M potassium phosphate and 0.2 M NaCl at pH 6.0, and eluted with the same buffer plus 1 M NaCl. Methylation and acetylation of RSL were performed as described.[42,43] Mass analysis of RSL-R$_8$ and MK-RSL was performed with an Agilent 6460 Triple Quadrupole LC/MS (Figure

S3, Table S1). Protein concentrations were determined spectrophotometrically with ε_{280} = 44.46 mM^{-1}cm^{-1} for the monomer.

Cocrystallization Trials. All experiments were performed with D-fructose bound RSL and variants. Protein-sclx$_8$ cocrystals were obtained at 20 °C by using commercial (JCSG++ HTS, Jena Bioscience) or homemade screens, applied with an Oryx 8 robot (Douglas Instruments). Generally, the crystals were reproduced manually by hanging-drop vapor diffusion in 24 well Greiner plates. Protein concentrations ranged from 0.8 to 1.8 mM. Crystals were obtained in JCSG++ HTS conditions **B1** (0.8 M ammonium sulfate and 0.1 M sodium citrate pH 4.0), **C11** (2.0 M ammonium sulfate and 0.1 M sodium acetate pH 4.6), **E2** (2.0 M ammonium sulfate, 0.2 M sodium chloride and 0.1 M MES pH 6.5), and **G11** (2.0 M ammonium sulfate and 0.1 M BIS-TRIS pH 5.5). Homemade screens included 0.8−2.4 M (NH$_4$)$_2$SO$_4$, 0.1 M buffer, and 0 or 0.2 M NaCl, (Li)$_2$SO$_4$ or MgCl$_2$. The buffers tested (pH values indicated in parentheses, not corrected for the presence of salts) were glycine-HCl (2.2), citrate (4.0), acetate (4.6), MES (6.8), Tris-HCl (8.5), or CAPS (9.5). Typically, screens included 0, 1, 2, 4, 8, 16, 32, or 64 mM sclx$_8$. In the case of RSL-R$_8$, the Jena screen was tested at 50 and 100 mM sclx$_8$. Crystals were obtained also by incubation of protein−ligand mixture in microcentrifuge tubes. In this simplified batch crystallization, protein-sclx$_8$ mixtures were prepared in 20 mM acetate or phosphate, 50 mM NaCl, pH adjusted to 4.0, and incubated at 4 °C. Crystallization drops were imaged using an Olympus SZX16 stereomicroscope and a Olympus DP25 digital camera.

X-ray Data Collection, Processing, and Model Building. Crystals were cryo-protected in the crystallization solution supplemented with 20−25% glycerol and cryo-cooled in liquid nitrogen. Diffraction data were collected at beamline PROXIMA-2A, SOLEIL synchrotron (Saint-Aubin, France) with an Eiger X 9M detector (Tables S2−S4). Data were processed using the autoPROC pipeline.[54] Data were integrated in XDS[55] and the integrated intensities were scaled and merged in AIMLESS[56] and POINTLESS[57] in CCP4 and assessed for pathologies in phenix.Xtriage.[58] Structures were solved by molecular replacement in PHASER,[59] using the RSL monomer (derived from PDB 2BT9[39]) as a search model. For *de novo* phasing experiments data were collected at beamline X06DA, Swiss Light Source (Villigen, Switzerland). To maximize anomalous scattering from sulfur atoms, diffraction data were collected at a wavelength of 2.07 Å (Table S3). A single data set collected at this energy sufficed for *de novo* structure determination.[60] Diffraction frames (deposited on Zenodo, DOI: 10.5281/zenodo.3944486) were integrated using XDS and scaled with AIMLESS and POINTLESS. Substructure determination, phasing, and model building were performed in SHELX.[61] A second high-resolution data set used for refinement was collected on the same crystal at a wavelength of 0.97 Å and processed using the autoPROC pipeline (Table S3). The coordinates for sclx$_8$ (PDB id EVB) and D-fructose (PDB id BDF) were added to each model. Iterative cycles of model building in COOT[62] and refinement in phenix.refine[58] were performed until no

further improvements in the R_{free} or electron density were obtained. All of the structures were validated in MolProbity.[63] High resolution refined coordinates and structure factors were deposited in the Protein Data Bank (Tables 1 and S2−S4). Protein−ligand interface areas were calculated in PDBe PISA.[64] Crystal porosity was analyzed in MAP_CHANNELS.[65] The molar protein concentration, [P], in each crystal form was calculated by

$$[P] = N/(N_A \times V \times 10^{-27})$$

where N is the number of molecules of P in the unit cell, N_A is Avogadro's number, and V is the unit cell volume (Å³).

NMR Characterization. Samples typically comprised 0.1−1.0 mM ^{15}N-labeled protein in 20 mM phosphate buffer, 50 mM NaCl, 5 mM D-fructose, and 10% D_2O. Samples in 20 mM acetate buffer (instead of phosphate) were tested also. Ligand titrations were performed with μL aliquot additions of 0.1 M **sclx₈**. 2D ^{1}H−^{15}N HSQC watergate spectra were acquired at 30 °C with 4 or 8 scans and 64 increments on a Varian 600 MHz spectrometer with a HCN cold probe. Data processing and analysis were performed in NMRpipe[66] and CCPN,[67] respectively. Binding isotherms were obtained by plotting the chemical shift perturbation (Δδ) as a function of the **sclx₈** concentration. Nonlinear least-squares fits to a one-site binding model were performed, with Δδ and [**sclx₈**] as the dependent and independent variables, respectively, and the dissociation constant (K_d) and maximum chemical shift change ($\Delta\delta_{max}$) as the fit parameters. pH titration curves were generated from ^{1}H−^{15}N HSQC spectra of RSL, in the presence of 0 or 5 mM **sclx₈**, pH adjusted in increments of 0.2 pH units. The pH was measured before and after each HSQC data acquisition. The pH dependence of the chemical shifts was analyzed using nonlinear least-squares fits of the data to the modified Henderson−Hasselbalch equation

$$\delta_{obs} = \delta_{low} - \frac{\delta_{low} - \delta_{high}}{1 + 10^{n(pK_a - pH)}}$$

where pK_a is the ionization constant, δ_{low} and δ_{high} are the low and high pH chemical shift plateaus, and n is the apparent Hill coefficient.[68]

Empirical pK_a Calculations. The pK_a values of the acidic residues measured by NMR were compared with empirical estimates obtained by using PROPKA3.2.[69] In this version of the pK_a predictor, the parameter sets account for noncovalent interactions with ligand groups. The PROPKA3.2 calculations were performed with the coordinates of the $P3$ RSL-**sclx₈** crystal structure (PDB 6ZQ5) with and without the **sclx₈** coordinates included in the model.

RESULTS

RSL-sclx₈ Cocrystal Forms. Despite its "neutral" character, RSL cocrystallized with the highly anionic **sclx₈** under a broad variety of conditions (Table 1 and Figure S4). Data collection at SOLEIL synchrotron revealed three distinct crystal forms with high-quality diffraction properties (Tables 1, 2, and S2−S4).

Sitting-drop vapor-diffusion experiments with a commercial screen yielded RSL-**sclx₈** cocrystals in 2.0 M ammonium

Table 2. RSL-sclx₈ Cocrystal Forms and Properties

Form	Space Group	RSL:sclx₈[a]	[P] (mM)[b]	S.C. (%)[c]	Pore Ø (nm)[d]
I	$P2_13$	1:1	76	36	1.7
II	$I23$	1:2	36	66	4.2
III	$P3$	1:1	50	59	2.8

[a]Protein:ligand ratio per RSL monomer. [b]Calculated protein concentration based on unit cell contents. [c]Solvent content estimated from total mass (protein plus **sclx₈**). [d]Diameter of widest pore, calculated in MAP_CHANNELS.[65]

sulfate with different buffers (see Experimental Section). This crystal form (I) grew at 40−80 equiv **sclx₈** over 7−10 days. Homemade screens extended these results to 1.6−2.4 M ammonium sulfate and buffers ranging from acetate pH 4.8 to CAPS pH 9.5 (Table 1). The growth of crystal form I was independent of the pH, occurring in acidic or basic conditions where the protein is cationic or anionic, respectively. This observation, together with the high ammonium sulfate concentration (Debye screening), suggests that attractive charge−charge interactions have a minor contribution to this cocrystallization process. While **sclx₈** is always anionic, RSL is cationic or anionic at low or high pH, respectively. Therefore, attractive charge−charge interactions between the protein and calixarene occur only at pH ≤ 5.

Crystal form I was solved with an asymmetric unit comprising one RSL monomer and one **sclx₈** in the cubic space group $P2_13$. An essentially identical structure was obtained with RSL-R₆. This tightly packed crystal form with 36% solvent content (Figure 1 and Table 2) involves two protein−protein crystal contacts and five protein-calixarene interfaces. The protein−protein interfaces bury 100 or 140 Å² per molecule and are typical crystal contacts. The protein-**sclx₈** interfaces bury ∼65% of the calixarene, while the remainder is solvent exposed. **sclx₈** adopts a highly puckered conformation and molds neatly to one RSL monomer (Figures 1B and 2A), burying 585 Å² and entrapping adjacent residues Val13 and Lys34 in niches formed by two and three proximal phenolic units, respectively. Another calixarene cavity interacts *exo* to RSL and masks several residues, including Asp32 (*vide infra*) and Tyr37, adjacent to the sugar binding site. The other side of the **sclx₈** forms an interface that is ∼3-fold smaller (180 Å²) and binds Lys83 and Ala85 of a second RSL monomer. Calixarene complexation of two aliphatic side chains (Val13 and Ala85) via CH-π bonds suggests that the hydrophobic effect is important to this crystal form that grows in high salt conditions.
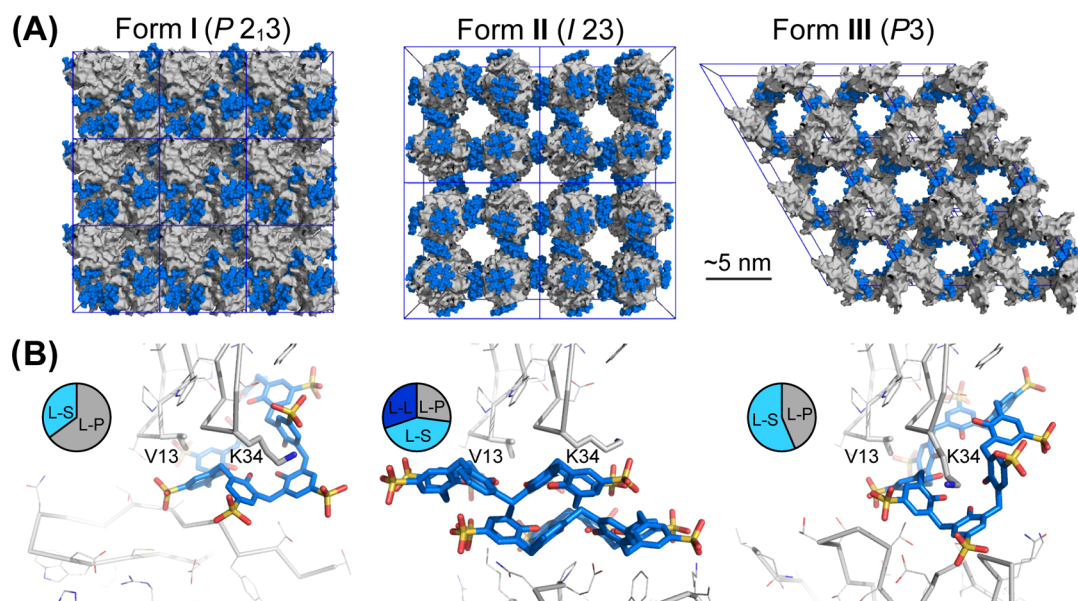
While the hydrophobic effect appears to be important in crystal form I, charge−charge interactions are likely to be inconsequential. Support for this argument arises from the growth of these crystals in both acidic and alkaline conditions (*vide supra*) and from the disorder at the Lys34 site (Figure 2). While the electron density maps clearly indicate an extended conformation for Lys83 across the pH range, Lys34 was modeled in two conformations with high temperature factors[70] (indicative of flexibility/motion) even in the pH 4.8 structure. The encapsulating portion of **sclx₈** has correspondingly high temperature factors (∼30 Å²) compared to the rest of the ligand (∼10 Å²). At pH ≥ 8.5, the Lys34 side chain is completely disordered beyond C^β and one monomer of **sclx₈** flicks between two opposing conformations (Figure 2B). These results suggest that the Lys34 side chain is deprotonated and does not engage **sclx₈** via the typical salt bridge interactions.[33−37] The replacement of Lys34 with Arg34 in RSL-R₆ yields a better-defined interface (Figure 2A). The high pK_a of Arg[71] ensures protonation and salt bridge formation with the sulfonic acids. Presumably, the increased bulk of Arg relative to Lys also contributes to stabilize the interface. Another aspect of the disorder around **sclx₈** concerns the N-terminal Ser1. This protonated residue is disordered despite its location at ∼5 Å from the nearest sulfonic acid, suggesting again that charge− charge interactions are minor in crystal form I.

The commercial screen also yielded crystals, form II, in 0.8 M ammonium sulfate and citrate pH 4.0. In this case, 10 equiv

Figure 1. (A) Crystal packing in RSL-**sclx₈** cocrystal forms I ($P2_13$), II ($I23$), and III ($P3$). Note the high porosity of the $I23$ and $P3$ forms, with nanometer-scale solvent channels. Proteins shown as gray surfaces, and **sclx₈** shown as blue spheres. (B) Details of the principal protein-**sclx₈**-protein interfaces in each crystal form, with RSL shown as the monomer for clarity. The Val13 and Lys34 binding patch is common to each crystal form. Pie charts show area proportions of **sclx₈**-mediated interfaces (see Table S4). L, ligand; P, protein; S, solvent.
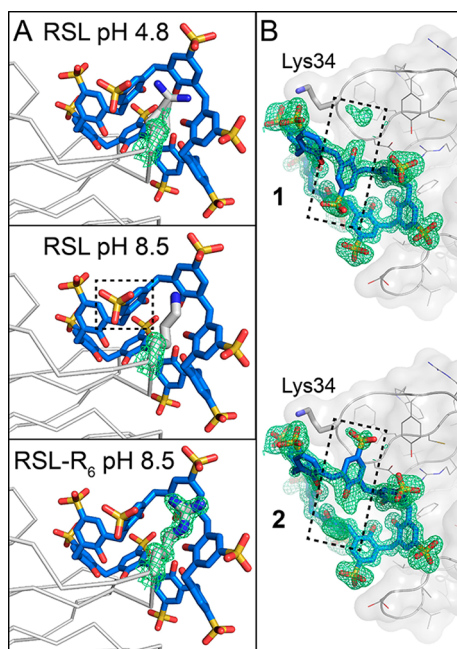

Figure 2. (A) Disorder at the Lys34-**sclx₈** site increases with increasing pH in the $P2_13$ form. In contrast, Arg34 is fully defined in the RSL-R_6 variant. (B) Electron density in the RSL-**sclx₈** crystal structure at pH 8.5 is consistent with alternate conformations for one calixarene monomer, highlighted with a dashed box. Alternate conformers 1 and 2 of **sclx₈** were modeled at 55 and 45% occupancy, respectively. 2Fo-Fc electron density maps (contoured at 1.0 σ) are shown as green mesh.

sclx₈ was sufficient and the crystals appeared within 4−5 days (Table 1). Homemade screens including acetate pH 4.0 or glycine-HCl pH 2.2 also yielded crystal form II. The requirement for acidic conditions, in which the protein is cationic, suggests that attractive charge−charge interactions are important for cocrystallization of form II. The structure was

solved in the cubic space group $I23$ with an asymmetric unit that contains two molecules of **sclx₈** per monomer of RSL (Table 2). The unit cell is a remarkable assembly of eight RSL trimers connected by **sclx₈** dimers (Figure 1). Here, **sclx₈** occurs in the fully extended, pleated loop conformation (ASA 1600 Å²) that presents four shallow cavities on each face. The dimer, composed of two identical **sclx₈**, is a staggered assembly that buries 480 Å² per calixarene and involves multiple CH−π, OH−π, π−π, and anion−π interactions (Figures 3 and S5). Two of the sulfonic acid substituents are encapsulated in shallow cavities of the partner calixarene. This highly anionic species (formal charge exceeding −16) dominates the crystal packing which is a porous assembly devoid of any protein−protein crystal contacts (66% solvent content, pore diameter ~4 nm, Table 2). Two distinct patches of RSL bind to either side of the **sclx₈** dimer. The larger protein-**sclx₈** interface (500 Å²) entraps Asn42, Pro44, and Trp74. Each side chain makes van der Waals contact with two phenolic rings and one methylene of **sclx₈**. This cluster of residues is flanked by Lys25 and Lys83, both of which form salt bridges to the calixarene. Glu43 (*vide infra*) is completely masked by the calixarene and forms hydrogen bonds to the lower rim phenols. The smaller protein-**sclx₈** interface (240 Å²) entraps Val13 and Lys34, with features similar to the principal interface in crystal form I ($P2_13$). In this case, the electron density was clear for Lys34 and both the side chain and the calixarene were modeled with low temperature factors. Charge interactions appear to be important for crystal form II as the calixarene is present as a dimer, all three lysines of RSL participate in **sclx₈** complexation, and the crystals grow only at pH ≤ 4.0.

The sulfur content of **sclx₈** was taken advantage of and the $I23$ structure was solved *de novo* by single-wavelength anomalous dispersion (SAD) of sulfur atoms.[60] The crystal architecture with two **sclx₈** per RSL monomer adds 16 sulfur atoms per asymmetric unit in addition to two cysteine residues (Figure 3). *De novo* phasing using the anomalous sulfur signal was straightforward and yielded a perfectly interpretable
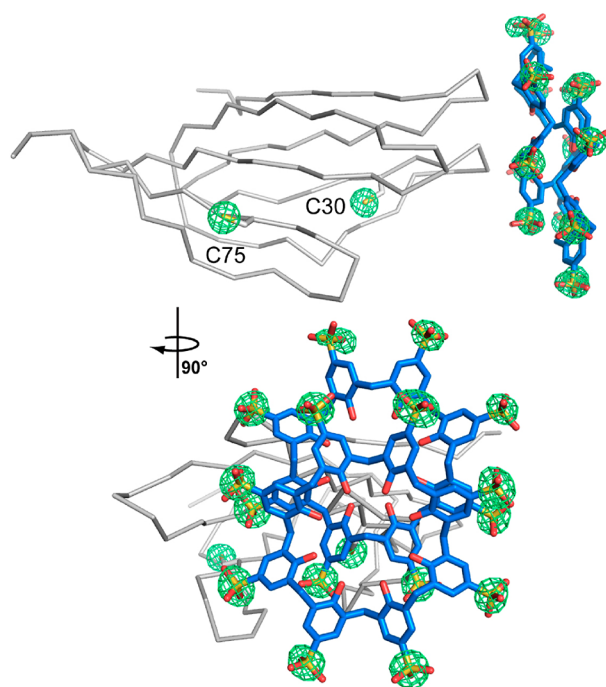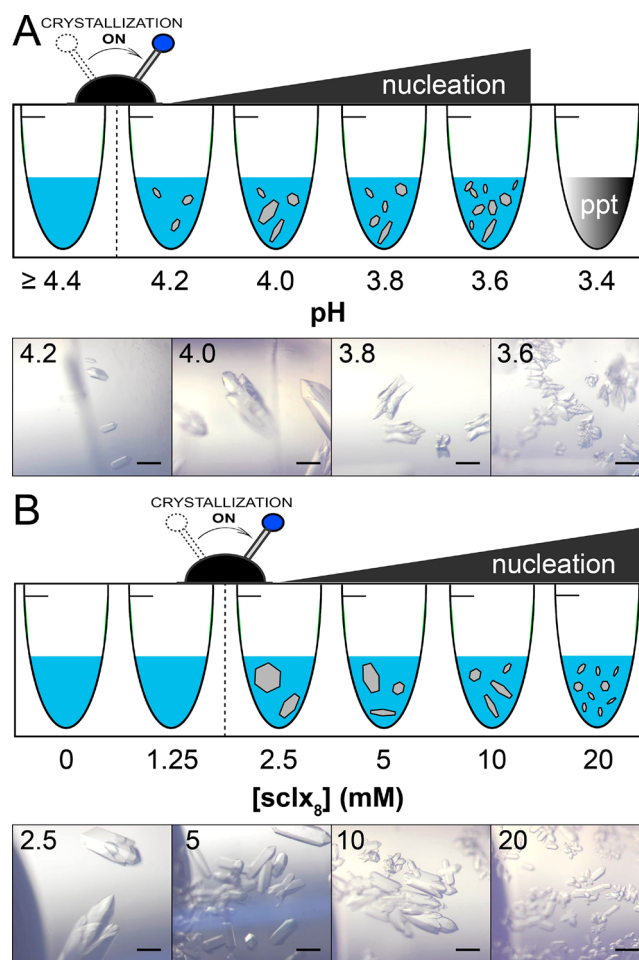
Figure 3. Fourier maps (green mesh) calculated from sulfur anomalous data and contoured at 4.0 σ in crystal form II (*I*23). The RSL monomer is shown as the C^α trace, the **sclx$_8$** dimer and cysteine side chains are shown as sticks.

Figure 4. RSL-**sclx$_8$** cocrystallization in the absence of precipitant at 4 °C. (A) pH dependence of crystal growth at 1 mM RSL, 10 mM **sclx$_8$** in 20 mM phosphate buffer, and 50 mM NaCl. Cocrystallization was triggered at pH ≤ 4.2. Microcrystalline precipitate occurred at pH 3.4 (Figure S5). (B) Cocrystallization dependence on the **sclx$_8$** concentration at pH 4.0. Cocrystallization was triggered at **sclx$_8$** ≥ 2.5 mM. Microscope images were acquired after 3 h incubation. Scale bars are 200 μm.

electron density map suited to completely automated model building (Table S3). These data demonstrate the utility of **sclx$_8$** for framework fabrication as well as crystal structure determination.

Crystal form III was discovered in the course of NMR experiments when overnight sample storage at 4 °C yielded crystals. Form III also requires acidic conditions, and grows rapidly (2−3 h) in the absence of precipitant. Protein−ligand mixtures in 20 mM phosphate or acetate buffer and 50 mM NaCl at pH 4.0, yielded crystals upon incubation at 4 °C (Figures 4, S6−S8). At 10 equiv **sclx$_8$**, cocrystallization is switched on at pH ≤ 4.2 (Figure 4A). Nucleation increased with decreasing pH and a fine microcrystalline precipitate formed at pH 3.4 (Figure S7). This precipitate was redissolved simply by raising the pH or by increasing the **sclx$_8$** concentration (Figure S8), consistent with reentrant condensation.[51,52] Compared to crystal forms I and II, the growth of form III was more sensitive to the protein:calixarene ratio. At pH 4.0, 2.5 equiv **sclx$_8$** was sufficient for cocrystallization. This form was distinguished also by its size, yielding millimeter-scale crystals overnight at low **sclx$_8$** concentrations (Figure 4). While the crystal growth was pH sensitive, crystals transferred to buffer at pH 7.4 were stable for 2−3 h. Above pH 7.4, the crystals dissolved within minutes. Considering the low pH and low ionic strength, charge−charge interactions dominate crystal form III, which was reproduced with RSL variants (RSL* and RSLex) under identical conditions, and with the highly cationic RSL-R$_8$ at pH 8.5 in ammonium sulfate (Table 1 and Figure S4).
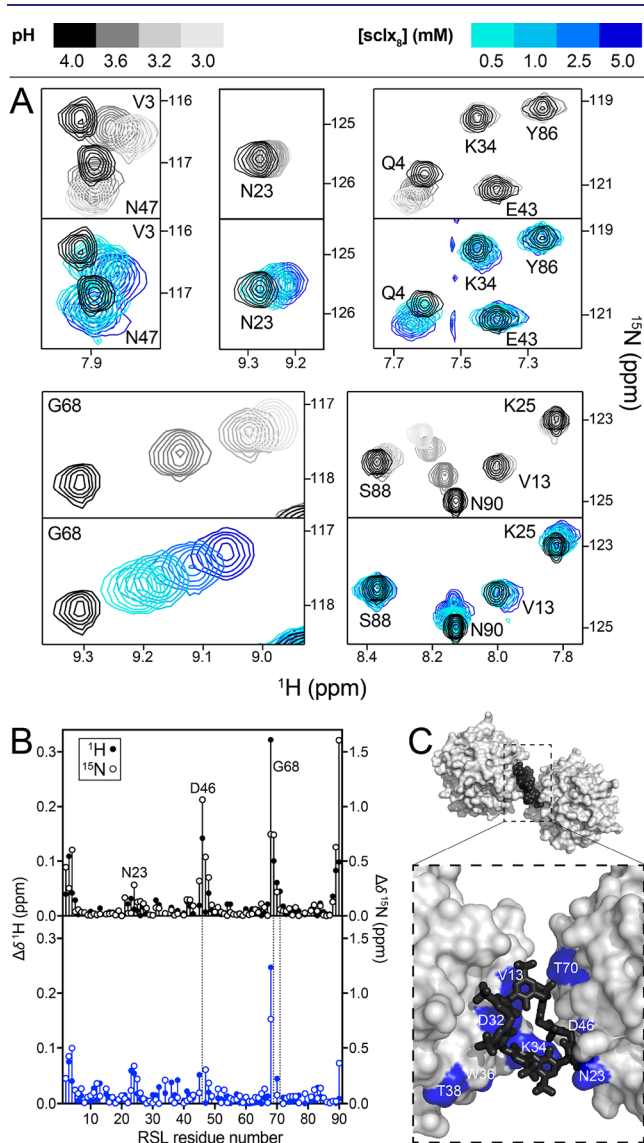
Crystal form III was solved in space group *P*3 with one **sclx$_8$** per RSL monomer. Similar to the *I*23 form, protein-calixarene contacts dominate the porous *P*3 assembly (59% solvent content, pore diameter ~3 nm, Table 2) and there are no protein−protein contacts (Figure 1A). **sclx$_8$** adopts a highly

puckered conformation, akin to that in form I (*P*2$_1$3), but it is substantially more solvent-exposed (Figure 1B and Table S5). Two protein-**sclx$_8$** interfaces are involved in this assembly. In the largest interface, 395 Å^2 of **sclx$_8$** is buried by the encapsulation of Val13 and Lys34. Surprisingly, **sclx$_8$** also accommodates Asp32 in a shallow cavity formed by two adjacent phenolic units. The Asp32 side chain is rotated slightly (relative to the *P*2$_1$3 and *I*23 forms) such that it no longer forms a salt bridge with Arg17, hinting at a change in its protonation state. The second interface buries 260 Å^2 of the calixarene and involves three short loops of the neighboring RSL monomer pressed up against the calixarene portion that encapsulates Lys34. These loops are capped by Asn23, Asp46, and Gly68. A pronounced conformational change of Asp46 in the central loop appears to facilitate the protein-calixarene interface. The Asp46 side chain rotates ~90° from the usual conformation (a hydrogen bond with the Gly24-C^α) to hydrogen bond with the amide NH of Gly68. This new conformation eliminates a potential steric clash with the bound **sclx$_8$** in crystal form III.

NMR Analysis of RSL-sclx$_8$ Complexation and the pH Trigger. The pH effects identified in the cocrystallization experiments were evident also from solution-state NMR experiments in 20 mM phosphate (or acetate) buffer and 50 mM NaCl. At pH 5.6, the ^{1}H−^{15}N HSQC spectrum of RSL was unchanged by titration with mM concentrations of sclx$_8$, suggesting that RSL and sclx$_8$ have negligible interactions under these conditions (Figure S9). However, at pH 4.0, specific RSL resonances were perturbed significantly by sclx$_8$, with saturation occurring at ∼10 equiv sclx$_8$ (Figures 5 and S9). Analysis of the chemical shift perturbations as a function of sclx$_8$ concentration yielded hyperbolic binding curves (Figure S10) with binding affinity in the mM range. Mapping the significantly affected resonances onto crystal form III ($P3$)

revealed binding patches consistent with the crystallographically defined protein-calixarene interfaces (Figure 5C). The occurrence of both binding sites (as per crystal form III) in the NMR experiments suggests transient sharing of bound sclx$_8$ between two RSL trimers. This inference is corroborated by line broadening evidence. The average ^{1}H^N line-widths of RSL increased by ∼8 Hz in the presence of sclx$_8$ (Table S6), consistent with complexation and assembly in solution.[34,35] Attempts to characterize RSL-sclx$_8$ at lower pH values were thwarted by precipitation at pH ≤ 3.4 and severe signal loss in the HSQC spectrum (Figure 6A). These observations agreed
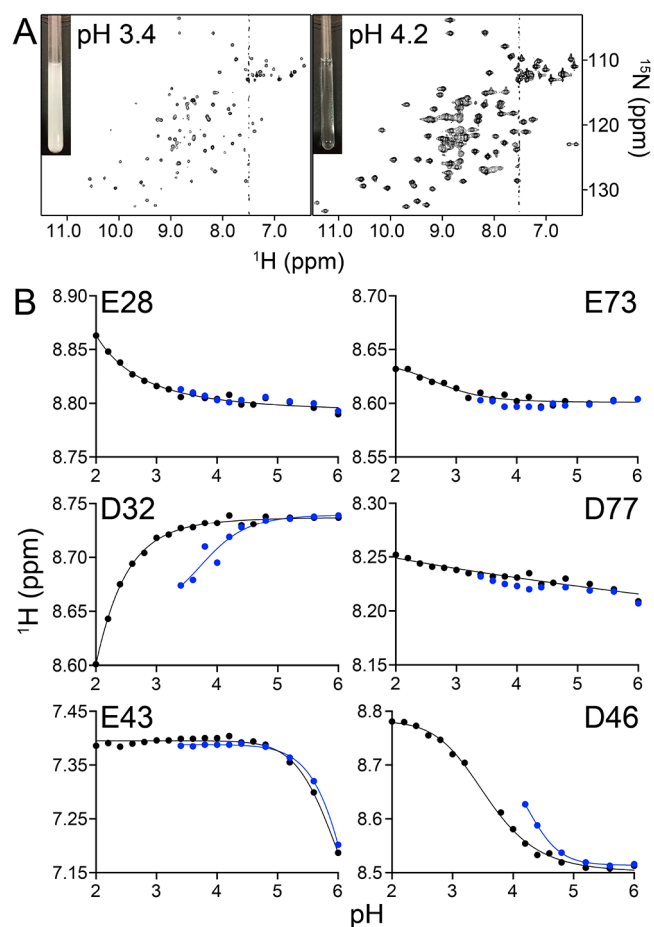

Figure 6. (A) RSL-sclx$_8$ mixtures precipitated at pH ≤ 3.4, precluding NMR data collection. Precipitation was pH-reversible. (B) pH titration curves for the 6 acidic residues of RSL obtained by monitoring the ^{1}H^N signal at 0 (black) or 5 mM (blue) sclx$_8$.

with the results of the cocrystallization experiments (Figure 4A). Precipitation in the NMR sample was fully reversible. Increasing the sample pH to 4.2 resulted in complete dissolution of the precipitate (within seconds) and concomitant restoration of the spectrum.

RSL-sclx$_8$ complexation at low pH is due to the increased cationic character of the protein. In the course of the NMR experiments, we observed a striking similarity between the chemical shift perturbations induced by sclx$_8$ and the perturbations produced by pH titration of the pure protein (Figure 5). The amide resonance of Gly68 shifted strongly, with the same direction and magnitude, due to a 1 pH unit change or to a 10-fold increase in sclx$_8$ concentration. Similar effects occurred for N-terminal residues (Val3, Gln4) and

Figure 5. (A) Regions from overlaid ^{1}H−^{15}N HSQC spectra of RSL during pH (gray scale) or sclx$_8$ (color scale) titrations. (B) Chemical shift perturbation plots of RSL backbone amides in response to pH adjustment (pH 4.0−3.0) or to the addition of 5 mM sclx$_8$ (at pH 4.0). Dashed lines indicate unassigned resonances due to overlap. (C) Detail of the RSL-sclx$_8$ $P3$ assembly (Figure 2) showing RSL trimers in surface representation and bridging sclx$_8$ as black spheres or sticks. Significant chemical shift perturbations at 5 mM sclx$_8$ are highlighted blue.

Asn47. Both Asn47 and Gly68 are reporters on Asp46, the resonance of which occurs in a crowded region of the spectrum. Other resonances such as Val13, Lys25, and Lys34, implicated in $sclx_8$ binding, were affected by the addition of $sclx_8$ but not by changes in pH. Yet others were strongly affected by pH (e.g., C-terminal Asn90) but less so by $sclx_8$. These effects were independent of the buffer, either phosphate or acetate, and suggest that calixarene binding involves protonation of RSL. Asp32 and Asp46 are the obvious candidates considering that they contribute to the protein-calixarene-protein interfaces in crystal form III (P3). Asp32 is buried completely by the calixarene while the solvent accessibility of Asp46 is unaffected, and both side chains undergo significant conformational changes (Figures 5C and 7B). In particular, the $sclx_8$-directed assembly appears to be contingent on a structural rearrangement of Asp46, which flips to form a new hydrogen bond with Gly68.
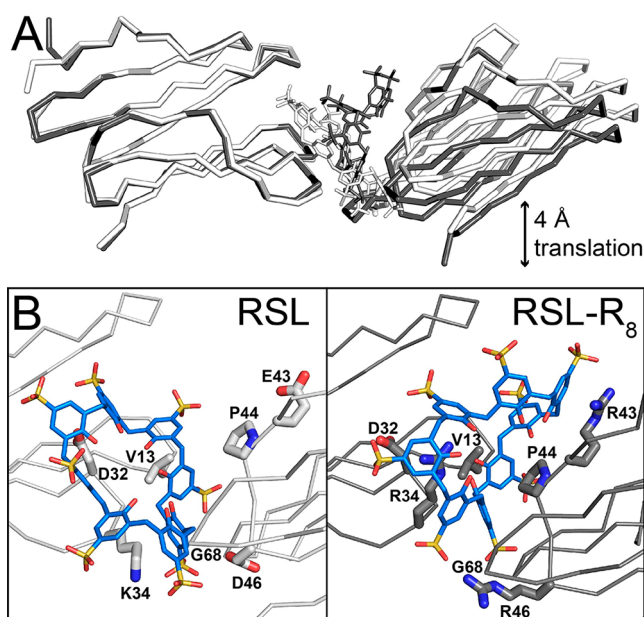
Figure 7. (A) Structural alignment (via chain A) of the main protein-$sclx_8$-protein assembly in the P3 crystal forms of RSL (white) and RSL-R_8 (gray). One of the proteins is translated by ~4 Å in the RSL-R_8 structure. (B) Detail of the protein-$sclx_8$-protein interfaces with key side chains shown as sticks.

Further evidence for a pH trigger for protein-calixarene assembly was provided by pK_a analysis of the acidic side chains in the absence and presence of $sclx_8$ (Figures 6 and S11). The NMR-derived pK_a values for the acidic residues of RSL were in good agreement with the values calculated in PROPKA3.2 (Table S7).[69] RSL has six acidic residues, including structurally equivalent pairs (Glu28/Glu73 and Asp32/Asp77) located in the homologous N- and C-terminal halves of the protein.[39,43] The pH titration curves of Glu28/Glu73, which are completely buried (Table S7) and do not participate in calixarene complexation, were unaffected by the addition of 5 mM $sclx_8$. Similarly, Glu43, which stacks coplanar with the Trp74 side chain, has a high pK_a value (5.9) that was unaffected by $sclx_8$. In the case of Asp32, the pK_a was elevated by 2 pH units to 3.7 in the presence of $sclx_8$. However, $sclx_8$ had no effect on the structurally equivalent but nonbinding Asp77 (Figure 6B and Table S7). The pK_a of Asp46 was similarly elevated, but a

value was not determined directly due to spectral overlap. Data for Asn23 and Gly68, which flank Asp46, yield a pK_a of ~4.2 (Figure S11). The significantly increased pK_a values of Asp32 and Asp46 are consistent with protonation upon $sclx_8$ binding at pH ~ 4. PROPKA3.2 calculations yielded similar pK_a elevations due to $sclx_8$ (Table S7). The aromatic shielding and the negative potential of the calixarene favor the protonated state of these side chains. Host-guest pK_a elevation is well-known in small molecule systems.[38,72–75] Now, we show that such effects apply to proteins also,[76] with consequences for macrocycle-mediated assembly. Considering pH effects, the charge-state of $sclx_8$ must be considered also. While the sulfonic acids are always deprotonated in water, 4 of the phenolic groups have pK_a values ranging from ~2 to ~8.5, at 0.1 M ionic strength.[38] At pH 4, $sclx_8$ carries a formal net charge of −10 while the RSL trimer is +11, accounting for calixarene-induced protonation of Asp32 and Asp46. At pH 5.6, where binding does not occur, the formal charge of $sclx_8$ is unchanged while RSL is +5 (Table S7).

Versatility of P3 Framework and Effect of Arginine Enrichment. The facile fabrication of crystal form III (P3) makes it a prime candidate for crystal engineering. To this end, the P3 cocrystallization method (incubation of protein-$sclx_8$ mixtures at pH 4.0 and 4 °C) was applied to several chemically modified or mutated RSL variants (Figures S6 and S7). The methylated protein (RSL*), bearing dimethylated lysines with increased affinity for $sclx_n$,[77,78] behaved in a manner similar to the native protein. Crystals grew within hours and were essentially identical to the native RSL-$sclx_8$ structure (Tables 1 and S4, Figure S12). Mean isotropic temperature factor comparisons[57] indicated a more rigid assembly for the dimethylated protein compared to the other variants (all of which were determined at similar resolution. Table S8). In contrast, the acetylated protein (RSL-Ac), bearing amide-terminated Lys side chains with negligible affinity for $sclx_n$,[77] did not cocrystallize. No localized precipitate was observed during the pH-adjustment of RSL-Ac and $sclx_8$ mixtures, which remained soluble even at pH 2.0. The introduction of an additional Lys residue into RSL was in some cases compatible with the P3 assembly. Previously, we generated the double mutant N79K/T82Y named RSLex (Figures S1 and S2).[43] These mutations introduce a C-terminal feature homologous to the macrocycle binding site at Lys34 in the N-terminal blade. Interestingly, up to 10 equiv $sclx_8$ caused precipitation of RSLex, indicating that this protein was more prone to macrocycle-mediated assembly than RSL. At 15 equiv or higher, the mixture was soluble suggesting that the protein is encapsulated by the calixarene.[35] Crystals grew within hours, and the structure was identical to the original P3 assembly (Table 1, Figure S12). The new potential binding site at Lys79 did not interact with the calixarene in the crystal, suggesting that additional features are required to switch on complexation. The other mutant tested, MK-RSL, contains a Met-Lys motif as the N-terminus. In crystal form III, $sclx_8$ has peripheral interactions with the N-terminus, which is disordered even though it is cationic. Despite the additional Lys, MK-RSL remained soluble in the presence of $sclx_8$ and growth of the crystal form III was switched off, possibly because of the N-terminal disorder or steric hindrance.

The arginine-enriched mutants RSL-R_6 and RSL-R_8 deviated distinctly from the behavior of the native protein, forming amorphous precipitates at ≥0.5 equiv $sclx_8$ (Figure S7). Soluble mixtures of RSL-R_n were obtained only at 20 equiv

$sclx_8$. Similar to RSLex, these observations suggest that low equiv $sclx_8$ induces aggregation due to the "molecular glue" effect, while high equiv $sclx_8$ yields soluble samples as each protein is encapsulated by the calixarene.[35] Attempts to characterize such assemblies by solution NMR were hindered by the high concentration of $sclx_8$ required and the concomitant ionic strength effects. The pH-dependent precipitation characterized for native RSL occurred similarly for RSL-R_n (Figure S7). The precipitated mixtures at 10 equiv $sclx_8$ and pH 4.0 could be solubilized by adjusting the pH to 4.6 in the case of RSL-R_6. However, RSL-R_8 remained partly precipitated even at pH 8.0 due to its high pI. While the RSL-$sclx_8$ crystal form III grew within hours, the RSL-R_n mutants exhibited no crystal growth in this condition. Conventional vapor-diffusion experiments in the presence of ammonium sulfate yielded cocrystals. RSL-R_6 and RSL-R_8 were cocrystallized with $sclx_8$ under near-identical conditions (Table 1) while RSL-R_6 resulted in crystal form I ($P2_13$, *vide supra*) and RSL-R_8 crystallized rapidly (~3 h) to yield crystal form III ($P3$, Figures 7 and S12). This remarkable result supports the concept of pH- or charge-controlled protein-macrocycle assembly. The charge properties of RSL at pH 4 are preserved in RSL-R_8, which is highly cationic across the pH range and thus amenable to calixarene-mediated assembly.

Although RSL-R_8 yielded crystal form III like RSL, the unit cell had a shorter c axis (60 Å compared to 64 Å, Table 1). Congruently, one protein chain in the asymmetric unit was translated by ~4 Å relative to the original structure (Figure 7A). Although the positions of RSL-R_8 and RSL are similar in the two structures, the bridging calixarenes have substantially different conformations and binding sites. A more puckered $sclx_8$ occurs bound to RSL-R_8, compared to RSL-$sclx_8$ (Figure S13). Interestingly, although similar protein-$sclx_8$ interface areas were formed in both structures, the principal interface was wholly modified. In RSL, the principal interface centers on Val13 and Lys34 (*vide supra*). In RSL-R_8, binding at Val13 and Arg34 remains crucial to the assembly but the interface is ~2-fold smaller than the corresponding interface in the native protein. Instead, the principal binding site involves the new arginines, Arg43 and Arg46, as well as Pro44. Apparently, this binding site fulfils a similar function to the protonated Asn23/Asp46/Gly68 patch in the native protein but with a ~2-fold larger surface area. Arg43 and Pro44[36] dominate this interface with cation-π and CH-π bonds to the calixarene. The RSL-R_8 structure also suggests why RSL-R_6 did not crystallize in crystal form III ($P3$). The calixarene conformation supported by Arg43 in RSL-R_8 cannot be maintained by Glu43 in RSL-R_6.

■ DISCUSSION

Anionic calixarenes are well-established "glues" for the assembly and crystallization of cationic proteins such as cyt$c$,[34,35,37] lysozyme,[33,78] and PAF.[36] Here, we have demonstrated the cocrystallization of "neutral" RSL with highly anionic $sclx_8$. Three distinct crystal forms were obtained across a broad pH range (2.2−9.5) raising the possibility of general applications of calixarene-mediated protein assembly. The tightly packed crystal form I ($P2_13$) grew at pH 4.8−9.5 and required >1.5 M ammonium sulfate, consistent with the hydrophobic effect, rather than charge−charge interactions, as the dominant contributor to assembly. At pH ≤ 4.2, two types of porous frameworks were obtained. At this pH, RSL is cationic, which favors calixarene binding. NMR experiments suggest that calixarene complexation alters the cationic

character of the protein by raising the pK_a values of two acidic residues (Asp32 and Asp46). $sclx_8$ complexation may also modulate the protonation of Lys34. The disorder of this side chain in the $P2_13$ structures suggests that it is deprotonated (Figure 2).

Crystal forms II and III, are highly porous frameworks with pore diameters of ~4 and ~3 nm, respectively (Figure 1 and Table 2). The most porous assembly (form II, $I23$) was obtained at pH 2.2−4.0 and <1 M ammonium sulfate. This framework is built from $sclx_8$ dimers and RSL trimers acting as nodes. Each node is a protein coated with six calixarenes in an octahedral arrangement (Figure 8). The resulting *primitive*
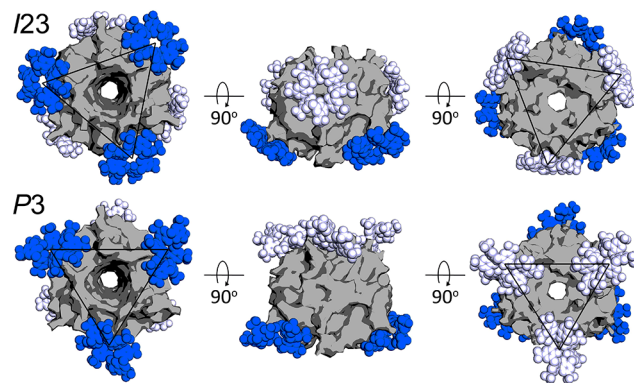
Figure 8. Calixarene encapsulation of RSL is evident in the building blocks of crystal forms II ($I23$) and III ($P3$). In each assembly, the RSL trimer (gray surface) is masked by six $sclx_8$ molecules. The ligands are rendered in blue or white corresponding to the large and small interfaces, respectively. Black triangles indicate C_3 planes that connect symmetry-related calixarenes.

cubic lattice has no protein−protein contacts and relies instead on protein-calixarene and calixarene-calixarene contacts (Figure 1). Protein assembly mediated by macrocycle dimers has been observed previously with porphyrins[44] and calixarenes.[34] For example, phosphonato-calix[6]arene forms a dimeric disc to mediate a porous assembly of cyt$c$ (PDB 5LYC).[34] Similar to crystal form II ($I23$), form III ($P3$) involves RSL trimers bridged together by six $sclx_8$ ligands arranged in triads on two C_3 planes (Figure 8). However, the positioning of the calixarene triads on opposite ends of the trimer disables the possibility of cubic assembly. Also like the $I23$ case, the $P3$ framework assembles only in acidic conditions (pH ≤ 4.0). Such pH-triggered assembly is an exciting new aspect of macrocycle-directed protein assembly. pH-dependent cationization as a driver of protein-macrocycle framework assembly was corroborated by the results with RSL-R_8. This permanently cationic protein also yielded the $P3$ framework but under mildly alkaline conditions (Figures 7 and S12) where arginine side chains remain protonated.

Depending on the conditions, RSL-$sclx_8$ mixtures remain soluble or form crystals ranging from microcrystalline precipitates to millimeter-scale crystals. For example, RSL and 10 equiv $sclx_8$ yields a microcrystalline precipitate at pH ≤ 3.4 and room temperature (Figure S7). The precipitate is dissolved by raising the pH to >4.2 or by increasing the $sclx_8$ concentration to 20 equiv (Figures 4A and S8). The effect of increasing pH is explained simply by the less cationic RSL having lower affinity for the calixarene. The interruption of precipitation at 20 equiv $sclx_8$ suggests calixarene-coating or encapsulation of the protein.[35] At low equiv, the calixarene

bridges two or more proteins resulting in aggregation. At high equiv $sclx_8$, each protein binding site is masked by calixarene and the "glue" activity is switched off as the interaction between protein-calixarene particles switches from attractive to repulsive. Both the $I23$ and the $P3$ frameworks, with significant macrocycle-masking of the symmetric protein building block (Figure 8), support the assumption of encapsulation[35] in solution. This behavior is analogous to the reentrant condensation phase behavior of proteins in the presence of multivalent counterions.[51,52] Protein encapsulation is currently a topic of intense development[35,50,79,80] and macrocycle-masking may be a simple approach to tackle this challenge.

The precipitation of RSL-R$_6$, RSL-R$_8$, and RSLex operated via similar condensed regimes, but unlike RSL, the variants formed amorphous precipitates at pH 4.0 (Figure S7). Apparently, arginine-enrichment and the increased pI leads to enhanced calixarene binding and rapid precipitation. RSL-R$_6$ and RSLex required pH 4.6 to switch off precipitation while the most cationic variant, RSL-R$_8$, precipitated even at pH 8.0. Lysine to arginine mutagenesis tends to reduce protein solubility, a potential contributing factor in the precipitation of the arginine-enriched variants.[81] At pH 4.0 and 10 equiv $sclx_8$, the lysine-enriched variant RSLex displayed the same precipitation behavior as the Arg-rich variants, suggesting that the assembly of all three variants was a consequence of higher $sclx_8$ affinity rather than reduced protein solubility. Unlike RSLex, soluble mixtures of RSL-R$_6$ and RSL-R$_8$ at high equiv $sclx_8$ and low ionic strength yielded no cocrystals. The versatility of guanidinium groups at protein–protein and protein–ligand interfaces is well-established.[53,82–84] We reason that the "stickiness" of the Arg-rich variants toward $sclx_8$ yielded highly encapsulated protein-calixarene particles incapable of self-association. This hypothesis is supported by cocrystallization of RSL-R$_6$ and RSL-R$_8$ occurring only at high salt concentration wherein counterions compete for $sclx_8$ complexation and lower its affinity for the protein surface.

With considering the importance of protein frameworks to both basic science and biotechnology, the availability of facile fabrication strategies is essential to stimulate progress. Vastly different strategies are currently in development and applicable in distinct settings. Significant progress has been achieved by harnessing naturally occurring frameworks such as protein cages,[6,15] capsids,[8,46] and crystals.[11,14] For example, the supramolecular assembly of ferritin cages can be directed by metal-coordination sites.[15,24] Protein crystals that grow naturally inside living cells have been used to capture cargo including dyes or enzymes.[11,14] In a particularly striking example, a 32 kDa lipase was trapped within the pores of Cry3Aa crystals that grow in *Bacillus thuringiensis* cells.[14] The strategies used to engineer frameworks include designed oligomerization,[28,47–49] charge–charge interactions,[25–27] and ligand-mediated assembly.[19,23,37,45] The latter stands out as the most promising since (1) protein engineering is not required and (2) the assembly process can be regulated by the ligand concentration and/or the presence of inhibitors.[35,37] Custom-made bivalent ligands comprising a sugar and rhodamine dye have been used to generate porous frameworks and other assemblies of concanavalin A.[19,23] With biotech applications in mind, inexpensive, highly water-soluble and nontoxic ligands are desirable. Anionic calixarenes fit these requirements and there is increasing evidence of their utility for framework assembly.[34–37] Here, we demonstrated a pH-triggered framework fabrication of a "neutral" protein with $sclx_8$, in the presence or absence of a precipitant. The pH dependence affords a second level of control over assembly. Finally, the host–guest interactions of $sclx_8$[29–32] confers the crystal with additional binding sites.[37] The different pore dimensions in the frameworks (Table 2) along with the varying solvent accessibilities of the calixarene (∼35, 45 and 55% solvent exposed in crystal forms I, II and III, respectively, Figure 1 and Table S5) result in materials with different ligand uptake capacities.

CONCLUSIONS

Various sophisticated frameworks have been described to date, usually with the requirement for protein engineering or specific assembly inducing ligands.[19,23,25–27,47–49] Sulfonato-calixarenes are commercially available: biocompatible ligands that can be combined with protein building blocks to yield frameworks in a manner reminiscent of MOF manufacture using off-the-shelf reagents.[12] The data presented here greatly widen the scope for $sclx_n$-assisted protein assembly and crystallization to include "neutral" target proteins. The primary advantage is facile pH-controlled framework fabrication. pH-induced cationization of the protein enhances calixarene complexation and results in porous assemblies. pK_a modulation, well-characterized in protein–ligand binding[76] and in small molecule host–guest systems,[72–75] is a contributing factor to macrocycle-mediated protein assembly.

Rapid, scalable production of protein-based frameworks is essential for their development as biomaterials and biologic alternatives to MOFs/COFs. While the $I23$ framework required days to grow, the $P3$ framework grew within hours in the absence of precipitant. This rapid fabrication of millimeter-scale crystals (Figure 4) under pH control and without the need for precipitants or specialized equipment is attractive for manufacturing processes, including the purification of therapeutic proteins.[85] Furthermore, the controlled nucleation and the formation of microcrystalline samples within minutes has applications in serial crystallography.[86,87] Complementing these valuable crystallization properties, the sulfur-rich $sclx_8$ is a promising phasing agent (Figure 3).

While calixarene-directed assembly is easily applied to proteins in their native state,[34–37] protein engineering can complement the supramolecular strategy. Here, the permanently cationic Arg-rich variant RSL-R$_8$ resulted in the same $P3$ framework as the native protein at low pH. Finally, both the $I23$ and the $P3$ frameworks provide evidence of protein encapsulation by $sclx_8$. These strategies for controlled protein assembly have potential applications in therapeutics and smart biomaterials. The opportunity for auxiliary host–guest chemistry within protein-macrocycle frameworks[37] makes these materials prime candidates for the uptake of biomolecules such as enzyme substrates or other proteins. Such developments boost the application of biosupramolecular chemistry as a general strategy for protein framework fabrication.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c10697.

Sequences and electrostatic surfaces of RSL variants, representative cocrystals and conditions, X-ray data

collection and refinement statistics, crystal structure details, NMR data and pK_a analysis (PDF)

AUTHOR INFORMATION

Corresponding Author

Peter B. Crowley − *School of Chemistry, National University of Ireland Galway, Galway H91 TK33, Ireland;* orcid.org/0000-0002-5365-0096; Phone: +353 91 49 24 80; Email: peter.crowley@nuigalway.ie

Authors

Kiefer O. Ramberg − *School of Chemistry, National University of Ireland Galway, Galway H91 TK33, Ireland*

Sylvain Engilberge − *School of Chemistry, National University of Ireland Galway, Galway H91 TK33, Ireland; Swiss Light Source, Paul Scherrer Institut, Villigen PSI 5232, Switzerland*

Tomasz Skorek − *School of Chemistry, National University of Ireland Galway, Galway H91 TK33, Ireland*

Complete contact information is available at:
https://pubs.acs.org/10.1021/jacs.0c10697

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Eder, M.; Amini, S.; Fratzl, P. Biological composites - complex structures for functional diversity. *Science* **2018**, *362*, 543−547.

(2) Abascal, N. C.; Regan, L. The past, present and future of protein-based materials. *Open Biol.* **2018**, *8*, 180113.

(3) Kuan, S. L.; Bergamini, F. R.; Weil, T. Functional protein nanostructures: a chemical toolbox. *Chem. Soc. Rev.* **2018**, *47*, 9069−9105.

(4) Beloqui, A.; Cortajarena, A. L. Protein-based functional hybrid bionanomaterials by bottom-up approaches. *Curr. Opin. Struct. Biol.* **2020**, *63*, 74−81.

(5) Lopez, S.; Rondot, L.; Lepretre, C.; Marchi-Delapierre, C.; Menage, S.; Cavazza, C. Cross-linked artificial enzyme crystals as heterogeneous catalysts for oxidation reactions. *J. Am. Chem. Soc.* **2017**, *139*, 17994−18002.

(6) Uchida, M.; McCoy, K.; Fukuto, M.; Yang, L.; Yoshimura, H.; Miettinen, H. M.; LaFrance, B.; Patterson, D. P.; Schwarz, B.; Karty, J. A.; Prevelige, P. E.; Lee, B.; Douglas, T. Modular self-assembly of protein cage lattices for multistep catalysis. *ACS Nano* **2018**, *12*, 942−953.

(7) Anaya-Plaza, E.; Aljarilla, A.; Beaune, G.; Timonen, J. V.; de la Escosura, A.; Torres, T.; Kostiainen, M. A. Phthalocyanine-virus nanofibers as heterogeneous catalysts for continuous-flow photo-oxidation processes. *Adv. Mater.* **2019**, *31*, 1902582.

(8) Falkner, J. C.; Turner, M. E.; Bosworth, J. K.; Trentler, T. J.; Johnson, J. E.; Lin, T.; Colvin, V. L. Virus crystals as nanocomposite scaffolds. *J. Am. Chem. Soc.* **2005**, *127*, 5274−5275.

(9) Abe, S.; Tsujimoto, M.; Yoneda, K.; Ohba, M.; Hikage, T.; Takano, M.; Kitagawa, S.; Ueno, T. Porous protein crystals as reaction vessels for controlling magnetic properties of nanoparticles. *Small* **2012**, *8*, 1314−1319.

(10) Ueno, T. Porous protein crystals as reaction vessels. *Chem. - Eur. J.* **2013**, *19*, 9096−9102.

(11) Abe, S.; Tabe, H.; Ijiri, H.; Yamashita, K.; Hirata, K.; Atsumi, K.; Shimoi, T.; Akai, M.; Mori, H.; Kitagawa, S.; Ueno, T. Crystal engineering of self-assembled porous protein materials in living cells. *ACS Nano* **2017**, *11*, 2410−2419.

(12) Chen, Y.; Li, P.; Modica, J. A.; Drout, R. J.; Farha, O. K. Acid-resistant mesoporous metal-organic framework toward oral insulin delivery: Protein encapsulation, protection, and release. *J. Am. Chem. Soc.* **2018**, *140*, 5678−5681.

(13) Li, M.; Qiao, S.; Zheng, Y.; Andaloussi, Y. H.; Li, X.; Zhang, Z.; Li, A.; Cheng, P.; Ma, S.; Chen, Y. Fabricating covalent organic framework capsules with commodious microenvironment for enzymes. *J. Am. Chem. Soc.* **2020**, *142*, 6675−6681.

(14) Heater, B. S.; Yang, Z.; Lee, M. M.; Chan, M. K. *In vivo* enzyme entrapment in a protein crystal. *J. Am. Chem. Soc.* **2020**, *142*, 9879−9893.

(15) Sontz, P. A.; Bailey, J. B.; Ahn, S.; Tezcan, F. A. A metal organic framework with spherical protein nodes: rational chemical design of 3D protein crystals. *J. Am. Chem. Soc.* **2015**, *137*, 11598−11601.

(16) Gu, C.; Chen, H.; Wang, Y.; Zhang, T.; Wang, H.; Zhao, G. Structural insight into binary protein metal-organic frameworks with ferritin nanocages as linkers and nickel clusters as nodes. *Chem. - Eur. J.* **2020**, *26*, 3016−3021.

(17) Howitz, W. J.; Wierzbicki, M.; Cabanela, R. W.; Saliba, C.; Motavalli, A.; Tran, N.; Nowick, J. S. Interpenetrating cubes in the X-ray crystallographic structure of a peptide derived from Medin$_{19−36}$. *J. Am. Chem. Soc.* **2020**, *142*, 15870−15875.

(18) Fegan, A.; White, B.; Carlson, J. C.; Wagner, C. R. Chemically controlled protein assembly: techniques and applications. *Chem. Rev.* **2010**, *110*, 3315−3336.

(19) Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y.; Chen, G.; Jiang, M. Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nat. Commun.* **2014**, *5*, 4634.

(20) McGovern, R. E.; Feifel, S. C.; Lisdat, F.; Crowley, P. B. Microscale crystals of cytochrome *c* and calixarene on electrodes: Interprotein electron transfer between defined sites. *Angew. Chem., Int. Ed.* **2015**, *54*, 6356−6359.

(21) Bosmans, R. P.; Briels, J. M.; Milroy, L. G.; de Greef, T. F.; Merkx, M.; Brunsveld, L. Supramolecular control over split-luciferase complementation. *Angew. Chem., Int. Ed.* **2016**, *55*, 8899−8903.

(22) de Vink, P. J.; Briels, J. M.; Schrader, T.; Milroy, L. G.; Brunsveld, L.; Ottmann, C. A binary bivalent supramolecular assembly platform based on cucurbit[8]uril and dimeric adapter protein 14−3-3. *Angew. Chem., Int. Ed.* **2017**, *56*, 8998−9002.

(23) Yang, G.; Ding, H. M.; Kochovski, Z.; Hu, R.; Lu, Y.; Ma, Y. Q.; Chen, G.; Jiang, M. Highly ordered self-assembly of native proteins into 1D, 2D, and 3D structures modulated by the tether length of assembly-inducing ligands. *Angew. Chem., Int. Ed.* **2017**, *56*, 10691−10695.

(24) Zhang, L.; Bailey, J. B.; Subramanian, R. H.; Groisman, A.; Tezcan, F. A. Hyperexpandable, self-healing macromolecular crystals with integrated polymer networks. *Nature* **2018**, *557*, 86−91.

(25) Künzle, M.; Eckert, T.; Beck, T. Binary protein crystals for the assembly of inorganic nanoparticle superlattices. *J. Am. Chem. Soc.* **2016**, *138*, 12731−12734.

(26) Simon, A. J.; Zhou, Y.; Ramasubramani, V.; Glaser, J.; Pothukuchy, A.; Gollihar, J.; Gerberich, J. C.; Leggere, J. C.; Morrow, B. R.; Jung, C.; Glotzer, S. C.; Taylor, D. W.; Ellington, A. D. Supercharging enables organized assembly of synthetic biomolecules. *Nat. Chem.* **2019**, *11*, 204−212.

(27) Minamihata, K.; Tsukamoto, K.; Adachi, M.; Shimizu, R.; Mishina, M.; Kuroki, R.; Nagamune, T. Genetically fused charged peptides induce rapid crystallization of proteins. *Chem. Commun.* **2020**, *56*, 3891−3894.

(28) Laniado, J.; Yeates, T. O. A complete rule set for designing symmetry combination materials from protein molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 31817.

(29) Shinkai, S.; Araki, K.; Manabe, O. NMR determination of association constants for calixarene complexes. Evidence for the formation of a 1:2 complex with calix[8]arene. *J. Am. Chem. Soc.* **1988**, *110*, 7215−7217.

(30) Hamuro, Y.; Calama, M. C.; Park, H. S.; Hamilton, A. D. A calixarene with four peptide loops: an antibody mimic for recognition of protein surfaces. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2680−2683.

(31) Zadmard, R.; Schrader, T. Nanomolar protein sensing with embedded receptor molecules. *J. Am. Chem. Soc.* **2005**, *127*, 904−915.

(32) Baldini, L.; Casnati, A.; Sansone, F. Multivalent and multifunctional calixarenes in bionanotechnology. *Eur. J. Org. Chem.* **2020**, *32*, 5056−5069.

(33) McGovern, R. E.; McCarthy, A. A.; Crowley, P. B. Protein assembly mediated by sulfonatocalix[4]arene. *Chem. Commun.* **2014**, *50*, 10412−10415.

(34) Rennie, M. L.; Doolan, A. M.; Raston, C. L.; Crowley, P. B. Protein dimerization on a phosphonated calix[6]arene disc. *Angew. Chem., Int. Ed.* **2017**, *56*, 5517−5521.

(35) Rennie, M. L.; Fox, G. C.; Pérez, J.; Crowley, P. B. Auto-regulated protein assembly on a supramolecular scaffold. *Angew. Chem., Int. Ed.* **2018**, *57*, 13764−13769.

(36) Alex, J. M.; Rennie, M. L.; Engilberge, S.; Lehoczki, G.; Dorottya, H.; Fizil, A.; Batta, G.; Crowley, P. B. Calixarene-mediated assembly of a small antifungal protein. *IUCrJ* **2019**, *6*, 238−247.

(37) Engilberge, S.; Rennie, M. L.; Dumont, E.; Crowley, P. B. Tuning protein frameworks via auxiliary supramolecular interactions. *ACS Nano* **2019**, *13*, 10343−10350.

(38) Suga, K.; Ohzono, T.; Negishi, M.; Deuchi, K.; Morita, Y. Effect of various cations on the acidity of *p*-sulfonatocalixarenes. *Supramol. Sci.* **1998**, *5*, 9−14.

(39) Kostlánová, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerová, M.; Imberty, A. The fucose-binding lectin from *Ralstonia solanacearum*. *J. Biol. Chem.* **2005**, *280*, 27839−27849.

(40) Antonik, P. M.; Volkov, A. N.; Broder, U. N.; Lo Re, D.; van Nuland, N. A.; Crowley, P. B. Anomer-specific recognition and dynamics in a fucose-binding lectin. *Biochemistry* **2016**, *55*, 1195−1203.

(41) Ribeiro, J. P.; Villringer, S.; Goyard, D.; Coche-Guerente, L.; Höferlin, M.; Renaudet, O.; Römer, W.; Imberty, A. Tailor-made Janus lectin with dual avidity assembles glycoconjugate multilayers and crosslinks protocells. *Chem. Sci.* **2018**, *9*, 7634−7641.

(42) Ramberg, K. O.; Antonik, P. M.; Cheung, D. L.; Crowley, P. B. Measuring the Impact of PEGylation on a protein-polysaccharide interaction. *Bioconjugate Chem.* **2019**, *30*, 1162−1168.

(43) Guagnini, F.; Engilberge, S.; Ramberg, K. O.; Perez, J.; Crowley, P. B. Engineered assembly of a protein-cucurbituril biohybrid. *Chem. Commun.* **2020**, *56*, 360−363.

(44) Goel, M.; Damai, R. S.; Sethi, D. K.; Kaur, K. J.; Maiya, B. G.; Swamy, M. J.; Salunke, D. M. Crystal structures of the PNA-porphyrin complex in the presence and absence of lactose: Mapping the conformational changes on lactose binding, interacting surfaces, and supramolecular aggregations. *Biochemistry* **2005**, *44*, 5588−5596.

(45) Vandebroek, L.; Noguchi, H.; Kamata, K.; Tame, J. R. H.; Van Meervelt, L.; Parac-Vogt, T. N.; Voet, A. R. D. Hybrid assemblies of a symmetric designer protein and polyoxometalates with matching symmetry. *Chem. Commun.* **2020**, *56*, 11601−11604.

(46) Van Eldijk, M. B.; Wang, J. C. Y.; Minten, I. J.; Li, C.; Zlotnick, A.; Nolte, R. J.; Cornelissen, J. J. L. M.; van Hest, J. C. Designing two self-assembly mechanisms into one viral capsid protein. *J. Am. Chem. Soc.* **2012**, *134*, 18506−18509.

(47) Lai, Y. T.; Tsai, K. L.; Sawaya, M. R.; Asturias, F. J.; Yeates, T. O. Structure and flexibility of nanoscale protein cages designed by symmetric self-assembly. *J. Am. Chem. Soc.* **2013**, *135*, 7738−7743.

(48) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D. Accurate design of co-

assembling multi-component protein nanomaterials. *Nature* **2014**, *510*, 103−108.

(49) Cannon, K. A.; Nguyen, V. N.; Morgan, C.; Yeates, T. O. Design and Characterization of an Icosahedral protein cage formed by a double-fusion protein containing three distinct symmetry elements. *ACS Synth. Biol.* **2020**, *9*, 517−524.

(50) Fujita, D.; Suzuki, K.; Sato, S.; Yagi-Utsumi, M.; Yamaguchi, Y.; Mizuno, N.; Kumasaka, T.; Takata, M.; Noda, M.; Uchiyama, S.; Kato, K.; Fujita, M. Protein encapsulation within synthetic molecular hosts. *Nat. Commun.* **2012**, *3*, 1−7.

(51) Asor, R.; Ben-nun-Shaul, O.; Oppenheim, A.; Raviv, U. Crystallization, reentrant melting, and resolubilization of virus nanoparticles. *ACS Nano* **2017**, *11*, 9814−9824.

(52) Bye, J. W.; Curtis, R. A. Controlling phase separation of lysozyme with polyvalent anions. *J. Phys. Chem. B* **2019**, *123*, 593−605.

(53) Kyne, C.; Crowley, P. B. Short arginine motifs drive protein stickiness in the *Escherichia coli* cytoplasm. *Biochemistry* **2017**, *56*, 5026−5032.

(54) Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O. S.; Paciorek, W.; Womack, T.; Bricogne, G. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 293−302.

(55) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.

(56) Evans, P. R.; Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204−1214.

(57) Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 282−292.

(58) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macro-molecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.

(59) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.

(60) Weinert, T.; Olieric, V.; Waltersperger, S.; Panepucci, E.; Chen, L.; Zhang, H.; Zhou, D.; Rose, J.; Ebihara, A.; Kuramitsu, S.; Li, D.; Howe, N.; Schnapp, G.; Pautsch, A.; Bargsten, K.; Prota, A. E.; Surana, P.; Kottur, J.; Nair, D. T.; Basilico, F.; Cecatiello, V.; Pasqualato, S.; Boland, A.; Weichenrieder, O.; Wang, B.; Steinmetz, M. O.; Caffrey, M.; Wang, M. Fast native-SAD phasing for routine macromolecular structure determination. *Nat. Methods* **2015**, *12*, 131−133.

(61) Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 479−485.

(62) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.

(63) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.

(64) Krissinel, E.; Henrick, K. J. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774.

(65) Juers, D. H.; Ruffin, J. MAP_CHANNELS: A computation tool to aid in the visualization and characterization of solvent channels in macromolecular crystals. *J. Appl. Crystallogr.* **2014**, *47*, 2105−2108.

(66) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. D. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277−293.

(67) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 687−696.

(68) Croke, R. L.; Patil, S. M.; Quevreaux, J.; Kendall, D. A.; Alexandrescu, A. T. NMR determination of pK_a values in α-synuclein. *Protein Sci.* **2011**, *20*, 256−269.

(69) Søndergaard, C. R.; Olsson, M. H.; Rostkowski, M.; Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK_a values. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.

(70) Sun, Z.; Liu, Q.; Qu, G.; Feng, Y.; Reetz, M. T. Utility of B-factors in protein science: interpreting rigidity, flexibility, and internal motion and engineering thermostability. *Chem. Rev.* **2019**, *119*, 1626−1665.

(71) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno, E. B.; McIntosh, L. P. Arginine: Its pK_a value revisited. *Protein Sci.* **2015**, *24*, 752−761.

(72) Saleh, N.; Koner, A. L.; Nau, W. M. Activation and stabilization of drugs by supramolecular pK_a shifts: Drug-delivery applications tailored for cucurbiturils. *Angew. Chem., Int. Ed.* **2008**, *47*, 5398−5401.

(73) Basílio, N.; Gago, S.; Parola, A. J.; Pina, F. Contrasting pK_a shifts in cucurbit[7]uril host-guest complexes governed by an interplay of hydrophobic effects and electrostatic interactions. *ACS Omega* **2017**, *2*, 70−75.

(74) Cai, X.; Kataria, R.; Gibb, B. C. Intrinsic and extrinsic control of the pK_a of thiol guests inside yoctoliter containers. *J. Am. Chem. Soc.* **2020**, *142*, 8291−8298.

(75) Paul, T. J.; Vilseck, J. Z.; Hayes, R. L.; Brooks, C. L., III Exploring pH dependent host/guest binding affinities. *J. Phys. Chem. B* **2020**, *124*, 6520−6528.

(76) Onufriev, A. V.; Alexov, E. Protonation and pK changes in protein-ligand binding. *Q. Rev. Biophys.* **2013**, *46*, 181−209.

(77) Beshara, C. S.; Jones, C. E.; Daze, K. D.; Lilgert, B. J.; Hof, F. A simple calixarene recognizes post-translationally methylated lysine. *ChemBioChem* **2010**, *11*, 63−66.

(78) McGovern, R. E.; Snarr, B. D.; Lyons, J. A.; McFarlane, J.; Whiting, A. L.; Paci, I.; Hof, F.; Crowley, P. B. Structural study of a small molecule receptor bound to dimethyllysine in lysozyme. *Chem. Sci.* **2015**, *6*, 442−449.

(79) Azuma, Y.; Bader, D. L. V.; Hilvert, D. Substrate sorting by a supercharged nanoreactor. *J. Am. Chem. Soc.* **2018**, *140*, 860−863.

(80) Chapman, R.; Stenzel, M. H. All wrapped up: Stabilization of enzymes within single enzyme nanoparticles. *J. Am. Chem. Soc.* **2019**, *141*, 2754−2769.

(81) Warwicker, J.; Charonis, S.; Curtis, R. A. Lysine and arginine content of proteins: computational analysis suggests a new tool for solubility design. *Mol. Pharmaceutics* **2014**, *11*, 294−303.

(82) Crowley, P. B.; Golovin, A. Cation-π interactions in protein-protein interfaces. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 231−239.

(83) Noresson, A. L.; Aurelius, O.; Öberg, C. T.; Engström, O.; Sundin, A. P.; Håkansson, M.; Stenström, O.; Akke, M.; Logan, D. T.; Leffler, H.; Nilsson, U. J. Designing interactions by control of protein−ligand complex conformation: tuning arginine-arene interaction geometry for enhanced electrostatic protein−ligand interactions. *Chem. Sci.* **2018**, *9*, 1014−1021.

(84) Mogaki, R.; Okuro, K.; Ueki, R.; Sando, S.; Aida, T. Molecular glue that spatiotemporally turns on protein-protein interactions. *J. Am. Chem. Soc.* **2019**, *141*, 8035−8040.

(85) Hekmat, D.; Huber, M.; Lohse, C.; von den Eichen, N.; Weuster-Botz, D. Continuous crystallization of proteins in a stirred classified product removal tank with a tubular reactor in bypass. *Cryst. Growth Des.* **2017**, *17*, 4162−4169.

(86) Schlichting, I. Serial femtosecond crystallography: the first five years. *IUCrJ* **2015**, *2*, 246−255.

(87) de Wijn, R.; Hennig, O.; Roche, J.; Engilberge, S.; Rollet, K.; Fernandez-Millan, P.; Brillet, K.; Betat, H.; Mörl, M.; Roussel, A.; Girard, E.; Mueller-Dieckmann, C.; Fox, G. C.; Olieric, V.; Gavira, J. A.; Lorber, B.; Sauter, C. A simple and versatile microfluidic device for efficient biomacromolecule crystallization and structural analysis by serial crystallography. *IUCrJ* **2019**, *6*, 454−464.