Facile Fabrication of Protein–Macrocycle Frameworks
Kiefer O. Ramberg, Sylvain Engilberge, Tomasz Skorek, and Peter B. Crowley*  

ABSTRACT: Precisely defined protein aggregates, as exemplified by crystals, have applications in catalytic devices.1−4 Porous crystals are of particular interest considering their capacity to capture (store) and transform biomolecules.6,8 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.7 This paper describes the facile fabrication of millimeter-scale, high-porosity, solid-state composites of precisely arrayed protein and synthetic macrocycle with variable cavity dimensions,29 a solvent accessible surface area (ASA) of ∼1600 Å², 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 (cytc, isoelectric point, pH ∼9.5) with, with 1 equiv sclx₈ forming a tetramer, and 3 equiv yielding a calixarene-coated (encapsulated) protein.35 The ligand-protein ratio influences also the formation of cytsclx₈ crystalline frameworks with varying porosities, ranging from 65 to 85% solvent content.37 The most porous framework is mediated exclusively by protein-calixarene contacts and requires at least 3 equiv sclx₈ with respect to the cationic protein.  

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.1−4 Porous crystals are of particular interest considering their capacity to capture (store) and transform biomolecules.6,8 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.7 This paper describes the facile fabrication of millimeter-scale, high-porosity, solid-state composites of precisely arrayed protein and synthetic macrocycle with variable cavity dimensions,29 a solvent accessible surface area (ASA) of ∼1600 Å², 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 (cytc, isoelectric point, pH ∼9.5) with, with 1 equiv sclx₈ forming a tetramer, and 3 equiv yielding a calixarene-coated (encapsulated) protein.35 The ligand-protein ratio influences also the formation of cytsclx₈ crystalline frameworks with varying porosities, ranging from 65 to 85% solvent content.37 The most porous framework is mediated exclusively by protein-calixarene contacts and requires at least 3 equiv sclx₈ 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 sclx₈.22−24 The 6-bladed β-propeller Ralstonia solanacearum lectin (RSL, pH ∼ 6.5) was selected as the model protein.39−43 RSL is a rigid, C₃-symmetric spheroid with high thermal stability making it an interesting candidate for protein-based frame-
Trimeric RSL possesses pseudo C₃-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, related β-propellers, ferritin, viral capsids, and engineered cages. The cocrystallization conditions for RSL and several variants including RSLex and MK-RSL (each containing one extra lysine), and two arginine-enriched mutants, RSL-R₆ and RSL-R₈ (Figures S1 and S2). In RSL-R₆, the three lysines of native RSL are replaced by arginine. RSL-R₈ includes these mutations as well as two acidic residues added in parallel. 

## EXPERIMENTAL SECTION

### Materials

Stock solutions of sclx₈ (Tokyo Chemical Industry) were prepared in water, and the pH was adjusted to 8.0. Unlabeled and ¹⁵N-labeled RSL samples were produced in E. coli BL21 transformed with the pET25rsl plasmid. The modified pET25s vectors that encode RSL-R₆ and RSLex were reported previously. The vectors encoding RSL-R₆ (K25R/K34R/E43R/D46R/K83R) and MK-RSL were produced by Genscript. All proteins, except RSL-R₈, were purified 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 (Table S1). Protein concentrations were determined spectrophoto metrically with ε₂₅₀ = 44.46 M⁻¹ cm⁻¹ for the monomer.

### Cocrystallization Trials

All experiments were performed with α-fructose bound RSL and variants. Protein-sclx₈ cocrystals were obtained at 20 °C by using commercial (JCSG++ HTS, Jena Bioscience) or homemade screens, applied with an Orixa 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 (0.6 to 1.0 M ammonium sulfate and 0.1 M sodium citrate pH 4.0). Crystals were obtained at 20 °C. Crystallographic data were collected using an Olympus SX161 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. Data were integrated in XDS and the integrated intensities were scaled and merged in AIMLESS and POINTLESS in CCP4 and assessed for pathologies in phenix.xtreme. Structures were solved by molecular replacement in PHASER using the RSL monomer (derived from PDB 2BT9) 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. Diffraction frames (deposited on Zenodo, DOI: 10.5281/zenodo.3944886) were integrated using XDS and scaled with AIMLESS and POINTLESS. Substructure determination, phasing, and model building were performed in SHEXL. 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 sclx₈ (PDB id EVB) and α-fructose (PDB id BDF) were added to each model. Iterative cycles of model building in COOT and refinement in phenix.refine were performed until no

### Table 1. Cocrystallization Conditions and Structure Properties

| Form | Protein | equiv sclx₈ | (NH₄)₂SO₄ (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 | Li₂SO₄ | P₂₁₃ | 64² | 1.2 | 6Z6O |
| RSL-R₆ | 80 | 1.6 | Tris-HCl | 8.5 | Li₂SO₄ | | | | | 6Z6L |
| RSL-R₈ | 50 | 1.6 | Tris-HCl | 8.5 | Li₂SO₄ | | | | | 6Z5Z |
| RSL | 50 | 2.4 | MES | 6.8 | | | | | | 6Z5W |
| RSL | 80 | 2.0 | acetate | 4.8 | | | | | | 6Z5X |
| II RSL | 50 | 0.8 | citrate | 4.0 | P₁ | 104 | 1.3 | 6Z5G |
| RSL | 10 | 0.8 | Gly-HCl | 2.2 | MgCl₂ | | | | | 6Z5M |
| III RSL | 10 | | acetate | 4.0 | P₁ | 60 x 60 x 64 | 1.3 | 6Z5Q |
| RSL | 10 | | acetate | 4.0 | | 60 x 60 x 64 | | 1.7ALF |
| RSLex | 15 | | acetate | 4.0 | | 60 x 60 x 64 | | 1.7ALG |
| RSL-R₈ | 50 | 1.3 | Tris-HCl | 8.5 | Li₂SO₄ | 60 | 1.4 | 6Z5P |

S3, Table S1). Protein concentrations were determined spectrophotometrically with ε₂₅₀ = 44.46 M⁻¹ cm⁻¹ for the monomer.
further improvements in the Rfree or electron density was obtained. All of the structures were validated in MolProbity.65 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.66 Crystal porosity was analyzed in MAP_CHANNELS.65 The molar protein concentration, [P], in each crystal form was calculated by

\[ [P] = \frac{N}{V} \times 10^{-27} \]

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 15N-labeled protein in 20 mM phosphate buffer, 50 mM NaCl, 5 mM n-fructose, and 10% D2O. 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 1H–15N 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 NMRpipe67 and CCPN,68 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 (Δδ_max) as the fit parameters. pH titration curves were generated from 1H–15N 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

\[ \Delta \delta = \Delta \delta_{\text{low}} - \frac{\Delta \delta_{\text{low}} - \Delta \delta_{\text{high}}}{1 + 10^{n(pK_a - \phiH)}} \]

where pK_a is the ionization constant, Δδ_{low} and Δδ_{high} are the low and high pH chemical shift plateaus, and n is the apparent Hill coefficient.69

Empirical pK_a Calculations. The pK_a values of the acidic residues measured by NMR were compared with empirical estimates obtained using PROPKA3.2.70 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 6ZQS) 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 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 P213. 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 for Lys34 across the pH range, Lys34 was modeled in two conformations with high temperature factors (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 Cα 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 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

| Form | Space Group | RSL-sclx8 (mM) | [P] (mM) | S.C. (%) | Pore Ø (nm) |
|------|-------------|---------------|----------|---------|-------------|
| I    | P213        | 1:1           | 76       | 36      | 1.7         |
| II   | P213        | 1:2           | 36       | 66      | 4.2         |
| III  | P213        | 1:1           | 50       | 59      | 2.8         |

Table 2. RSL-sclx8 Cocrystal Forms and Properties

“Protein:ligand ratio per RSL monomer. bC Calculated protein concentration based on unit cell contents. Solvent content estimated from total mass (protein plus sclx8). dDiameter of widest pore, calculated in MAP_CHANNELS.65
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 $I\overline{2}3$ 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 1. (A) Crystal packing in RSL-sclx₈ cocrystal forms I ($P2_13$), II ($I\overline{2}3$), and III ($P3$). Note the high porosity of the $I\overline{2}3$ 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-R₆ 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.
electron density map suited to completely automated model building (Table S3). These data demonstrate the utility of sclx8 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 sclx8, 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 sclx8 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 sclx8 was sufficient for cocrystallization. This form was distinguished also by its size, yielding millimeter-scale crystals overnight at low sclx8 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-R8 at pH 8.5 in ammonium sulfate (Table 1 and Figure S4).

Crystal form III was solved in space group P3 with one sclx8 per RSL monomer. Similar to the I23 form, protein-calixarene contacts dominate the porous P3 assembly (59% solvent content, pore diameter ~3 nm, Table 2) and there are no protein–protein contacts (Figure 1A). sclx8 adopts a highly puckered conformation, akin to that in form I (P213), but it is substantially more solvent-exposed (Figure 1B and Table S5). Two protein-sclx8 interfaces are involved in this assembly. In the largest interface, 395 Å2 of sclx8 is buried by the encapsulation of Val13 and Lys34. Surprisingly, sclx8 also accommodates Asp32 in a shallow cavity formed by two adjacent phenolic units. The Asp32 side chain is rotated slightly (relative to the P213 and I23 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 Å2 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-Cα) to hydrogen bond with the amide NH of Gly68. This new conformation eliminates a potential steric clash with the bound sclx8 in crystal form III.
NMR Analysis of RSL-sclx₈ Complexation and the pH Trigger. The pH effects identified in the cocryostallization 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 ¹H-¹⁵N HSQC spectrum of RSL was unchanged by titration with mM concentrations of sclx₈, suggesting that RSL and sclx₈ have negligible interactions under these conditions (Figure S9). However, at pH 4.0, specific RSL resonances were perturbed significantly by sclx₈, with saturation occurring at ~10 equiv sclx₈ (Figures 5 and S9). Analysis of the chemical shift perturbations as a function of sclx₈ 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 sclx₈ between two RSL trimers. This inference is corroborated by line broadening evidence. The average ¹H line-widths of RSL increased by ~8 Hz in the presence of sclx₈ (Table S6), consistent with complexation and assembly in solution.

Attempts to characterize RSL-sclx₈ 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 with the results of the cocryostallization 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-sclx₈ 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 sclx₈ 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 sclx₈ concentration. Similar effects occurred for N-terminal residues (Val3, Gln4) and

---

Figure 5. (A) Regions from overlaid ¹H-¹⁵N HSQC spectra of RSL during pH (gray scale) or sclx₈ (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 sclx₈ (at pH 4.0). Dashed lines indicate unassigned resonances due to overlap. (C) Detail of the RSL-sclx₈ P₃ assembly (Figure 2) showing RSL trimers in surface representation and bridging sclx₈ as black spheres or sticks. Significant chemical shift perturbations at 5 mM sclx₈ are highlighted blue.

Figure 6. (A) RSL-sclx₈ 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 ¹H signal at 0 (black) or 5 mM (blue) sclx₈.
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 sclx₈ binding, were affected by the addition of sclx₈ but not by changes in pH. Yet others were strongly affected by pH (e.g., C-terminal Asn90) but less so by sclx₈. 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 undergoes significant conformational changes (Figures 5C and 7B). In particular, the sclx₈-directed assembly appears to be contingent on a structural rearrangement of Asp46, which flips to form a new hydrogen bond with Gly68.

Further evidence for a pH trigger for protein-calixarene assembly was provided by pKₐ analysis of the acidic side chains in the absence and presence of sclx₈ (Figures 6 and S11). The NMR-derived pKₐ values for the acidic residues of RSL were in good agreement with the values calculated in PROPKA3.2 (Table S7). 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. The pH titration curves of Glu28/Glu73, which are completely buried (Table S7) and do not participate in calixarene binding, were a function of Asp46, which makes it a prime candidate for crystal engineering. To this end, 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-sclx₈ 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 sclx₈ behaved in a manner similar to the native protein. Crystals grew within hours and were essentially identical to the native RSL-sclx₈ structure (Tables 1 and S4, Figure S12). Mean isotropic temperature factor comparisons 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 sclx₈ did not cocrystallize. No localized precipitate was observed during the pH-adjustment of RSL-Ac and sclx₈ 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). 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 sclx₈ 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. 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-RSL₄ contains a Met-Lys motif as the N-terminus. In crystal form III, sclx₈ has peripheral interactions with the N-terminus, which is disordered even though it is cationic. Despite the additional Lys, MK-RSL₄ remained soluble in the presence of sclx₈ 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-R₇ and RSL-R₉ deviated distinctly from the behavior of the native protein, forming amorphous precipitates at ≥0.5 equiv sclx₈ (Figure S7). Soluble mixtures of RSL-R₉ were obtained only at 20 equiv

Figure 7. (A) Structural alignment (via chain A) of the main protein-sclx₈-protein assembly in the P₃ crystal forms of RSL (white) and RSL-R₈ (gray). One of the proteins is translated by ∼4 Å in the RSL-R₈ structure. (B) Detail of the protein-sclx₈-protein interfaces with key side chains shown as sticks.
sclx8. Similar to RSLex, these observations suggest that low equiv sclx8 induces aggregation due to the “molecular glue” effect, while high equiv sclx8 yields soluble samples as each protein is encapsulated by the calixarene. Attempts to characterize such assemblies by solution NMR were hindered by the high concentration of sclx8 required and the concomitant ionic strength effects. The pH-dependent precipitation characterized for native RSL occurred similarly for RSL-R6 (Figure S7). The precipitated mixtures at 10 equiv sclx8 and pH 4.0 could be solubilized by adjusting the pH to 4.6 in the case of RSL-R8. However, RSL-R6 remained partly precipitated even at pH 8.0 due to its high pI. While the RSL-sclx8 crystal form III grew within hours, the RSL-R6 mutants exhibited no crystal growth in this condition. Conventional vapor-diffusion experiments in the presence of ammonium sulfate yielded cocrystals. RSL-R6 and RSL-R8 were cocrystallized with sclx8 under near-identical conditions (Table 1) while RSL-R6 resulted in crystal form I (P213, vide supra) and RSL-R8 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-R6, which is highly cationic across the pH range and thus amenable to calixarene-mediated assembly.

Although RSL-R6 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-R6 and RSL are similar in the two structures, the bridging calixarenes have substantially different conformations and binding sites. A more puckered sclx8 occurs bound to RSL-R6 compared to RSL-sclx8 (Figure S13). Interestingly, although similar protein-sclx8 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-R6, 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 fulfills 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 dominate this interface with cation-π and CH-π bonds to the calixarene. The RSL-R6 structure also suggests why RSL-R6 did not crystallize in crystal form III (P3). The calixarene conformation supported by Arg43 in RSL-R6 cannot be maintained by Glu43 in RSL-R8.

![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 sclx8 molecules. The ligands are rendered in blue or white corresponding to the large and small interfaces, respectively. Black triangles indicate C3 planes that connect symmetry-related calixarenes.](https://dx.doi.org/10.1021/jacs.0c10697)

**DISCUSSION**

Anionic calixarenes are well-established “glues” for the assembly and crystallization of cationic proteins such as cytc, lysozyme, and PAF. Here, we have demonstrated the cocryrstallization of “neutral” RSL with highly anionic sclx8. 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 (P23) 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 pKₐ values of two acidic residues (Asp32 and Asp46). sclx8 complexation may also modulate the protonation of Lys34. The disorder of this side chain in the P213 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 sclx8 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 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 porphyrins44 and calixarenes.45 For example, phosphonato-calix[6]arene forms a dimeric disc to mediate a porous assembly of cyt (PDB 5LYC).46 Similar to crystal form II (I23), form III (P3) involves RSL trimers bridged together by six sclx8 ligands arranged in triads on two C3 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 assemblies 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-R8. 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-sclx8 mixtures remain soluble or form crystals ranging from microcrystalline precipitates to millimeter-scale crystals. For example, RSL and 10 equiv sclx8 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 sclx8 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 sclx8 suggests calixarene-coating or encapsulation of the protein.55 At low equiv, the calixarene...
bridges two or more proteins resulting in aggregation. At high equiv sclxn 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 in solution. This behavior is analogous to the reentrant condensation phase behavior of proteins in the presence of multivalent counterions. Protein encapsulation is currently a topic of intense development and macrocycle-masking may be a simple approach to tackle this challenge.

The precipitation of RSL-R6, RSL-R8, 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-R8 and RSLex required pH 4.6 to switch off precipitation while the most cationic variant, RSL-R6, 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. At pH 4.0 and 10 equiv sclxn, the lysine-enriched variant RSLx displayed the same precipitation behavior as the Arg-rich variants, suggesting that the assembly of all three variants was a consequence of higher sclxn affinity rather than reduced protein solubility. Unlike RSLx, soluble mixtures of RSL-R6 and RSL-R8 at high equiv sclxn and low ionic strength yielded no cocrystals. The versatility of guanidinium groups at protein–protein and protein–ligand interfaces is well-established. We reason that the “stickiness” of the Arg-rich variants towards sclxn yielded highly encapsulated protein-calixarene particles incapable of self-association. This hypothesis is supported by cocrystallization of RSL-R8 and RSL-R8 occurring only at high salt concentration wherein counterions compete for sclxn 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, capsids, and crystals. For example, the supramolecular assembly of ferritin cages can be directed by metal-coordination sites. Protein crystals that grow naturally inside living cells have been used to capture cargo including dyes or enzymes. In a particularly striking example, a 32 kDa lipase was trapped within the pores of Cry3Aa crystals that grow in Bacillus thuringiensis cells. The strategies used to engineer frameworks include designed oligomerization, charge-charge interactions, and ligand-mediated assembly. 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. Custom-made bivalent ligands comprising a sugar and rhodamine dye have been used to generate porous frameworks and other assemblies of concanavalin A 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. Here, we demonstrated a pH-triggered framework fabrication of a “neutral” protein with sclxn 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 sclxn confers the crystal with additional binding sites. 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. 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. The data presented here greatly widen the scope for sclxn-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 modulation, well-characterized in protein–ligand binding and in small molecule host–guest systems, 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 under pH control and without the need for precipitants or specialized equipment is attractive for manufacturing processes, including the purification of therapeutic proteins. Furthermore, the controlled nucleation and the formation of microcrystalline samples within minutes has applications in serial crystallography. Complementing these valuable crystallization properties, the sulfur-rich sclxn is a promising phasing agent.

While calixarene-directed assembly is easily applied to proteins in their native state, protein engineering can complement the supramolecular strategy. Here, the permanently cationic Arg-rich variant RSL-R6 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 sclxn. 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 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 Englberger — 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.

## ACKNOWLEDGMENTS

We are grateful for the financial support of NUI Galway, NUI (Travelling Studentship to KOR), Royal Society of Chemistry (Undergraduate Bursary to TS), Swiss National Science Foundation (grant 200021-182369) and Science Foundation Ireland (grants 13/ CDA/ 2168 and 12/ RC/ 2275 P2). We thank SOLEIL synchrotron for beam time allocation, and the staff at beamline PROXIMA-2A for their assistance with data collection. We acknowledge the technical assistance of R. Doohan and B. Harhen.

## 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) Belouqui, 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.; Cavaza, 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.; Kostianen, M. A. Phthalocyanine-virus nanoillers 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.; Iijiri, 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, B. 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, 11598–11603.

(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.; Wierzbiicki, M.; Cabanela, R. W.; Saliba, C.; Motavalli, A.; Tran, N.; Nowick, J. S. Interpreting cubes in the X-ray crystallographic structure of a peptide derived from Medin19. 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.; Feifei, S. C.; Lisdat, F.; Crowley, P. B. Microscale crystals of cytochrome c and calixarenes 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, 9898–9902.

(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.; Grosman, A.; Tezcan, F. A. Hyperexpendable, self-healing macromolecular crystals with integrated polymer networks. Nature 2018, 557, 86–91.

(25) Künze, 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.; Pothukuchi, A.; Golillar, J.; Gerberich, J. C.; Leggere, J. C.; Mastro, 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.
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.; Kumaska, 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) Vounire, 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.; Bankoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Huang, 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 macromolecular 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.; Kamisaki, S.; Li, D.; Howe, N.; Schopp, G.; Pautsch, A.; Schnapp, G.; Pautsch, A.; Raston, C. L.; Crowley, P. B. Measuring the impact of PEGylation on a protein-poly saccharide interaction. Bioconjugate Chem. 2019, 30, 1162–1168.

(61) Sheldrick, G. M. Experimental phasing with SHELXE/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 assembly from crystalline state. J. Mol. Biol. 2007, 372, 774.

(65) Emsley, P.; Lohkamp, W.; Cowtan, K. M. MAP_CHANNELS: a contribution 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) Franken, 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 pKα values in α-synuclein. *Protein Sci.* 2011, 20, 256−269.

(69) Sondergaard, C. R.; Olsson, M. H.; Rostkowski, M.; Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKα 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 pKα value revisited. *Protein Sci.* 2015, 24, 752−761.

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

(73) Basilio, N.; Gago, S.; Parola, A. J.; Pina, F. Contrasting pKα 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 pKα of thiol guests inside yoctoliter containers. *J. Am. Chem. Soc.* 2020, 142, 8291−8298.

(75) Paul, T. J.; Vilecek, 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 pKα 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.; Engelberge, 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.