Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Natural and artificial protein cages: design, structure and therapeutic applications

Jonathan Gardiner Heddle[1,*], Soumyananda Chakraborti[1,*] and Kenji Iwasaki[2]

Advanced electron microscopy techniques have been used to solve many viral capsid structures. The resulting detailed structural knowledge contributes to understanding of the mechanisms of self-assembly, maturation pathways and virion–host cell interactions. It also acts as inspiration for design and production of capsid-like artificial protein cages. Both natural and artificial cages have potential uses in medicine including as vaccines and in drug delivery. For vaccines, virus-like particles formed only from outer virion shells, lacking genetic material, offer the simplest basis for development, while encapsulation of target molecules inside protein cages is potentially more challenging. Here we review advances in cryo-electron microscopy with particular reference to viral capsid structures. We then consider why knowledge of these structures is useful, giving examples of their utilization as encapsulation and vaccine agents. Finally we look at the importance of structural techniques including cryo-EM in the rapidly progressing field of designed protein cages.

Addresses
[1] Bionanoscience and Biochemistry Laboratory, Malopolska Centre of Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland
[2] Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Japan

Corresponding author: Iwasaki, Kenji (ikenji@protein.osaka-u.ac.jp)
[*] Contributed equally.

Current Opinion in Structural Biology 2017, **43**:148–155

This review comes from a themed issue on **Macromolecular assemblies**

Edited by **Toshiyuki Shimizu** and **Florence Tama**

For a complete overview see the Issue and the Editorial

Available online 27th March 2017

http://dx.doi.org/10.1016/j.sbi.2017.03.007

Introduction

In nature, spherical protein cages fulfill a number of important roles. Perhaps the most well known cages are the virus capsids, macromolecular machines which act as protective containers and delivery vehicles for viral genetic material [1]. Capsid proteins are the main constituents of the outer shell of viruses (which may also include membrane components) forming the interface between viruses and host cells and playing a vital role in cell invasion. Therefore the study of virus protein cages, in particular their structures, is of importance for understanding pathogenicity and, possibly, for producing therapeutics that inhibit capsid–host interactions to limit virus entry. Cryo-electron microscopy (cryo-EM) has proved a vital tool for understanding capsid structures and continues to advance in terms of speed, capability and achievable resolution.

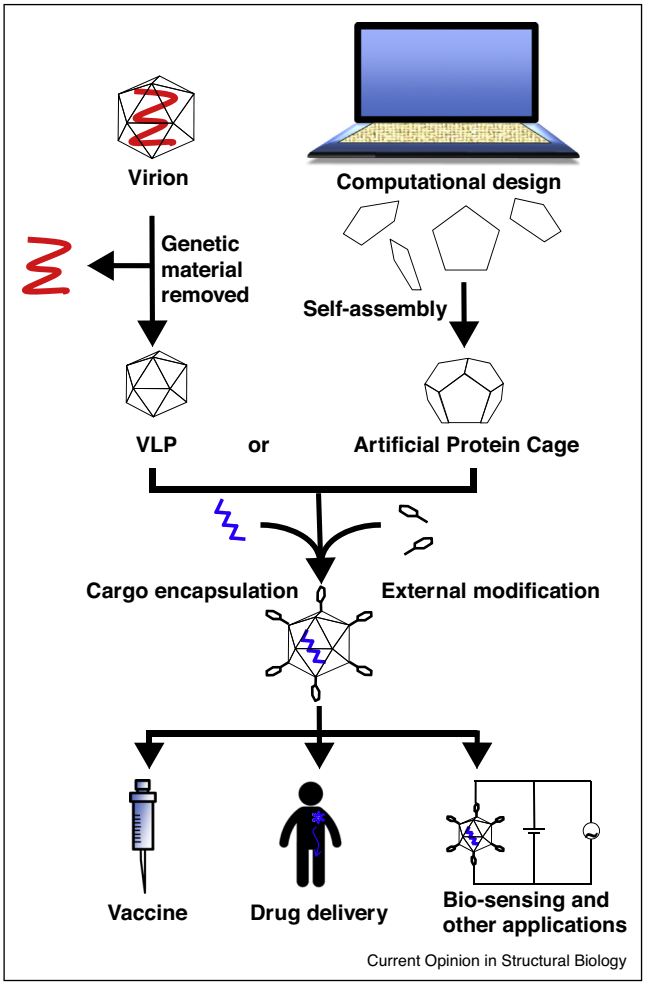
Virus cages also have applications apart from their natural role: Non-infectious particles can be constructed from viral structural proteins produced using recombinant technology to yield virus-like particles (VLPs). These can self-assemble to produce an exact reproduction of a viral capsid, devoid of genetic material. The outer surface of such VLPs can be modified either chemically or genetically to produce vaccines with a large number of attachment points to which antigens can be connected, resulting in high immunogenicity. VLPs have been utilized as vaccines both clinically and experimentally [2,3]. Clinically, HEPTAVAX-B (Merck) was the first commercial vaccine developed using VLPs, and the HPV vaccines Gardasil1 (Merck and Co., Inc.) and Cervarix1 (GlaxoSmithKline) have also been launched. VLP cavities can also be filled and their natural cell penetrating abilities taken advantage of in order to deliver cargoes. These may be genetic (gene therapy [4]) or a host of other substances including drugs, inorganic materials and even functional proteins [5].

Finally, synthetic biology approaches are increasingly allowing us to design bespoke protein cages with desired properties. This represents the frontier of protein cage research and offers the prospect of a widening of protein cage capabilities to include new structures with novel medical applications. A visual summary is shown in Figure 1.

Advanced cryo-EM for understanding protein cage structures

Producing symmetrical and spherical structures from multiple copies of the same protein is an efficient way to build a container to protect a genome if the genome itself is small and so can only encode a limited number of proteins. Many viruses adopt this idea to build a protein cage (known as a capsid) with the most common capsid

Figure 1



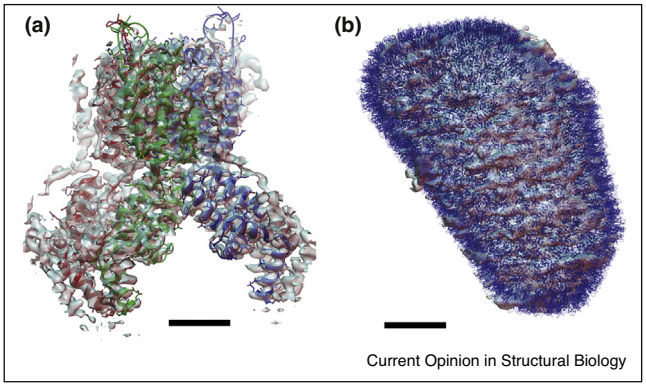
Production, modification and application of natural and artificial protein cages. Virus-like particles (VLPs) are generated by removing genetic material, making them non-infectious. However, they retain the antigenicity of the native virus, meaning that they remain useful as vaccines. Non-antigenic VLPs can be converted to useful vaccines by attachment of antigens from other viruses. Artificial protein cages can be designed from component proteins *in silico* and then assembled *in vitro*. Both VLPs and artificial cages can be used to encapsulate useful molecules (*e.g.*, therapeutics) and can be surface-modified with antigens, cell-targeting/penetrating groups or with groups that may provide a detectable signal depending on interaction with the environment. Such assembled VLPs and cages have applications as vaccines, drug delivery systems and bio-sensors, amongst others.

geometry being icosahedral. These icosahedral virions have been used as a target of single-particle reconstruction since the early days of the technique because the produced images are easy to process. In fact, the first near-atomic resolution structure exploiting single-particle reconstruction was achieved using icosahedral viruses [6–9]. The first *de novo* model building was also carried out using an icosahedral virus [6]. The advantages of using cryo-EM instead of X-ray crystallography is that a small amount of virus particles is often enough to achieve 3D structures. Rhinovirus C is a good example [10,11] where, in spite of a relatively low concentration of the virions, a 2.8 Å resolution structure was obtained by cryo-EM and single-particle reconstruction. As a result of not requiring a potentially time-consuming crystallization step, timely structural analysis of emerging viruses such as Zika virus becomes possible [12]. This will allow important viruses in terms of pathology to be studied in more detail. For instance, differences between genogroups can be obtained, leading to vaccine designs for viruses with various kinds of serotype.

Whereas the high-resolution structures of icosahedral and non-enveloped viruses have been successfully pursued, the next and more formidable challenges are the non-symmetrical virus particles such as enveloped viruses. For example, the whole structure of herpes simplex virus type 1 (HSV-1) was elucidated using cryo-electron tomography (cryo-ET) instead of single-particle reconstruction [13]. HSV is composed of three major components, a glycoprotein-containing envelope, a proteinaceous layer between the capsid and the envelope called the tegument, and a nucleocapsid. Cryo-ET was essential for understanding the outline structure of HSV because the components (excepting the icosahedral nucleocapsid shell [14]) are non-symmetrical.

Recent advanced cryo-ET has revealed very high resolution structures of non-symmetrical virus particles such as enveloped viruses. Among them, a historical event is that of a subtomogram averaging technique reaching a resolution of 3.9 Å, which was reported for the atomic model of the capsid domain and spacer peptide 1 of Gag polyprotein of HIV-1 in the immature state [15••] (Figure 2a). This is notable as most of the structures that had been obtained by subtomogram averaging up until this point were at a resolution of around 20 Å at best, whereas this work succeeded in building an atomic model. Here, a brand-new method to acquire tomographic tilt-series; a dose-symmetric tilt scheme developed by the authors, was used. With the development of direct electron detectors and appearance of well-controlled electron microscopes equipped with a stable stage, for example, Titan KRIOS (FEI), the method of acquiring images has improved. In fact, the implementation of the direct electron detector is a major factor in the recent dramatic increase in near atomic resolution structures accomplished using single-particle electron cryomicroscopy [16]. Before its appearance, charge-coupled device (CCD) sensors having phosphor screens with fiber-optic coupling were popular. These type of cameras are more convenient for the microscopist to check the examined specimen compared to photographic film. However, since the detective quantum efficiency (DQE) that represents the loss of input signal of a detector due to noise or the signal conversion systems in the detector, is worse than photographic film at high frequency, they are not

Figure 2



(a) The CA-SP1 atomic model (PDB 5L93) is visualised with a 3D map of CA-SP1 (EMD-4015) using UCSF Chimera [51]. The script file used for UCSF Chimera was downloaded from EM-Navigator [52] (http://pdbj.org/emnavi/?&lang=en). Scale bar, 2 nm. **(b)** HIV-1 capsid model produced using a computer-assisted hybrid approach. The atomic model, composed of 186 hexamers and 12 pentamers (PDB-3j3y), was fitted into a 3D map obtained by cryo-ET (EMD-5639) and is visualized by UCSF Chimera following the same method as for (a). Scale bar, 20 nm.

particularly suitable for high resolution work. This is because converting electron signals to photons via a phosphor scintillator and transmitting the photons to the CCD sensors thorough fiber optics results in blurring, especially in the scintillator. On the other hand, direct electron detectors, as the name suggests, can detect incident electrons directly without conversion to light, leading to higher DQEs in either high or low frequency regions compared to CCD cameras and photographic film. In addition, high-speed readout enables us to correct the beam-induced motion that is a phenomenon notorious for causing significant blurring of the recorded images. As a result, direct electron detection is presently the gold standard detection method for near-atomic resolution work [17,18].

The retroviral virions like HIV commonly change their shape largely through their maturation pathway. In immature virions, a quasi-spherical shell composed of Gag is formed. However, size and morphology is not uniform. Therefore, cryo-ET and the subtomogram averaging technique is essential to achieve a high-resolution structure. Even in the mature virion, the capsid core is not symmetrical but shows a cone-like shape, which is revealed by cryo-ET [19]. This shape is formed by the 216 hexamers and 12 pentamers composed of six and five capsid proteins respectively. The non-symmetrical and cone-like shell shape is related to the locations of the 12 pentamers [20]. Here, a computer-assisted hybrid approach was used to obtain an atomic model of the mature capsid shell using X-ray crystallography, NMR, cryo-ET and molecular dynamics [21••] (Figure 2b). This integrative modeling approach certainly will become mainstream thanks to the need to acquire more complicated and non-symmetrical capsid structures and the increasing sophistication and resolution of the techniques allows us to progressively modify and exploit VLPs, in particular in the medical field.

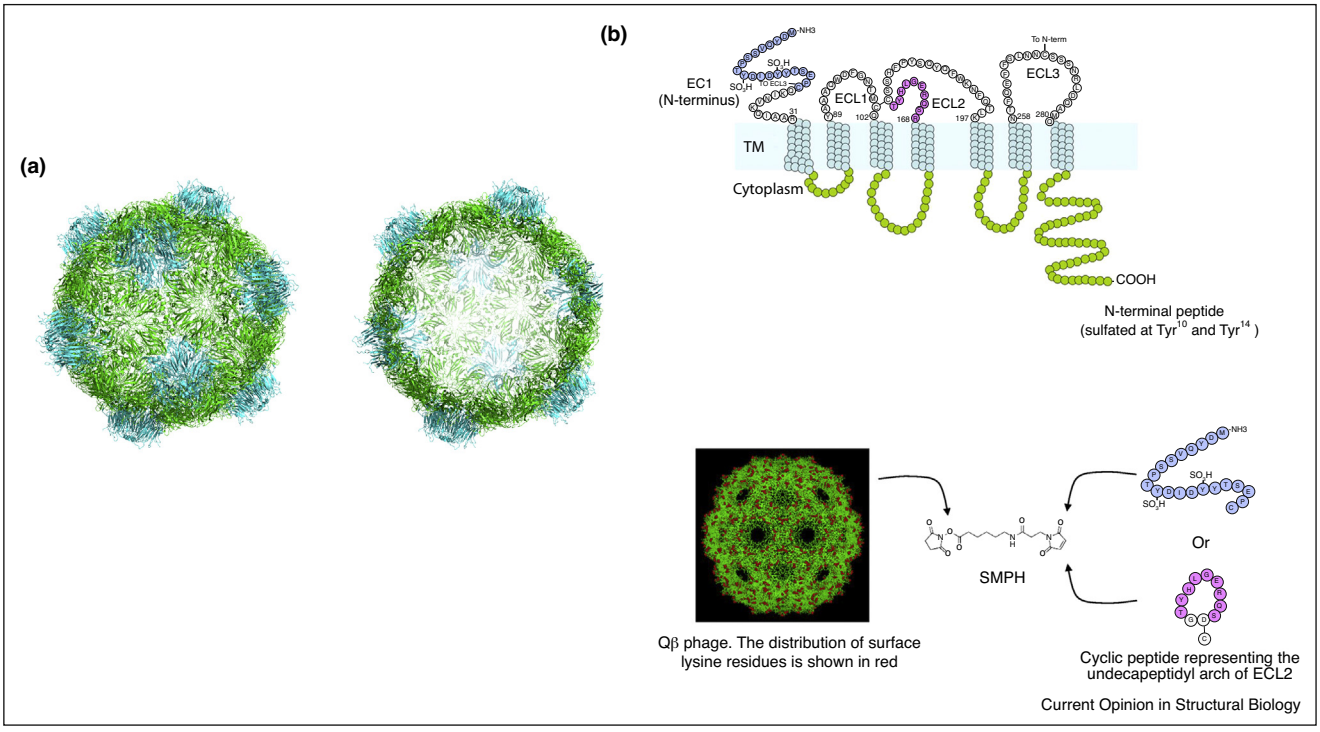
Encapsulation and VLP vaccines: exploiting protein cages in medicine

For medical applications VLPs are attractive as they are devoid of any kind of nucleic material, and normally are non-infectious. They can be derived either from viral sources or synthetically from non-viral cage forming systems by heterologous expression [22••]. When free of the natural nucleic acid contents, the VLP cage interior provides an appealing compartment for non-native cargo encapsulation (Figure 3a). VLPs gained attention because of their large cargo loading abilities, biocompatibility, *in-vivo* targeting and stability [22••,23,24] and they are often able to withstand harsh conditions required for chemical modification which can be carried out both internally and externally. The most common modification technique employs genetic engineering, which is highly effective for insertion, replacement or deletion of residues, allowing systematic addition of functional groups. One of the best known examples is mutation to cysteine, which can be used for disulfide linkages, association with metal, and bio-conjugation with thiol-selective moieties [25]. Insertion of unnatural amino acids is also possible using this technique [26].

At present a number of VLP-based drugs are in clinical trials: T-VEC manufactured by Amgen was recently approved for oncotherapy against melanoma [27]. Other promising roles for VLPs as cargo carriers include in tissue/cell specific delivery of contrast and imaging agents, mainly in the field of magnetic resonance imaging (MRI), and positron emission tomography (PET) [28]. VLPs also have great potential as vaccines: Traditional vaccines have been used against a range of diseases, often highly successfully with notable examples including diphtheria, polio and mumps. However, there have been failures and challenges such as HIV, influenza, Ebola and some cancers. One of the major problems is to elicit robust and persistent immune responses, and VLPs are a promising platform to achieve this as they are capable of multivalent antigen presentation. For this, the three-dimensional structure must be known in some detail in order to decide where the antigens can/should be attached. One example of successful structure-based design is Qβ VLP carrying a peptide based on CCR5 (a cytosolic self-replicating protein functioning as a co-receptor for replication and pathogenesis of HIV; Figure 3b) which boosted adaptive immunity in rhesus monkeys [29].

Cryo-EM along with X-ray crystallography, has contributed to this field, by providing fine structural details of

Figure 3



(a) Example of an icosahedral virus protein cage. The crystal structure of cowpea mosaic virus, widely exploited in bionano science, is shown (PDB-5fmo [53]). The structure is shown in cartoon format with each copy of the constituent large and small proteins shown in green and cyan respectively. The exterior of the VLP is shown on the left and revealing the interior cavity on the right. The particle is approximately 28 nm in diameter. **(b)** Preparation and conjugation strategy of two CCR5 (a cytosolic self-replicating protein functioning as a co-receptor for replication and pathogenesis of HIV) vaccine constructs with Qβ virus. A linear peptide representing the N terminus (EC1) and a cyclic peptide representing the undecapeptidyl arch of the second extracellular loop (ECL2) of macaque CCR5 (top), representing two different regions that interact with HIV gp120 during infection, were synthesized and then conjugated at high density to Qβ bacteriophage using the bifunctional cross-linker SMPH to obtain Qβ.EC1 and Qβ. ECL2 particles, respectively (bottom). Adapted by permission from reference [36], copyright American Society of Microbiology, 2014.

VLP-based vaccines. The structure of the papillomavirus VLP, which constitutes the HPV vaccine, has been solved using cryo-EM to 3.6 Å resolution [30,31]. In a separate study, cryo-EM structures determined from a VLP based on enterovirus 71 (EV71) and coxsackievirus A16 (responsible for hand, foot, and mouth disease) at ~5 Å resolution provided a structural basis for the development of an improved vaccine candidate [32]. Recently a cryo-EM structure of chikungunya virus-antibody complex has also been determined to 5.3 Å resolution, which helped in determining the details of how the antibody neutralizes the virus [33]. Vaccines against two other important viruses use recombinant VLPs: hepatitis B virus and hepatitis E virus. Interestingly, in both these cases, cryo-EM has been used for the acquisition of high-resolution structural data, allowing progress in understanding structure–function relationships and demonstrating a major role for the technique in development of therapeutics [34]. It is also proving highly useful in the new field of artificial protein cage production.

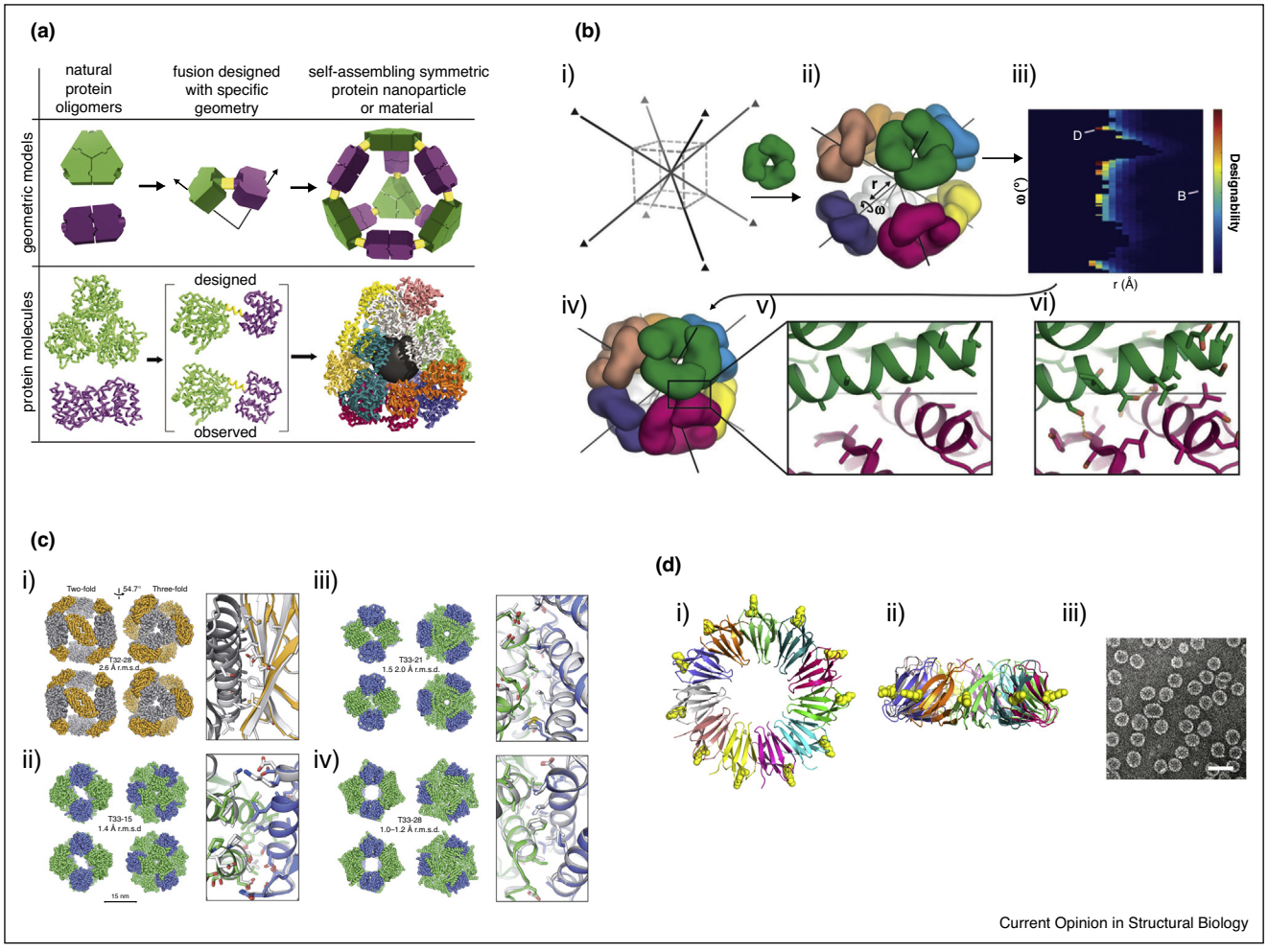
Artificial protein cages and their potential therapeutic uses

Artificial virus-like cages are currently under investigation using a wide variety of building block materials [35]. Synthetic biology approaches allow us to design and produce artificial protein cages, permitting the production of structures with bespoke properties free from evolutionary constraints. Cages produced in this way could have a number of potential uses including as drug delivery systems or artificial vaccines and may be desirable because their structural features can be more precisely controlled. In recent years advances in computational and experimental techniques have resulted in a number of artificial cage proteins being designed, produced and investigated structurally [36,37••,38,39,40••,41,42]. However, there are still constraints: The large size of artificial cages means that, due to limitations inherent in protein expression and folding, they, like their natural counterparts, must be constructed of large numbers of identical, individual protein subunits arranged with high symmetry.

This can present challenges in their design. The approaches used to overcome these challenges have been reviewed in depth elsewhere [37••,41,43•] and are covered briefly here.

The role of cryo-EM in artificial protein cage production compared to naturally occurring cages is focused more on confirming the success of the design and highlighting any unexpected deviations from the predicted structure. This is important as the designs are made at the atomic scale with specific amino acid side chain interactions making up the protein–protein interfaces. While other biophysical techniques such as dynamic light scattering may be able to give an approximate size of the produced cage

Figure 4



Protein cage assembly. **(a)** Symmetry engineering via protein fusion. i) In this example a protein with threefold rotational symmetry (green) is fused to a protein with twofold rotational symmetry (purple) via a linker (yellow) [40••]. The proteins assemble into a tetrahedral cage as shown schematically in the top row of the figure. The bottom row shows the structures of the individual proteins used and the crystal structure of the formed cage [39]. The central grey sphere is artificially added to highlight the presence of a cavity. From reference [39]. Reprinted with permission from AAAS. **(b)** Protein interface design showing i) Illustration of octahedral symmetry with threefold axes shown. ii) Alignment of building block proteins, identical and with threefold rotational symmetry, along the axes. Radial displacement and rotation around the axes can then be varied until a 'best fit' can be found. iii) Designability of structures is assessed. iv) A docked structure is produced. v) Close up of a protein–protein interface showing backbone in cartoon format and carbon beta atoms as sticks. vi) Amino acid side chains are included in the final stage of design to ensure an appropriate low energy interface including hydrogen bonds (dotted lines). From reference [37••]. Reprinted with permission from AAAS. **(c)** Examples of complex designed protein cages (i–iv). Each cage is made from multiple copies of two different proteins, shown in different colours. For each cage, the top row shows designs and the bottom row shows determined X-ray crystal structures. Two views for each structure are shown from left to right along the twofold and threefold tetrahedral symmetry axes respectively. RMSD between designed and solved structures are shown. Boxes on the right show the designs (white) overlaid onto crystal structures (other colours) for protein–protein interfaces. Adapted by permission from Macmillan Publishers Ltd. [Nature] [36], copyright 2014 **(d)** TRAP protein (PDB 1qaw [54]) shown i) looking along the 11-fold rotation axes and ii) orthogonal to it. The structure is shown in ribbon format, with each monomer coloured differently. The yellow spheres denote the locations of residue 35. When this residue is mutated to cysteine and the protein is mixed with gold nanoparticles, the rings form a protein cage as shown in the transmission electron micrograph iii), scale bar = 50 nm.

they cannot confirm in detail that the produced structure matches the design.

Typically, artificial cage proteins are made using one of two methods, one is a 'symmetry engineering' approach that takes advantage of the fact that the number of protein structures now known is huge (as of August 2016 the pdb database (www.rcsb.org [44]) consists of over 120 000 entries). Utilizing this structural repository, protein monomers which are known to assemble with particular symmetries can be fused together *in silico* to provide hybrid molecules. Satisfying the symmetries of the individual fused monomers results in formation of geometrically regular cage structures (Figure 4a). For example, a protein trimer (providing threefold rotational symmetry) and a second protein, a dimer (providing twofold rotational symmetry) can be fused together to produce a tetrahedral cage [40••] with four trimers forming the four vertices of a tetrahedron and the dimers forming the six edges (Figure 4a).

Another approach is to carefully design the interfaces between proteins such that new interactions (*e.g.*, hydrophobic packing, hydrogen bonds) are introduced to encourage self-assembly in desired conformations (Figure 4b). This is a computationally intense approach and software such as Rosetta has proved useful in providing advanced modeling of the crucial protein–protein interfaces between the protein subunits. This approach can result in quite stable assemblies, with the 60-subunit icosahedral cage structure recently reported by Hsia *et al*. remaining stable up to ~80°C [42]. Protein cages made from copies of two different proteins have even proved possible to design (Figure 4c). Most recently the interface design approach has led to spectacular success with Bale *et al*. reporting the production and structure of several different protein cages, each made from two different proteins and spanning a range of sizes and masses (up to almost 3 MDa) [45••]. In many of these examples cryo-EM was used to confirm the structure of the produced cages.

Still other novel approaches to protein cage construction remain to be exploited: One of us (Heddle) has developed an alternative technique using gold nanoparticles to promote protein cage formation from a homo-11mer ring protein known as TRAP (trp RNA-binding attenuator protein [46,47], Figure 4d). The resulting cage [48,49•] has unusual properties including disassembly in relatively mild reducing conditions and it may have unusual structure features which will likely require cryo-EM to fully resolve.

Design and production of artificial protein cages is still in its infancy but its potential power is now beginning to be understood and a large number of applications are envisaged [43•] including, ultimately, as modular components of more complex protein machines [50] and cryo-EM analysis will remain a key technique as the field moves forward.

Conclusions

This review focuses on the design, structure-determination and (medical) applications of protein cages, both natural (virus capsids) and artificial. The ability to produce a hollow, nanoscale 'ball' is extremely powerful; the resulting structure having internal and external faces that can be addressed independently. They demonstrate the desirable feature of having an enclosed space separated from the surrounding environment that can be used to transport/react cargoes. Building such protein cages requires structural data and the atomic models of VLPs or viruses themselves provide useful information for vaccine design as well as for understanding virus infection mechanisms. Sometimes structural information is important for understanding the difference in effectiveness of treatments between different genogroups. In such cases, obtaining atomic resolution structures quickly even in virus genogroups that are hard to propagate, is important. Cryo-EM techniques with direct-electron detectors and sophisticated image processing software are powerful tools in such cases. Cryo-EM, along with X-ray crystallography, also provides high-resolution structural data of building block proteins used to construct artificial cages. *In silico* design processes use this structural data as a start point for software-based cage design. The same structural techniques are then used to validate the produced structures. Overall, the interplay between advancing structure-determination techniques, deepening knowledge of natural protein cage structures and growing capabilities in artificial cage design are leading to the production of increasingly useful and sophisticated protein cages for medical use and beyond.

Acknowledgements

K.I. acknowledges the Platform Project for Supporting in Drug Discovery and Life Science Research (Platform for Drug Discovery, Informatics, and Structural Life Science) from Japan Agency for Medical Research and development (AMED). J.G.H. and S.C. were funded by the Malopolska Centre of Biotechnology and J.G.H. was funded by the National Science Centre (NCN, Poland) grant no. 2016/20/W/NZ1/00095 (Symfonia-4).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Iwasaki K, Omura T: **Electron tomography of the supramolecular structure of virus-infected cells**. *Curr Opin Struct Biol* 2010, **20**:632-639.

2. Liu J, Dai S, Wang M, Hu Z, Wang H, Deng F: **Virus like particle-based vaccines against emerging infectious disease viruses**. *Virol Sin* 2016, **31**:279-287.

3. Chroboczek J, Szurgot I, Szolajska E: **Virus-like particles as vaccine**. *Acta Biochim Pol* 2014, **61**:531-539.

4. Naldini L: **Gene therapy returns to centre stage**. *Nature* 2015, **526**:351-360.

5. Yildiz I, Shukla S, Steinmetz NF: **Applications of viral nanoparticles in medicine**. *Curr Opin Biotechnol* 2011, **22**:901-908.

6. Yu X, Jin L, Zhou ZH: **3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy**. *Nature* 2008, **453**:415-419.

7. Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W: **Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy**. *Nature* 2008, **451**:1130-1134.

8. Zhang X, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, Grigorieff N: **Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction**. *Proc Natl Acad Sci U S A* 2008, **105**:1867-1872.

9. Grigorieff N, Harrison SC: **Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy**. *Curr Opin Struct Biol* 2011, **21**:265-273.

10. Liu Y, Hill MG, Klose T, Chen Z, Watters K, Bochkov YA, Jiang W, Palmenberg AC, Rossmann MG: **Atomic structure of a rhinovirus C, a virus species linked to severe childhood asthma**. *Proc Natl Acad Sci U S A* 2016, **113**:8997-9002.

11. Earl LA, Subramaniam S: **Cryo-EM of viruses and vaccine design**. *Proc Natl Acad Sci U S A* 2016, **113**:8903-8905.

12. Sirohi D, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, Kuhn RJ: **The 3.8 Å resolution cryo-EM structure of Zika virus**. *Science* 2016, **352**:467-470.

13. Grunewald K, Desai P, Winkler DC, Heymann JB, Belnap DM, Baumeister W, Steven AC: **Three-dimensional structure of herpes simplex virus from cryo-electron tomography**. *Science* 2003, **302**:1396-1398.

14. Zhou ZH, Dougherty M, Jakana J, He J, Rixon FJ, Chiu W: **Seeing the herpesvirus capsid at 8.5 Å**. *Science* 2000, **288**:877-880.

15. •• Schur FK, Obr M, Hagen WJ, Wan W, Jakobi AJ, Kirkpatrick JM, Sachse C, Krausslich HG, Briggs JA: **An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation**. *Science* 2016, **353**:506-508.
This work reveals the 3.9 Å structure of the capsid and spacer peptide 1 region of the viral Gag polyprotein within the assembled immature HIV-1. This is the highest resolution structure obtained by subtomogram averaging.

16. McMullan G, Faruqi AR, Henderson R: **Direct electron detectors**. *Methods Enzymol* 2016, **579**:1-17.

17. Brilot AF, Chen JZ, Cheng A, Pan J, Harrison SC, Potter CS, Carragher B, Henderson R, Grigorieff N: **Beam-induced motion of vitrified specimen on holey carbon film**. *J Struct Biol* 2012, **177**:630-637.

18. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, Cheng Y: **Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM**. *Nat Methods* 2013, **10**:584-590.

19. Perilla JR, Gronenborn AM: **Molecular architecture of the retroviral capsid**. *Trends Biochem Sci* 2016, **41**:410-420.

20. Nguyen TT, Bruinsma RF, Gelbart WM: **Continuum theory of retroviral capsids**. *Phys Rev Lett* 2006, **96**:078102.

21. •• Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C *et al.*: **Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics**. *Nature* 2013, **497**:643-646.
The large, asymmetrical capsid structure of HIV-1 was revealed by a hybrid approach employing molecular dynamics, X-ray crystallography and cryo-electron tomography.

22. •• Wen AM, Steinmetz NF: **Design of virus-based nanomaterials for medicine, biotechnology, and energy**. *Chem Soc Rev* 2016, **45**:4074-4126.
A comprehensive and exhaustive review of literature on VLPs that are currently being used for medical and biotechnology applications.

23. Deshayes S, Gref R: **Synthetic and bioinspired cage nanoparticles for drug delivery**. *Nanomedicine* 2014, **9**:1545-1564.

24. Uchida M, Klem MT, Allen M, Suci P, Flenniken M, Gillitzer E, Varpness Z, Liepold LO, Young M, Douglas T: **Biological containers: protein cages as multifunctional nanoplatforms**. *Adv Mater* 2007, **19**:1025-1042.

25. Miller RA, Presley AD, Francis MB: **Self-assembling light-harvesting systems from synthetically modified tobacco mosaic virus coat proteins**. *J Am Chem Soc* 2007, **129**:3104-3109.

26. Carrico ZM, Romanini DW, Mehl RA, Francis MB: **Oxidative coupling of peptides to a virus capsid containing unnatural amino acids**. *Chem Commun* 2008:1205-1207.

27. Mullard A: **Regulators approve the first cancer-killing virus**. *Nat Rev Drug Discov* 2015, **14**:811.

28. Allen M, Bulte JW, Liepold L, Basu G, Zywicke HA, Frank JA, Young M, Douglas T: **Paramagnetic viral nanoparticles as potential high-relaxivity magnetic resonance contrast agents**. *Magn Reson Med* 2005, **54**:807-812.

29. Van Rompay KK, Hunter Z, Jayashankar K, Peabody J, Montefiori D, LaBranche CC, Keele BF, Jensen K, Abel K, Chackerian B: **A vaccine against CCR5 protects a subset of macaques upon intravaginal challenge with simian immunodeficiency virus SIVmac251**. *J Virol* 2014, **88**:2011-2024.

30. Wolf M, Garcea RL, Grigorieff N, Harrison SC: **Subunit interactions in bovine papillomavirus**. *Proc Natl Acad Sci U S A* 2010, **107**:6298-6303.

31. Trus BL, Roden RB, Greenstone HL, Vrhel M, Schiller JT, Booy FP: **Novel structural features of bovine papillomavirus capsid revealed by a three-dimensional reconstruction to 9 Å resolution**. *Nat Struct Biol* 1997, **4**:413-420.

32. Gong M, Zhu H, Zhou J, Yang C, Feng J, Huang X, Ji G, Xu H, Zhu P: **Cryo-electron microscopy study of insect cell-expressed enterovirus 71 and coxsackievirus a16 virus-like particles provides a structural basis for vaccine development**. *J Virol* 2014, **88**:6444-6452.

33. Long F, Fong RH, Austin SK, Chen Z, Klose T, Fokine A, Liu Y, Porta J, Sapparapu G, Akahata W *et al.*: **Cryo-EM structures elucidate neutralizing mechanisms of anti-chikungunya human monoclonal antibodies with therapeutic activity**. *Proc Natl Acad Sci U S A* 2015, **112**:13898-13903.

34. Zhao Q, Li S, Yu H, Xia N, Modis Y: **Virus-like particle-based human vaccines: quality assessment based on structural and functional properties**. *Trends Biotechnol* 2013, **31**:654-663.

35. Lamarre B, Ryadnov MG: **Self-assembling viral mimetics: one long journey with short steps**. *Macromol Biosci* 2011, **11**:503-513.

36. King NP, Bale JB, Sheffler W, McNamara DE, Gonen S, Gonen T, Yeates TO, Baker D: **Accurate design of co-assembling multi-component protein nanomaterials**. *Nature* 2014, **510**:103-108.

37. •• King NP, Sheffler W, Sawaya MR, Vollmar BS, Sumida JP, André I, Gonen T, Yeates TO, Baker D: **Computational design of self-assembling protein nanomaterials with atomic level accuracy**. *Science* 2012, **336**:1171-1174.
An overview describing how designing and modeling of protein–protein interfaces can be used to design and build protein cage assemblies.

38. Lai Y-T, Reading E, Hura GL, Tsai K-L, Laganowsky A, Asturias FJ, Tainer JA, Robinson CV, Yeates TO: **Structure of a designed protein cage that self-assembles into a highly porous cube**. *Nat Chem* 2014, **6**:1065-1071.

39. Lai YT, Cascio D, Yeates TO: **Structure of a 16-nm cage designed by using protein oligomers**. *Science* 2012, **336**:1129.

40. •• Padilla JE, Colovos C, Yeates TO: **Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments**. *Proc Natl Acad Sci U S A* 2001, **98**:2217-2221.

Gives a general strategy for designing fusion proteins from individual proteins of defined rotational symmetries such that they subsequently self assemble into protein cages.

41. Yeates TO, Padilla JE: **Designing supramolecular protein assemblies**. *Curr Opin Struct Biol* 2002, **12**:464-470.

42. Hsia Y, Bale JB, Gonen S, Shi D, Sheffler W, Fong KK, Nattermann U, Xu C, Huang PS, Ravichandran R *et al.*: **Design of a hyperstable 60-subunit protein icosahedron**. *Nature* 2016, **535**:136-139.

43. • Lai YT, King NP, Yeates TO: **Principles for designing ordered protein assemblies**. *Trends Cell Biol* 2012, **22**:653-661.
An excellent review of the various methods for designing and constructing protein cages and other protein assemblies.

44. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The protein data bank**. *Nucleic Acids Res* 2000, **28**:235-242.

45. •• Bale JB, Gonen S, Liu Y, Sheffler W, Ellis D, Thomas C, Cascio D, Yeates TO, Gonen T, King NP *et al.*: **Accurate design of megadalton-scale two-component icosahedral protein complexes**. *Science* 2016, **353**:389-394.
A tour de force demonstrating the successful design, production and structure determination of a number of different protein cages, up to ~3 MDa, each made from two different proteins.

46. Gollnick P, Babitzke P, Antson A, Yanofsky C: **Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis***. *Annu Rev Genet* 2005, **39**:47-68.

47. Yanofsky C: **RNA-based regulation of genes of tryptophan synthesis and degradation, in bacteria**. *RNA* 2007, **13**:1141-1154.

48. Imamura M, Uchihashi T, Ando T, Leifert A, Simon U, Malay AD, Heddle JG: **Probing structural dynamics of an artificial protein cage using high-speed atomic force microscopy**. *Nano Lett* 2015, **15**:1331-1335.

49. • Malay AD, Heddle JG, Tomita S, Iwasaki K, Miyazaki N, Sumitomo K, Yanagi H, Yamashita I, Uraoka Y: **Gold nanoparticle-induced formation of artificial protein capsids**. *Nano Lett* 2012, **12**:2056-2059.
This article describes how a cage protein assembly can be made from an 11-membered ring protein with the assembly process being switched on by the addition of gold nanoparticles.

50. Glover DJ, Clark DS: **Protein calligraphy: a new concept begins to take shape**. *ACS Cent Sci* 2016, **2**:438-444.

51. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE: **UCSF Chimera—a visualization system for exploratory research and analysis**. *J Comput Chem* 2004, **25**:1605-1612.

52. Kinjo AR, Suzuki H, Yamashita R, Ikegawa Y, Kudou T, Igarashi R, Kengaku Y, Cho H, Standley DM, Nakagawa A *et al.*: **Protein Data Bank Japan (PDBj): maintaining a structural data archive and resource description framework format**. *Nucleic Acids Res* 2012, **40**:D453-460.

53. Huynh NT, Hesketh EL, Saxena P, Meshcheriakova Y, Ku Y-C, Hoang LT, Johnson JE, Ranson NA, Lomonossoff GP, Reddy VS: **Crystal structure and proteomics analysis of empty virus-like particles of Cowpea mosaic virus**. *Structure* 2016, **24**:567-575.

54. Chen X, Antson AA, Yang M, Li P, Baumann C, Dodson EJ, Dodson GG, Gollnick P: **Regulatory features of the trp operon and the crystal structure of the trp RNA-binding attenuation protein from *Bacillus stearothermophilus***. *J Mol Biol* 1999, **289**:1003-1016.