Self-Assembly of Electrostatic Cocrysalts from Supercharged Fusion Peptides and Protein Cages

Antti Korpi,* Chao Ma,† Kai Liu,‡ Nonappa,§ Andreas Herrmann,** Olli Ikkala,* and Mauri A. Kostiainen*†‡§

†Biohybrid Materials, Department of Bioproducts and Biosystems, Aalto University, FI-00076 Aalto, Finland
‡Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands
§Molecular Materials, Department of Applied Physics, Aalto University, FI-00076 Aalto, Finland

Supporting Information

ABSTRACT: Self-assembly is a convenient process to arrange complex biomolecules into large hierarchically ordered structures. Electrostatic attraction between the building blocks is a particularly interesting driving force for the assembly process, as it is easily tunable and reversible. Large biomolecules with high surface charge density, such as proteins and protein cages, are very promising building blocks due to their uniform size and shape. Assemblies of functional molecules with well-defined nanostructures have wide-ranging applications but are difficult to produce precisely by synthetic methods. Furthermore, obtaining highly ordered structures is an important prerequisite for X-ray structure analysis. Here we show how negatively charged ferritin and viral protein cages can adopt specific cocrysal structures with supercharged cationic polypeptides (SUPs, K72) and their recombinant fusions with green fluorescent protein (GFP-K72). The cage structures and recombinant proteins self-assemble in aqueous solution to large ordered structures, where the structure morphology and size are controlled by the ratio of oppositely charged building blocks and the electrolyte concentration. Both ferritin and viral cages form cocrysal structures with face centered cubic structure and lattice constants of 14.0 and 28.5 nm, respectively. The crystals are porous and the cationic recombinant proteins occupy the voids between the cages. Such systems resemble naturally occurring occlusion bodies and may serve as protecting agents as well as aid the structure determination of biomolecules by X-ray scattering.

Mimicking the highly evolved functionalities of native biomolecules has been in the focus of research efforts, especially over the past decade.1 Besides chemical composition, functionalities of natural systems are typically based on the three-dimensional position of the molecules. Additionally, biomolecules are often large but can still adopt specific hierarchical structures with great selectivity. Production of synthetic materials that could achieve the same level of structural sophistication has, however, been challenging.2

Another way to harvest the designs of nature is to extract the molecules from natural sources and incorporate them into nanostructured materials.3 The restrictions of top-down methods to produce fine-structures can simultaneously be overcome, as many biological molecules form organized systems via self-assembly processes.4 The procedure is the basis of many natural phenomena like protein folding5 and can be used to produce functional materials with well-defined nanostructures.6 Self-assembly is typically carried out in liquid media, which allows the building blocks to diffuse without restraints.7 Noncovalent self-assembly is preferred in many cases as it is typically reversible, easy to control, and applicable to a large pool of molecules, allowing production of assemblies with varying chemical composition and physical dimensions.8 The assemblies can additionally be tuned by chemical modification of the assembling particles or changing the environment of the assembly. Several bottom-up synthesis methods have been recently studied to produce such nanostructured materials.9−15 Practical applications include tissue engineering,16 drug delivery,17,18 catalysis,19 and nanopatterning.20

Protein cages have been utilized as part of self-assembling systems due to their ability to retain functionality while complexed.21−25 They often possess uniform size and shape, making them ideal building blocks for crystalline assemblies.24,26 Many protein cages additionally carry an overall electric charge,27,28 which enables them to assemble via electrostatic interactions. Such assemblies are reversible and responsive to changes in both pH and salinity of the solution, allowing additional control over the system.29−31 To form complexes, the charged particles require counterparts with opposing charge. Polyelectrolytes are a noteworthy option as they possess high charge density.32 They also have the ability to provide proteins and enzymes with additional stability and have therefore been used in delivery systems.33 Copolymers enable even more possibilities for optimizing such systems, as block
copolymers composed of oppositely charged blocks have been reported to further enhance the stability of protein complexes.\(^{34,35}\) However, this method restrains the amount of protein binding with the polymers as only a part of the polymer chains can interact with the protein. Introducing additional charges onto the particles or initially selecting proteins with higher charge densities has been found to be an effective way to increase system stability.\(^{36}\)

We have previously shown that positively charged avidin proteins and negatively charged protein cages can form ordered structures through electrostatic self-assembly.\(^{30}\) These structures could be further functionalized with different biotin-tagged moieties. However, this approach requires an additional biotinylation step. To overcome this, we wanted to study whether fusion proteins that are directly produced with a cationic peptide could be incorporated into the crystals.

In this study, we focus on the self-assembly properties of two native protein cages: apoferritin from *Pyrococcus furiosus* (aFT) and cowpea chlorotic mottle virus (CCMV) from *Vigna unguiculata*. The two cages are complexed with cationic supercharged polypeptides (SUPs) composed of 72 consecutive lysine-containing repeating units (K\(^{72}\)) as well as green fluorescent protein (GFP) produced as a recombinant fusion with the same SUP tag (GFP-K\(^{72}\)). The SUPs are derived from elastin-like polypeptides, consisting of pentapeptide repeats (GVGXP) where the fourth position X was substituted with a lysine (K) residue by molecular cloning.\(^{37,38}\)

![Figure 1](image1.png)

**Figure 1.** Building blocks used in the study. Negatively charged (a) CCMV (pI ≈ 3.8) and (b) aFT (pI ≈ 4.5). Calculated crude vacuum electrostatic potential of the full cages (upper) and solution electrostatic surface potential of protein trimer subunits (lower) are presented for both cages. Red and blue colors represent negative and positive electrostatic potential, respectively. Values range from 0 \(k_B T^{-1}\) (blue) to \(-9 k_B T^{-1}\) (red), where \(k_B\) is the Boltzmann constant, \(T\) absolute temperature, and \(e\) elementary charge. c) GFP-K\(^{72}\) and the chemical structure of K\(^{72}\), where amino acids G (glycine), V (valine), and P (proline) are marked in blue. Both K\(^{72}\) homopolymer and recombinant GFP-K\(^{72}\) were used in this study.

![Figure 2](image2.png)

**Figure 2.** DLS and agarose gel EMSA data: (a) CCMV solution (20 mg L\(^{-1}\)) titrated with K\(^{72}\) and GFP-K\(^{72}\), (b) aFT solution (100 mg L\(^{-1}\)) titrated with K\(^{72}\) and GFP-K\(^{72}\), (c) electrolyte (NaCl) induced disassembly of the CCMV complexes, and (d) electrolyte induced disassembly of the aFT complexes. Volume-average size distribution profiles of CCMV with (e) K\(^{72}\), (f) GFP-K\(^{72}\), and aFT with (g) K\(^{72}\), (h) GFP-K\(^{72}\) at different stages of titration (panels a) and b), respectively). (i) Agarose gel EMSA demonstrating the effect of increasing K\(^{72}\) and GFP-K\(^{72}\) concentration on the electrophoretic mobility of CCMV.
the building blocks, including approximate dimensions and the electrostatic surface potential of the cages are presented in Figure 1. The oppositely charged systems were found to self-assemble in aqueous solution at zero or minor electrolyte concentration, but an excessive addition of electrolyte caused the particles to disassemble back into individual molecules, as expected for electrostatically interacting systems. The size and structure morphology of the assemblies were studied and most of them were crystalline with face centered cubic (fcc) morphology. The presence of GFP appeared to hinder the formation of crystalline assemblies, especially in the case of small aFT cage. It should also be noted that a variety of materials (protein, nanoparticle, synthetic small molecule, etc.) that we have tried to coassemble with protein cages, fail to give ordered structures even after thorough optimization. This provides additional support for the benefits of the SUPs studied in this work.

The self-assembly process was first studied using dynamic light scattering (DLS) by titrating aqueous aFT or CCMV solution with K72 or GFP-K72. The formation of the assemblies was followed by monitoring the scattering count rate and the hydrodynamic diameter ($D_h$). In the case of CCMV, the count rate increases together with the amount of added K72 or GFP-K72 and reached a plateau when $c_{K72}^{CCMV} > 0.5$, indicating the formation of large assemblies in the solution (Figure 2a). Count rate did not decrease even if titration was continued further. With aFT, a distinct difference in the count rate behavior was observed. The count rate increased first to high values, after which it descended quickly until it reached and maintained a constant level when $c_{K72}^{aFT} > 0.55$ (Figure 2b). This indicates the system in question forms a large number of small dense assemblies, which are highly scattering. The small assemblies merge once they pass a critical concentration and the count rate drops even though the size of the assemblies is constantly growing. The formed structures were disassembled by titration with aqueous sodium chloride (NaCl) solution, as a sufficient electrolyte concentration screened the electrostatic interactions between K72 and the protein cages. Both CCMV complexes disassembled uniformly when titrated with NaCl solution (Figure 2c). In the case aFT, a steady decrease in the count rate was observed for K72 complex, but aFT–GFP-K72 underwent an increase in the count rate at the beginning of the NaCl titration (Figure 2d). This indicates that the latter structures did not disassemble uniformly throughout the solution but broke first into numerous smaller assemblies. These small particles also disassembled when $c_{NaCl} > 100$ mM, and the count rate of the system settled to approximately the same values that were measured from the aqueous solution of free aFT.

$D_h$ of the complexes was monitored throughout the K72 and GFP-K72 titrations to follow the increase in particle size.
The complexes were studied (1) at the beginning of the titration, (2) at the concentration where count rate peaked, and (3) at the concentration where the count rate leveled. For CCMV complexes, neither K72 nor GFP-K72 complex grew significantly when using an excess of the protein (Figure 2e,f), suggesting they did not undergo a step with a large number of small particles as aFT complexes did. With aFT, both K72 and GFP-K72 complex kept increasing in size throughout the titration, confirming the hypothesis that the mid titration sharp increase in count rate was due to the amount of the particles, not their size (Figure 2g,h). The final D₅₀ of all four assemblies was close to 1 μm.

The self-assembly of CCMV with K72 and GFP-K72 was further demonstrated using agarose gel electrophoresis mobility shift assay (EMSA). CCMV was complexed with increasing concentrations of the cationic species, causing a loss in electrophoretic mobility as larger assemblies were formed. This was indicated by a tail, which followed the main band containing the smaller and more mobile particles (Figure 2i). The assemblies lost all mobility as the polycation concentration was increased high enough. GFP-K72 complex lost its mobility in lower concentrations than their K72 counterparts, which is in good agreement with the DLS measurements. Zeta potential measurements were conducted to investigate the surface charge of the assemblies. None of the complexes presented significant electrophoretic mobility, indicating surface charge close to zero (Figure S4). The morphology of the formed assemblies was studied using small-angle X-ray scattering (SAXS). The measurements were conducted in 10 mM NaCl solutions. The measured curves for CCMV complexes with both K72 and GFP-K72 (Figure 3a) as well as aFT–K72 complex (Figure 3b) implicated crystalline structures with fcc packing (space group Fm3m; number 225, (hkl) = (111), (200), (220), (311), (420), (422); q/q* = 1, √(4/3), √(8/3), √(11/3)). Face-centered cubic (fcc) structures are typical for aFT systems, but CCMV has been reported to adopt both fcc and body-centered cubic (bcc) configurations. aFT complexed with GFP-K72 was not crystalline, but broad signals were detected at the regions where aFT–K72 showed narrow well-resolved peaks. The assemblies were mostly amorphous and the broad signals were caused by weakly ordered regions. The different morphology explains the difference in the DLS curves between aFT complexes of K72 and GFP-K72, when the complexes were disassembled with NaCl. The amorphous structure of the aFT–GFP-K72 complex is most likely due to the size mismatch of the building blocks. The size of GFP-K72 is too large to fit into the voids between fcc packet aFT particles, which hinders the formation of an ordered structure.

Lattice constants of the K72 complexes with both aFT and CCMV were calculated using a linear fit to the peak positions obtained by SAXS plotted against the quadratic Miller indices of assigned reflections (Figure 3c). The lattice constant (a) for a cubic lattice can be obtained through \( a = 2\pi \sqrt{h^2 + k^2 + l^2} / q_{\text{max}} \) and was calculated to be 40.3 for CCMV–K72 and 19.8 nm for aFT–K72. By using these values, the center-to-center distance (d₁₂) of both complexes was calculated by using the equation \( d_{12} = a / \sqrt{2} \). For CCMV–K72, d₁₂ was 28.5 nm and for aFT–K72 14.0 nm (Figure 3d). These values correspond well with the sizes of aFT and CCMV.

Cryogenic transmission electron microscopy (cryo-TEM) was used to image the nanostructure of the assemblies. As seen in Figure 3, CCMV formed crystalline lattices with both K72 (Figure 3e) and GFP-K72 (Figure 3f). This is evident from the spherical shape and beveled pattern of the assemblies. The individual virus particles, which are ordered into small crystallites, can also be clearly observed. The observed crystallite size varied from approximately 300 nm to 1 μm. As DLS and SAXS measurements suggested, aFT–K72 complexes were crystalline (Figure 3g), whereas aFT–GFP-K72 complexes were amorphous (Figure 3h) and lacked both a distinguishable shape and the beveled pattern.

At optimized conditions, the particles could reach diameters large enough to be imaged using optical microscopy (Figure 3f, inset). The effect of electrolyte concentration on the formed assemblies was studied by preparing salinity series from 0 to 150 mM NaCl in which K72 and GFP-K72 were left to form assemblies with both aFT and CCMV over the course of 10 days at 6 °C. K72 complexes with both aFT and CCMV were the largest ones observed and had a diameter from 30 to 100 μm. The assemblies were heavily branched and irregularly shaped crystals. aFT–GFP-K72 complex had varying shape and size, as expected due to its amorphous morphology. CCMV – GFP-K72 complex formed the clearest crystalline structures. All of the large structures could be disassembled by 100 mM NaCl concentration.

Occlusion body mimicking protection of the complexed GFP was studied using fluorescence spectroscopy. Trypsin, an effective protease, was introduced into solutions of GFP-K72 with CCMV or aFT, and quenching of the fluorescence of GFP was investigated. Without the presence of CCMV or aFT, the fluorescence decreased as trypsin digested the GFP, but in the complexes this was not observed (Figure S6). For crystalline CCMV-GFP-K72 systems this was to be expected, as the GFP moieties were likely to be contained within the crystal lattices, and the used trypsin concentration could not efficiently digest the outer layers of the assemblies within the time frame of the measurement. Interestingly, fluorescence spectroscopy measurements also suggest that the GFP moieties are sufficiently safeguarded in the aFT complexes as well, regardless of the amorphous nature of the systems.

In conclusion, K72 and GFP-K72 underwent reversible self-assembly in aqueous solutions with both CCMV and aFT via electrostatic interactions. Electrolyte concentration that exceeded a critical point (∼100 mM NaCl) screened the interactions and caused the structures to disassemble. The same effect can most likely be achieved by adjusting the pH of the solution. The complexes adopted fcc packed crystalline morphology except for the aFT–GFP-K72 complex, which had an amorphous structure. This is most likely due to steric hindrance caused by GFP and is not present in CCMV complex because the cavities between the protein cages are large enough to house GFP. All of the complexes assembled into macroscopic structures, demonstrating that additional functionalities can be embedded into the systems without preventing self-assembly. Such structures resemble occlusion bodies found in nature and could find potential applications for maintaining the long-term stability of delicate biomolecules.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsmacrolett.8b00023.

Materials, protein details, characterization methods, and instrumentation (PDF)
Versatile Building Block for Bionanotechnology.

REFERENCES

Acknowledgments are acknowledged.

Author Contributions

Mauri A. Kostiainen: 0000-0002-8282-2379
Andreas Herrmann: 0000-0002-8886-0894

Author Information

ACS Macro Letters

AUTHOR INFORMATION

Corresponding Authors
*E-mail: mauri.kostiainen@aalto.fi.
*E-mail: a.herrmann@rug.nl.

ORCID

Nonappa: 0000-0002-6804-4128
Andreas Herrmann: 0000-0002-8886-0894
Olli Ilkala: 0000-0002-0470-1889
Mauri A. Kostiainen: 0000-0002-8282-2379

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work made use of the Aalto University Nanomicroscopy Center (Aalto-NMC) and Bioeconomy premises. Funding from Academy of Finland, Centre of Excellence in Molecular Engineering of Biosynthetic Hybrid Materials (HYBER Center (Aalto-NMC) and Bioeconomy premises. Funding from

REFERENCES


