J|A|C|S
COMMUNICATIONS

Published on Web 11/12/2008

Controlled Assembly of Bifunctional Chimeric Protein Cages and Composition Analysis Using Noncovalent Mass Spectrometry

Sebyung Kang,[†,§] Luke M. Oltrogge,[†,§] Chris C. Broomell,[†,§] Lars O. Liepold,[†,§] Peter E. Prevelige,[⊥] Mark Young,[*,‡,§] and Trevor Douglas[*,†,§]

Department of Chemistry & Biochemistry, Department of Plant Sciences, and Center for BioInspired Nanomaterials, Montana State University, Bozeman, Montana 59717, and Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, 35294

Received October 1, 2008; E-mail: myoung@montana.edu; tdouglas@chemistry.montana.edu

Protein cage architectures, such as viral capsids, ferritins, and heat shock proteins (Hsp), have been extensively used as model systems to study the self-assembly of macromolecular complexes[1−5] and as nanoreactors for materials synthesis.[6−13] However, it is still challenging to manipulate their self-assembly in a controlled way and to analyze their assembled products precisely at the molecular level.[14,15]

In this study, we have generated two different individual mutants of a protein cage with functional groups either inside or outside of the cage (Figure 1A). We chemically modified different cages and reconstructed chimeric cages with a controlled ratio of two subunit types (Figure 1B). Using mass spectrometry, we were able to determine the compositions of the ensemble population and also of the individual chimeric cages within the population at the molecular level. A model based on a binomial distribution suggested chimeric cages are assembled by random incorporation of the two individual subunits.

Mass spectrometry has been used to monitor multicomponent systems, because it can simultaneously resolve individual molecular masses present in a mixture.[16−18] Using a combination of electrospray ionization (ESI)[19] and a time-of-flight (TOF) mass analyzer, it is possible to determine the masses of individual protein components of a noncovalently associated macromolecular complex as well as the mass of the intact macromolecular complex without disturbing the structures.[16,20−23]

The Dps (DNA binding protein from starved cells) from the Gram-positive bacterium *Listeria innocua* (Li) is a member of the ferritin superfamily and prevents oxidative damage of DNA by accumulating iron atoms within its central cavity to produce an iron oxide core similar to that of ferritins.[24,25] The LiDps consists of 12 identical 18 kDa subunits that self-assemble into a hollow protein cage having tetrahedral 23 symmetry (Figure 1A).[24] The LiDps has an outer diameter of 9 nm and an inner cavity diameter of 5 nm with 0.8 nm pores at the 3-fold axis where molecules can pass through to the interior (Figure 1A).[24] The LiDps has been used as a template for nanomaterials synthesis of metal oxides of iron[26] and cobalt[27] as well as cadmium sulfide[28] and platinum[29] with or without modifications. The small number of subunits, robustness at high temperature,[26,27] and intrinsic biomineralizing capability[25] of the LiDps protein cage make it an attractive modifiable nanoreactor for nanomaterials syntheses. In addition, the defined small cavity size[24] allows synthesis of extremely small nanostructured materials.[29]

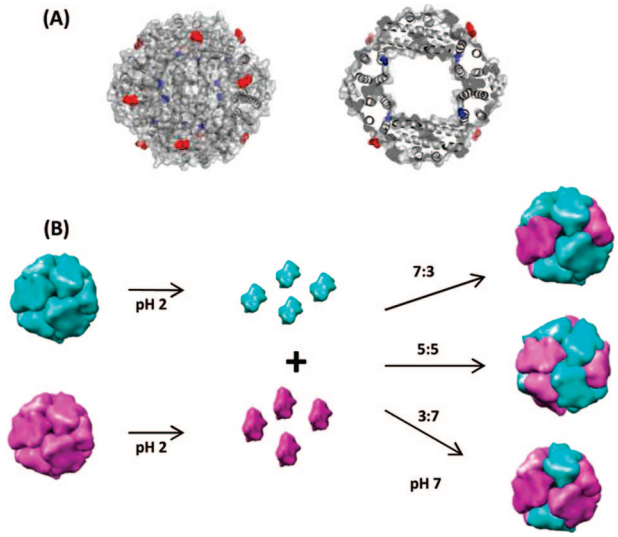

Figure 1. (A) Surface and ribbon diagram representations of LiDps (PDB 1QGH) looking down the 2-fold symmetry axis (left) and a clipped view showing the interior space of the cage (right). Serine 138 residues (blue) and C-termini (red) are indicated. (B) Chimeric cage construction scheme.

Two individual cysteine mutants, one exposed on the interior surface and the other on the exterior surface, were generated to adapt the LiDps for selective chemical modifications. The serine residue at position 138 located in the middle of helix E where it is directed toward the inside cavity was substituted with cysteine (S138C)[29] (Figure 1A, blue). Alternatively, for the exterior modification, four extra residues (KLFC) were added to the C-terminus (Figure 1A, red). Cysteine reactive maleimide and iodoacetamide agents were used to conjugate chemical moieties to the LiDps. We conjugated a metal chelating chemical moiety (5-iodoacetamido-1,10-phenanthroline (iodo-phen)) on the interior (S138C) and a chromophore (fluorescein-5-maleimide (F5M)) on the exterior (KLFC) surfaces of the LiDps protein cage. Chemically modified mutants eluted on size exclusion chromatography (SEC) at the same position as wild-type (wt) LiDps suggesting chemical modifications of the mutant cages did not alter protein architectures (Supporting Information(SI), Figure F1). While the hydrodynamic diameter of phen-S138C LiDps was the same as unmodified or wt LiDps (9 nm), that of F5M-KLFC LiDps was slightly larger (10.7 nm) than wt LiDps probably due to mass addition to the exterior (extra four residues and fluorescein, ∼1 kDa) and the flexible nature of the C-terminus (SI F2).[24]

The extent of iodo-phen and F5M modification to each protein cage was evaluated by mass spectrometry.[17,30] Mass analysis

[†] Department of Chemistry and Biochemistry, Montana State University.
[‡] Department of Plant Science, Montana State University.
[§] Center for BioInspired Nanomaterials, Montana State University.
[⊥] University of Alabama at Birmingham.

10.1021/ja807655t CCC: $40.75

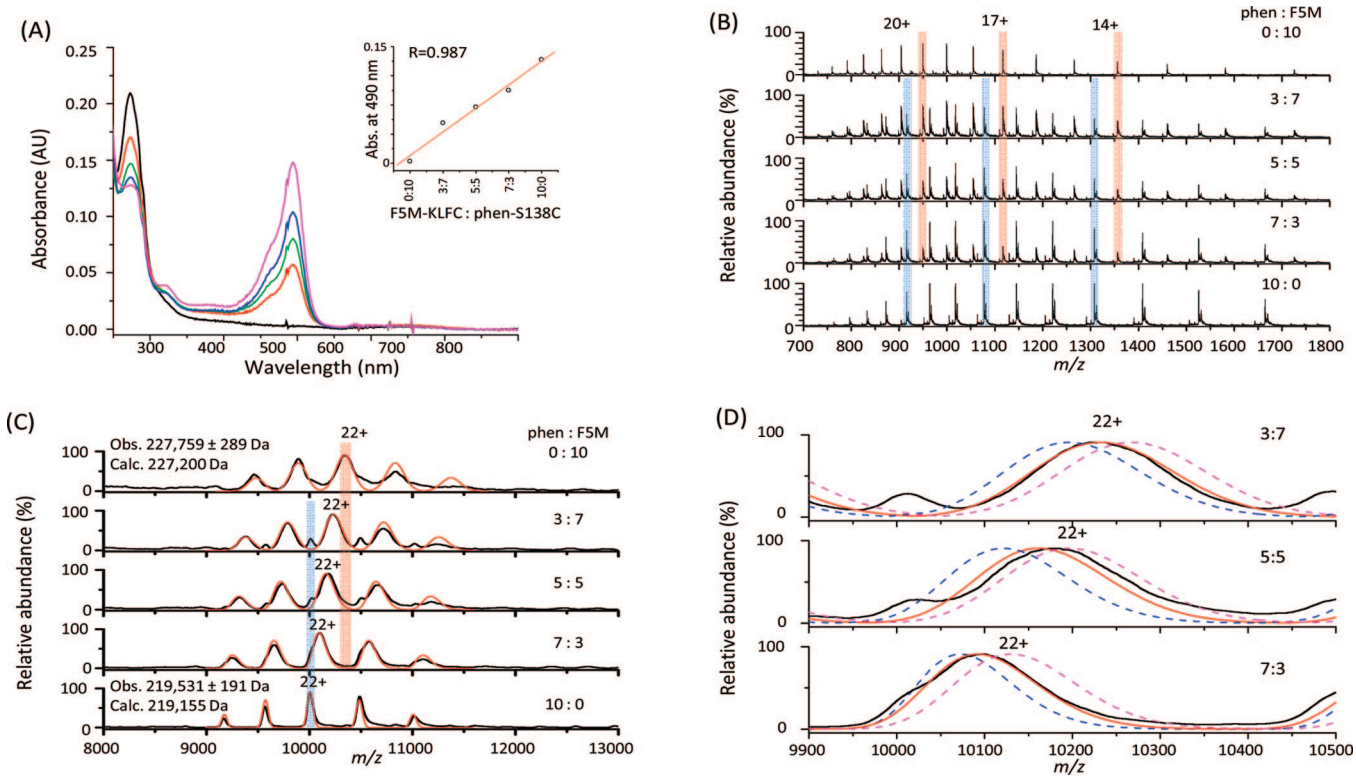

Figure 2. (A) UV−vis spectra of reassembled LiDps cages according to the initial mixing ratios. Phen-S138C/F5M-KLFC = 0:10 (magenta), 3:7 (blue), 5:5 (green), 7:3 (red), and 10:0 (black). Absorbances at 490 nm were plotted against the input ratio of phen-S138C and F5M-KLFC LiDps subunits (inset). Mass spectra of reassembled subunits (B) and whole cages (C). Charged peaks of subunits (14, 17, and 20+) and cages (22+) are indicated and marked by blue (phen-S138C) and red (F5M-KLFC) transparent bars. The experimental data were modeled (red lines) using a binomial distribution model[34] (SI, experimental section). (D) Comparison of modeled mass spectra generated by varying the composition of F5M-KLFC. Exact input amount (solid red), 10% higher (dashed magenta), and 10% lower (dashed blue) amounts.

revealed that 10 and 11 subunits per cage (12 subunits) of S138C and KLFC mutants were chemically modified with phen (phen-S138C) and F5M (F5M-KLFC), respectively (SI F3).

To construct chimeric LiDps with distinct internal and external functionalities in well-defined ratios, we dissociated both genetically and chemically modified LiDps by lowering the pH.[31] Dissociated subunits were separated by SEC and analyzed by DLS, mass spectrometry, and analytical ultracentrifugation to determine oligomeric states of dissociated cages (SI F4). Mass spectrometry was carried out at pH 2.0 and both dissociated phen-S138C and F5M-KLFC LiDps were detected predominantly as monomers (SI F4C). However, it is possible that modified LiDps subunits are weakly associated to each other in solution at pH 2.0 and become dissociated inside the mass spectrometer, because it is necessary to use higher temperature and voltage than physiological conditions to ionize analytes and to apply external energy to the ionized analytes allowing them reach to the detector[20,32] (SI, experimental section). Therefore, analytical ultracentrifugation was employed to probe the oligomeric state of the dissociated cages in solution. While the previously reported *s* value of intact LiDps cage is 10.5 *s*,[33] we obtained sedimentation coefficients of 1.6 *s* and 1.8 *s* from dissociated phen-S138C and F5M-KLFC LiDps, respectively, with a minor population sedimenting at approaximately 2.3 *s* (SI F4D). Chemically oxidized KLFC LiDps subunits, which form covalent dimers (SI F4E), run in parallel sedimented at 2.3 *s* consistent with a 40 kDa protein (SI F4D). Taken together, the mass spectrometry and velocity sedimentation data suggest that both dissociated phen-S138C and F5M-KLFC LiDps exist predominantly as monomers with a small amount of dimers in solution at pH 2.0.

Chimeric LiDps protein cages were constructed by mixing individually modified and dissociated subunits in various desired ratios and reassembled by slowly raising the pH to 7 (Figure 1B). The initial mixing ratio was determined by a combination of UV−vis spectra and relative intensities of subunit mass spectrometry analyses. Reassembled chimeric cages were isolated by SEC and characterized by UV−vis spectroscopy and mass spectrometry. The majority of the sample eluted as intact cages, although some large aggregates were observed (∼15%, SI F5). The absorbance at 490 nm, which is the absorption maxima of fluorescein, increased linearly as the proportions of F5M-KLFC LiDps subunits increased (Figure 2A, inset) indicating mixed incorporations of F5M-KLFC and phen-S138C subunits. The subunit composition was also analyzed by mass spectrometry (Figure 2B). The relative intensities of peaks assigned to the F5M-KLFC subunits (red transparent bars) increased as the input proportion of F5M-KLFC subunits increased (Figure 2B). Both UV−vis spectra and mass spectrometric analyses of subunits agreed well with the initial mixing ratios. However, these approaches provide information only about subunit composition in the whole population rather than for the individual cages. Therefore, it is possible that each subunit type reassembles to form two individual phen-S138C and F5M-KLFC homododecameric cages instead of reassembling into chimeric cages with a given mixing ratio.

Noncovalent mass spectrometric analyses of reassembled cages allowed us to distinguish between these two possibilities. Increased ratio of F5M-KLFC subunits, the heavier one, resulted in shifts of charge state distribution to the higher *m*/*z* (Figure 2C, black lines) clearly demonstrating mixed incorporation of both phen-S138C and

F5M-KLFC subunits within the population of reassembled cages. To rule out the possibility of subunit exchange between cages as has been observed with Hsp,[34,35] mixed populations, without dissociation, were monitored by noncovalent mass spectrometry. Two distinct masses corresponding to each homododecameric cage were consistently detected with different intensity ratio according to the amount of the initial input (SI F6) suggesting subunit mixing only occurs during dissociation/reassembly process.

The experimental mass spectra were modeled using a modified binomial distribution model[34] (SI, experimental section) to test whether the formation of chimeric cages is achieved by the random incorporation of two initial components. In general, a binomial distribution describes the probability of observing a specific composition for an ensemble composed of two different species if the ensembles are assembled randomly. In our case, the ensemble represents each cage structure which is a collection of 12 subunits and the two species are phen-S138C and F5M-KLFC subunits. The probabilities of observing a protein complex with each hybrid composition were calculated and then used to scale the area of Gaussian functions whose peak widths were calculated as a linear combination of the input complexes' peak widths (SI F7). This approach allowed us to determine the contributions of each particular chimeric cage in the whole population (SI T1). At the initial mixing subunit ratios of 7:3 and 3:7 (phen-S138C:F5M-KLFC), 47% of each population was contributed by hybrid species of 9 subunits of phen-S138C and 3 subunits of F5M-KLFC (9 + 3, 24.0%) and 8 + 4 (23.1%), and 3 + 9 (24.0%) and 4 + 8 (23.1%), respectively (SI F7 and T1).

Theoretical mass spectra were generated by integrating all the calculated probability-weighted Gaussian functions (Figure 2C,D, red lines). The experimental mass spectra of reassembled chimeric cages showed an excellent agreement with theoretical spectra (Figure 2C,D) suggesting that the formation of chimeric cages can be well described by a simple binomial distribution where the assembly unit is a monomer.[35] The uniqueness of the theoretical spectra was tested by varying the composition ±10% from the actual experimental input composition (Figure 2D, SI F8). These results suggest that both phen-S138C and F5M-KLFC LiDps subunits incorporate randomly into a chimeric cage obeying the initial mixing ratios instead of making two individual cage populations (Figure 2). However, a subpopulation corresponding to the phen-S138C only dodecameric cage was also observed (Figure 2C). Dissociated phen-S138C subunits have a slightly greater tendency to form dimers (SI 4D) and small amounts of preformed dimers may facilitate homododecameric cage formation.

In conclusion, we have constructed chimeric protein cages having dual functionalities inside and outside of LiDps by reassembling dissociated subunits with desired ratios and determined compositions of chimera at the molecular level using mass spectrometry. Binomial distribution analysis of mass spectra supported that dissociated subunits reassemble randomly into a dodecameric cage and allowed us to determine the contributions of each chimeric cage in the population. The disassembly/reassembly approach described here may give us greater insight into constructing finely tuned multifunctional protein cages and its biomedical and nanomaterials applications.

Acknowledgment. The authors wish to thank Dr. Peter Suci for critical discussion. This research was supported in part by grants from the National Science Foundation (Grant CBET-0709358), Office of Naval Research (Grant N00014-03-1-0692), the Department of Energy (Grant DE-FG02-07ER46477), and Human Frontiers of Science Program (Grant RGP61/2007).

Supporting Information Available: Details on experimental procedures and data of SEC, DLS, mass spectra simulation, and mass spectrometry analyses of chemically modified mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Casjens, S. *Virus Structure and Assembly*; Jones and Bartlett: Boston, MA, 1985.
(2) Johnson, J.; Speir, J. *J. Mol. Biol.* **1997**, *269*, 665–675.
(3) Prevelige, P. E., Jr.; King, J. *Prog. Med. Virol.* **1993**, *40*, 206–21.
(4) Steven, A. C.; Heymann, J. B.; Cheng, N.; Trus, B. L.; Conway, J. F. *Curr. Opin. Struct. Biol.* **2005**, *15*, 227–36.
(5) Wiedenheft, B.; Mosolf, J.; Willits, D.; Yeager, M.; Dryden, K. A.; Young, M.; Douglas, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10551–10556.
(6) Douglas, T.; Young, M. *Nature* **1998**, *393*, 152–155.
(7) Douglas, T.; Young, M. *Science* **2006**, *312*, 873–5.
(8) Lewis, J. D.; Destito, G.; Zijlstra, A.; Gonzalez, M. J.; Quigley, J. P.; Manchester, M.; Stuhlmann, H. *Nat. Med.* **2006**, *12*, 354–60.
(9) Klem, M. T.; Willits, D.; Solis, D. J.; Belcher, A. M.; Young, M.; Douglas, T. *Adv. Funct. Mater.* **2005**, *15*, 1489–1494.
(10) Mann, S.; Ozin, G. A. *Nature* **1996**, *382*, 313–318.
(11) Uchida, M.; Klem, M. T.; Allen, M.; Suci, P.; Flenniken, M.; Gillitzer, E.; Varpness, Z.; Liepold, L. O.; Young, M.; Douglas, T. *Adv. Mater.* **2007**, *19*, 1025–1042.
(12) Flenniken, M. L.; Liepold, L. O.; Crowley, B. E.; Willits, D. A.; Young, M. J.; Douglas, T. *Chem. Commun.* **2005**, 447–9.
(13) Klem, M. T.; Resnick, D. A.; Gilmore, K.; Young, M.; Idzerda, Y. U.; Douglas, T. *J. Am. Chem. Soc.* **2007**, *129*, 197–201.
(14) Gillitzer, E.; Suci, P.; Young, M.; Douglas, T. *Small* **2006**, *2*, 962–966.
(15) Klem, M. T.; Willits, D.; Young, M.; Douglas, T. *J. Am. Chem. Soc.* **2003**, *125*, 10806–10807.
(16) Benjamin, D. R.; Robinson, C. V.; Hendrick, J. P.; Hartl, F. U.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 7391–7395.
(17) Kang, S.; Lander, G. C.; Johnson, J. E.; Prevelige, P. E. *Chembiochem.* **2008**, *9*, 514–518.
(18) McKay, A. R.; Ruotolo, B. T.; Ilag, L. L.; Robinson, C. V. *J. Am. Chem. Soc.* **2006**, *128*, 11433–11442.
(19) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64–71.
(20) Benesch, J. L.; Robinson, C. V. *Curr. Opin. Struc. Biol.* **2006**, *16*, 245–251.
(21) Bothner, B.; Siuzdak, G. *Chembiochem* **2004**, *5*, 258–60.
(22) Heck, A. J. R.; van den Heuvel, R. H. H. *Mass Spectrom. Rev.* **2004**, *23*, 368–389.
(23) van den Heuvel, R. H.; Heck, A. J. *Curr. Opin. Chem. Biol.* **2004**, *8*, 519–526.
(24) Ilari, A.; Stefanini, S.; Chiancone, E.; Tsernoglou, D. *Nat. Struct. Biol.* **2000**, *7*, 38–43.
(25) Stefanini, S.; Cavallo, S.; Montagnini, B.; Chiancone, E. *Biochem. J.* **1999**, *339*, 775–775.
(26) Allen, M.; Willits, D.; Mosolf, J.; Young, M.; Douglas, T. *Adv. Mater.* **2002**, *14*, 1562–1565.
(27) Allen, M.; Willits, D.; Young, M.; Douglas, T. *Inorg. Chem.* **2003**, *42*, 6300–6305.
(28) Iwahori, K.; Enomoto, T.; Furusho, H.; Miura, A.; Nishio, K.; Mishima, Y.; Yamashita, I. *Chem. Mater.* **2007**, *19*, 3105–3111.
(29) Kang, S., Lucon, J., Varpness, Z. B., Liepold, L., Uchida, M., Willits, D., Young, M., Douglas, T. Angew. Chem., Int. Ed. **2008**, *47*, 7845–7848.
(30) Kang, S.; Hawkridge, A. M.; Johnson, K. L.; Muddiman, D. C.; Prevelige, P. E. *J. Proteome Res.* **2006**, *5*, 370–377.
(31) Chiaraluce, R.; Consalvi, V.; Cavallo, S.; Ilari, A.; Stefanini, S.; Chiancone, E. *Eur. J. Biochem.* **2000**, *267*, 5733–5741.
(32) Chernushevich, T. B., IV. *Anal. Chem.* **2004**, *76*, 1754–60.
(33) Ilari, A.; Latella, M. C.; Ceci, P.; Ribacchi, F.; Su, M. H.; Giangiacomo, L.; Stefanini, S.; Chasteen, N. D.; Chiancone, E. *Biochemistry* **2005**, *44*, 5579–5587.
(34) Sobott, F.; Benesch, J. L.; Vierling, E.; Robinson, C. V. *J. Biol. Chem.* **2002**, *277*, 38921–9.
(35) Painter, A. J.; Jaya, N.; Basha, E.; Vierling, E.; Robinson, C. V.; Benesch, J. L. P. *Chem. Biol.* **2008**, *15*, 246–253.

JA807655T