# Nanoscale Biological Icosahedral Structures

by
Mark A. Martin, Ph.D.
August 2006

## 1. Introduction

In nature, identical, spontaneously assembling protein subunits tend to form symmetrical, minimum-energy shapes. Certain shapes are well-suited for microbiological tasks and persist within organisms. One shape that occurs in several contexts is the icosahedron. A common feature among the icosahedral structures found in microbiology is that they control permeability. Some of these structures, in viruses, for example, may be nearly impermeable, protecting their contents from damaging outside influences and allowing small entities with limited genetic resources to construct viable outer casings. In other cases, icosahedral forms screen molecules, allowing only molecules of a certain size access to vital cellular resources and mechanisms. Other icosahedral structures do not limit access but form a scaffolding to support and guide important cellular functions. As I will describe below, structures with this shape seem to be relatively common and some of them play pivotal roles in processes that are critical to nearly all organisms.

## 2. Viruses

My first example is the quintessential and prototypical natural icosahedral structure – the viral capsid. The genetic material of a virus is enclosed in a protein shell called a "capsid". Many viral capsids are composed of identical protein subunits and are icosahedral. The smallest functionally equivalent component of a capsid is called a "structure unit". The structure units usually cluster and form larger morphological units called "capsomers". The capsids of some viruses are further enclosed in an "envelope" of material that may contain components taken from the host cells that the virus infects. Viruses that have outer envelopes are called "enveloped" whereas those in which the capsid forms the outermost container are said to be "naked". Additional molecules or structures may be embedded in the envelope or capsid and may protrude from the central virus particle.

Many families of viruses have icosahedral capsids. Some of these, listed roughly in order of increasing size, are the Parvoviridae, Picornaviridae, Caliciviridae, Hepadnaviridae, Flaviviridae, Papovaviridae, Retroviridae, Birnaviridae, Togaviridae, Reoviridae, Adenoviridae, Herpesviridae, and Iridoviridae. Additional families that also form icosahedra are the Comoviridae, Microviridae, Tymoviridae, Tombusviridae, Togaviridae, and many of the Bacteriophages. Viruses with icosahedral capsids range approximately from 20-300 nm in diameter and have genomes that range approximately from 5-350 kb in length.

Several constraints lead to the icosahedral shape in viral capsids. Viral genomes are small and only provide a very small amount of space for the coding of capsid proteins. As a result, capsids are generally constructed from a small number of identical units, which assemble themselves into symmetrical structures.

Triangulated icosahedra are minimum free energy structures and, hence, are energetically natural and favorable. Furthermore, capsids must protect the genetic material that they contain and therefore must be relatively impermeable. Icosahedra fit this criterion whereas other symmetric structures such as tetrahedra and octahedra that may form in the same ways tend to be too permeable and result in viruses that are not viable.

Clefts (or "canyons") on the surfaces of viruses contain sites that attach to receptors on the surfaces of host cells to initiate infection. Antibodies to a virus also bind to these sites in order to prevent infection. The icosahedral shape typically provides deep clefts, which helps hide the sensitive binding sites from the host immune system.

The individual proteins that comprise the capsid tend to be small. This has several advantages. First, small proteins require less space in the genome. Smaller proteins and smaller corresponding genes also reduce the chances of error during protein synthesis. Furthermore, there is less waste if a smaller protein is synthesized or folded incorrectly. In addition, the larger number of intermolecular bonds required to hold a larger number of smaller proteins together increases the stability and flexibility of the capsid while decreasing its permeability. Finally, as the number of subunits increases, the size of the capsid increases, within limits, which allows the capsid to accommodate a larger, more complex and more adaptable genome.

These constraints are more important for viruses with smaller genomes. Larger viruses, such as members of the Poxviridae family, tend to have capsids that more closely resemble cell membranes and are less rigid and composed of a greater variety of molecular components. The abandonment of the icosahedral shape in larger organisms indicates that icosahedral forms are likely to be preferred as outer casings only under the severe constraints outlined above.

(See VIPERdb for images of viral capsids, Lodish *et al* Section 6.3, The Big Picture Book of Viruses, Virus Structure, and Virus Taxonomy.)

## 3. Structural Components

The next two examples are only erstwhile examples of biological icosahedral structures. The chemical constituents involved provide form within the cell on a relatively large scale. The first example provides a mechanism whereby the cell incorporates outside material on a large scale. The second example is a way of subdividing space within the cell. Both of these mechanisms result in compartmentalization, which is one of the most fundamental requirements of living organisms.

## 3.1 Clathrin

Clathrin is a trimeric protein with a triskelion shape. It spontaneously assembles into pentagonal and hexagonal shapes that combine to form spherical polyhedra. It generally forms vesicles that engulf extracellular material in the process of endocytosis at the cell surface. The polyhedra that clathrin forms are not typically icosahedral. But some structures constructed using clathrin have sometimes been identified as icosahedral. These are probably cases of misidentification.

(See Kentsis & Borden and Pauloin *et al*.)

## 3.2 Inclusion Bodies

Inclusion bodies are compartments that form within cells, usually during certain growth phases or under particular environmental conditions. They are often polyhedral and are usually found in prokaryotic organisms but may also be found in some eukaryotes. Their purpose is generally to store, transport, or sequester material within cells or to sequester certain reactions within the cytoplasm. This isolates the contents from the rest of the cytoplasm and may also regulate any encapsulated reactions. The walls of inclusion bodies may be composed of proteins or lipid bilayers. Sometimes these bodies may actually be viruses or virus capsids that are infectious or symbiotic. Groups of inclusion bodies are sometimes enclosed in membranous envelopes that often contain enzymes related to the processing of the contents of the bodies. Inclusion bodies may be icosahedral but also assume a wide variety of other forms. They are generally on the order of 50-100 nm in diameter.

There are many different types of inclusion bodies. Some interesting examples are carboxysomes, gas vesicles, magnetosomes, and symbiotic virus particles. Carboxysomes are centers for carbon fixation. They are usually filled with the enzyme RuBisCO, which catalyzes the key reaction of the Calvin-Benson-Bassham cycle. The Calvin-Benson-Bassham cycle comprises the dark reactions of photosynthesis. Gas vesicles are cylindrical compartments containing gas that cyanobacteria use to control their buoyancy and thereby control their positions in the water column. Magnetosomes are found in magnetotactic bacteria and magnetotactic eukaryotic algae and are filled with the magnetic materials that these organisms use to navigate toward favorable environments. The are usually arranged in chains, which orients their magnetic dipole moments parallel to each other. Their walls are lipid bilayers formed from the inner cell membrane and they obtain their polyhedral appearances from the shapes of the ferrite crystals that they contain. Symbiotic virus particles are often either intact bacteriophages that are used to attack other bacterial species or defective bacteriophages that contain compounds that can be used for similar purposes.

(See Lengeler *et al* pp. 39-41, Cannon *et al*, Michal pp. 199-200, Komieli *et al*.)

## 4. Enzyme Complexes

Enzymes greatly accelerate biochemical reactions that would otherwise be exceedingly slow. There are very few chemical processes that occur quickly enough to meet physiological needs within organisms without the intervention of enzymes. Three enzyme complexes are known to have icosahedral forms *in vivo* – the riboflavin synthase/lumazine synthase complex, pyruvate dehydrogenase, and tricorn protease. Two of these play pivotal roles in metabolic pathways that are critical for most organisms. The low number of known examples and the importance of two of them is probably a reflection of the scrutiny that very important biochemical pathways have received rather than an indication the scarcity of icosahedral enzyme complexes in nature or of the importance of an icosahedral shape for enzymatic functions. As a set of authors who studied tricorn protease pointed out, such large quaternary protein structures are delicate and so large that they often either disintegrate or are inadvertently discarded during analysis. Consequently, other examples probably exist but have yet to be discovered. In the examples that are known, the icosahedral shape seems to have two primary roles – it prevents inappropriate substrates from reaching the active sites and/or provides a structure for channeling substrates and thereby enhances enzyme activity. These two roles sometimes have opposite effects and may, at times, interfere with each other.

### 4.1 Riboflavin Synthase/Lumazine Synthase

Riboflavin (vitamin $B_2$) is produced in plants, fungi, and microorganisms. All organisms require it, since it is the basis of coenzymes that participate in over 100 redox reactions, but animals and some microorganisms cannot produce it. The enzymes lumazine synthase and riboflavin synthase catalyze the final steps in riboflavin production. In some organisms, these enzymes form a complex consisting of an icosahedral capsid constructed from 60 lumazine synthase subunits that encloses a trimeric module of riboflavin synthase. In other organisms, the capsid forms but does not contain riboflavin synthase. In both of these cases, free trimers of riboflavin synthase reside in the cytoplasm and are responsible for most of the riboflavin synthase activity. The active sites of the lumazine synthase subunits lie near the inside of the capsid. The narrow gaps in the capsid appear to restrict the passage of substrates into the capsid and the passage of products out of the capsid. In some organisms, lumazine synthase forms a pentameric structure instead. The active site for the enzymes has the same topology and is located in the same general position in both forms. The icosahedral form seems to enhance activity via substrate channeling, particularly under low substrate conditions, but the functional differences between the two forms are not well understood.

(See Michal pp. 109-10, Bacher *et al*, Mörtl *et al*, and Fornasari *et al*.)

## 4.2 Pyruvate Dehydrogenase

The pyruvate dehydrogenase enzyme complex (PDH) plays a critical role in carbon metabolism. It connects the primary anaerobic metabolic process of glycolysis to the central process in aerobic metabolism, the citrate cycle (otherwise known as the Krebs or TCA cycle), by catalyzing the reactions required to convert pyruvate to acetyl-CoA. It also governs a large number of important anabolic processes and is present in virtually all organisms.

PDH consists of the three enzymes pyruvate dehydrogenase (E1), dihydrolipoyl acetyltransferase (E2), and dihydrolipoyl dehydrogenase (E3). The complex has an interior icosahedron formed from 60 E2 tetramers that is encapsulated in an exterior icosahedron formed from E1 subunits and a relatively small number of E3 subunits. There is a space between the two icosahedra that facilitates the positioning of metabolic intermediates for the next reaction and increases the activity of the enzyme. Flexible linker molecules span the space and connect the inner and outer icosahedra. The entire assembly is approximately 475 Å in diameter. The reaction begins when substrate reacts at the active site of an E1 subunit. The intermediate product then swings down to the reaction site on an E2 subunit for further processing. Once the reaction is complete, an E3 subunit resets the mechanism to prepare for the next substrate molecule.

The structure of PDH is flexible and the enzyme is functional over a range of structures as long as the subunits are chemically coupled. Different spatial configurations of the enzyme complex lead to different values for the activity of the enzyme. The icosahedral structure greatly increases the activity of the enzyme since it positions adjacent enzymes in the chemical pathway together and concentrates the available reaction sites. Other structures typically exhibit significantly lower activities.

(See Michal pp. 27-44; Milne *et al*; and Kentsis & Borden.)

## 4.3 Tricorn Protease

Tricorn protease is an enzyme that was discovered in the archaeon Thermoplasma acidophilum and later found in certain bacteria. In these organisms, it is part of the system that breaks down defective or unneeded proteins so that their components can be recycled. The proteosome or other proteases perform the initial stages of demolition, breaking proteins down into chains 7-12 amino acids long. Tricorn protease breaks these chains down further, into peptides consisting of 2-4 amino acids. Afterwards, aminopeptidases convert these small peptides into free amino acids, which are then used to synthesize new proteins.

Tricorn protease forms an icosahedral capsid within the cell. The capsid is comprised of 20 copies of the tricorn hexameric toroid and is approximately 55 nm in diameter. The active site for the enzyme lies on the inside of the icosahedron and the capsid contains small openings that are at most 2.5 nm across. The openings appear to only admit small unfolded polypeptides into the inner reaction chamber. The enzyme is equally active when arranged in an icosahedral capsid or in individual hexamers. So the icosahedral form does not enhance enzyme activity through substrate channeling.

(See Brandstetter *et al*, Walz *et al*., and Tamura *et al*.)

## 5. Photosynthetic Complexes

For the vast majority of food webs on Earth, photosynthesis is the most important physiological process. It is the transducer that provides power to most food webs using the most powerful and plentiful source in the solar system – the sun. As it turns out, pieces of the biochemical machinery that carry out photosynthesis assume icosahedral forms under certain conditions. These conditions may or may not occur within living organisms. But the propensity of these complexes to form icosahedra is another indicator of how widespread this energetically-favorable form might be. Below, I discuss the two examples where this is currently known to occur.

### 5.1 Bacteriorhodopsin-Lipid Complexes

Bacteriorhodopsin is a proteinaceous pigment that archaea in the class halobacteria use as a proton pump in the process of photosynthesis. The photosynthetic process in these organisms is unrelated to the more familiar process that occurs in higher plants. Bacteriorhodopsin is a variant of the pigment rhodopsin that is central in the chemical processes that provide vertebrates with low-light vision via the rod cells of their retinas. Within halobacteria, it forms a purple two-dimensional crystal called "purple membrane", which is responsible for the reddish purple color of the organisms under certain conditions. It is a trimeric protein that forms icosahedral protein-lipid vesicles when incubated at high temperature (30-35°C) with a small amount of detergent and a high concentration of a precipitant. Each vesicle is composed of 420 bacteriorhodopsin trimers and is approximately 460 Å in diameter. The vesicles arrange themselves in hexagonal lattices, resulting in a three-dimensional hexagonal crystal composed of nested layers of hexagonal lattices. This crystal belongs to space group $P6_322$. The vesicles may also form octahedral crystals when allowed to crystallize at low temperatures. In the octahedral crystals, the icosahedral vesicles are arranged in a face-centered cubic structure belonging to space group F23. A slightly different process produces hexagonal crystals composed of layers in which the vesicles have been flattened out into planar honeycomb lattices. These hexagonal crystals are birefringent and are members

of the space group P622. Although the icosahedral vesicles do not form under physiological conditions, this provides another example of a protein with the propensity to form icosahedra under relatively moderate conditions.

(See Kouyama *et al*; Takeda *et al*; Hino *et al*; Michal p. 199 and p. 219; and Hall & Rao.)

5.2 Light-Harvesting Chlorophyll a/b Protein Complexes (LHC-II)

Light harvesting chlorophyll-protein complexes are key components in the photosynthetic mechanisms of plants. For most photosynthetic organisms, photosynthesis consists of three different overall reactions, two of which occur in response to light and a third that is independent of illumination. The two light reactions are photosystem I and photosystem II. Photosystem I is associated with a reduction reaction and photosystem 2 is associated with an oxidation reaction. The net result of the light reactions is the oxidation of water and the production of the energetic compounds ATP and $NADPH_2$. The energetic compounds produced are used in the dark reactions to convert $CO_2$ to carbohydrates.

The light harvesting chlorophyll-protein complex associated with photosystem II, LHC-II, is a protein-pigment complex that contains both chlorophyll a and chlorophyll b. The primary protein component of LHC-II has a stable trimeric structure under physiological conditions, similar to that of bacteriorhodopsin. Also like bacteriorhodopsin, LHC-II forms hexagonal and octahedral crystals. The hexagonal crystal seems to be made of flat layers like the P622 crystal of bacteriorhodopsin and there are two types of octahedral crystal that are regular arrangements of icosahedral vesicles similar to the octahedral crystal formed by bacteriorhodopsin. The octahedral crystals form at low temperatures (approximately 10°C), which are considerably lower than those required for the crystallization of bacteriorhodopsin. Their formation also requires the presence of potassium chloride and lipids naturally occurring near LHC-II in the cell. These conditions are close to those commonly found *in vivo*. Vesicles form when a particular lipid (phosphatidylglycerol) fills the gaps between proteins and allows the creation of a closed proteoliposome. Each icosahedron consists of 20 LHC-II protein trimers and is approximately 250 Å in diameter. In one of the octahedral crystals, the vesicles are arranged in a structure belonging to space group F23 that is merohedrally twinned. In the other, the vesicles are arranged in a structure belonging to space group $P2_13$. Icosahedral vesicles have not been observed under physiological conditions but it is hypothesized that they might form during the development of thylakoid membranes or anytime LHC-II is located in low-lipid regions of the cell.

(See Hino *et al*; Michal pp. 196-9; Hall & Rao; and Falkowski p.2 and p. 175.)

## 6. Conclusions

As described above, an icosahedral shape is important in several physiological contexts. It forms naturally and relatively easily from many biochemical compounds and has numerous features that have proven their usefulness through the evolutionary process of selection. Using icosahedra to divide space and control access is very efficient and seems to be preferred in many contexts where resources are severely limited.

Several processes where icosahedra play a role are very important and the crucial features of important processes tend to be conserved. Since the substrate channeling capabilities of the nested icosahedra in the pyruvate dehydrogenase complex are so pronounced, the icosahedral forms of the complex are probably quite ancient and are unlikely to deviate much in different organisms. In contrast, even though riboflavin synthesis is a critical process, the icosahedral form of lumazine synthase seems to confer less of a benefit in this case. Consequently, its icosahedral form is less conserved and lumazine synthase is not icosahedral in many organisms. As this illustrates, perpetuation of an icosahedral form does not depend solely on the importance of the process where it plays a role. Icosahedra are likely to be more universal when they participate in important processes and provide greater benefit.

The general characteristics required for a biochemical complex to assume an icosahedral shape are currently uncertain. Nonetheless, since icosahedral forms seem to provide substantial benefits at relatively low costs in many contexts, it seems likely that icosahedral forms are relatively common within organisms and that more examples will be uncovered as more systems are studied and more sensitive methods of investigating cellular function are developed.

## 7. References


- Bacher, Adelbert, Eberhardt, Sabine, Eisenreich, Wolfgang, Fischer, Markus, Herz, Stefan, Illarionov, Boris, Kis, Klaus, and Richter, Gerald, "Biosynthesis of Riboflavin", Vitamins and Hormones, v. 61, pp. 1-49, 2001.

- The Big Picture Book of Viruses, URL: http://www.virology.net/Big_Virology/BVHomePage.html

- Brandstetter, H., Jeong-Sun, K., Groll, M. & Huber, R., Crystal structure of the tricorn protease reveals a protein disassembly line. Nature (London), 414 (2001), pp 466-470.

- Cannon, G. C., Bradburne, C E., Aldrich, H. C., Baker, S. H., Heinhorst, S., & Shively, J. M. Microcompartments in Prokaryotes: Carboxysomes and Related Polyhedra. Appl Environ Microbiol. 67(12) (2001), pp 5351-61.


 

URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=93316

- Falkowski, Paul G. and Raven, John A. *Aquatic Photosynthesis*. Blackwell Science, Malden, MA. 1997.

- Fornasari, María Silvina, Laplagne, Diego A., Frankel, Nicolás, Cauerhff, Ana A., Goldbaum, Fernando A., and Echave, Julián. 2004. Sequence determinants of quaternary structure in lumazine synthase. Mol Biol Evol. 21(1):97-107. URL: http://mbe.oxfordjournals.org/cgi/content/full/21/1/97

- Hall, D. O. and Rao, K. K. *Photosynthesis, 5th Ed.* Cambridge University Press, Cambridge, U.K. 1994.

- Hino T, Kanamori E, Shen JR, Kouyama T. An icosahedral assembly of the light-harvesting chlorophyll a/b protein complex from pea chloroplast thylakoid membranes. Acta Crystallogr D Biol Crystallogr. 2004 May;60(Pt 5):803-9. Epub 2004 Apr 21.

- Kentsis, A. & Borden, K. L. B. Physical Mechanisms and Biological Significance of Supramolecular Protein Self-Assembly. Current Protein and Peptide Science, v. 5, n. 2 (2004). 125-34.

- Komieli, A., Zhuo Li and D. K. Newman "Magnetosomes Are Cell Membrane Invaginations Organized by the Actin-Like Protein MamK" Science, 311, Jan. 2006, p.242-245.

- Kouyama, T., Yamamoto, M., Kamiya, N., Iwasaki, H., Ueki, T. & Sakurai, I. Polyhedral Assembly of a Membrane Protein in its Three-dimensional Crystal. J. Mol. Biol. 236 (1994), 990-994.

- Lengeler, Joseph W., Drews, Gerhart, and Schlegel, Hans G., Eds., *Biology of the Prokaryotes*, Georg Theime Verlag, Stuttgart, Germany, 1999.

- Lodish, Harvey, Berk, Arnold, Zipursky, Lawrence S., Matsudaira, Paul, Baltimore, David, Darnell, James. Molecular Cell Biology, 4th Ed. W. H. Freeman and Company, New York, 2000. URL: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=mcb

- Michal, Gerhard, Ed., *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley & Sons, Inc., New York, 1999.

- Milne, J. L. S., Shi, D., Rosenthal, P. B., Sunshine, J. S., Domingo, G. J., Wu, X., Brooks, B. R., Perham, R. N., Henderson, R., & Subramaniam, S., Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: a multifunctional catalytic machine. The EMBO

Journal, 21 (2002), 5587-98. URL: http://www.nature.com/emboj/journal/v21/n21/full/7594791a.html

- Mörtl, Simone, Fischer, Markus, Richter, Gerald, Tack, Jens, Weinkauf, Sevil, and Bacher, Adelbert. 1996. Biosynthesis of Riboflavin: LUMAZINE SYNTHASE OF ESCHERICHIA COLI. Journal of Biological Chemistry. v 271, n 52, Issue of December 27, 1996 pp. 33201-33207. URL: http://www.jbc.org/cgi/content/full/271/52/33201

- Pauloin A, Bernier I, eJolles P. Presence of cyclic nucleotide-Ca2+ independent protein kinase in bovine brain coated vesicles. Nature. 1982 Aug 5;298(5874):574-6.

- Persson, Karina, Schneider, Gunter, Jordan, Douglas B., Viitanen, Paul V., and Sandalova, Tatyana. Crystal structure analysis of a pentameric fungal and an icosahedral plant lumazine synthase reveals the structural basis for differences in assembly. Protein Science (1999), 8:2355-2365.

- Shively, J M, Bock, E, Westphal, K, and Cannon, G C. Icosahedral inclusions (carboxysomes) of Nitrobacter agilis. J Bacteriol. 1977 November; 132(2): 673–675. URL: http://www.pubmedcentral.gov/articlerender.fcgi?tool=pubmed&pubmedid=199579

- Takeda, K., Sato, H., Hino, T., Kono, M., Fukuda, K., Sakurai, I., Okada, T. & Kouyama, T. (1998). J. Mol. Biol. 283, 463-474.

- Tamura, N., Pfeifer, G., Baumeister, W., & Tamura, T., Tricorn Protease in Bacteria: Characterization of the Enzyme from Streptomyces coelicolor. Biol. Chem. 382(3) (2001), pp 449-58.

- VIPERdb – Virus Particle Explorer, URL: http://viperdb.scripps.edu/

- Virus Structure, URL: http://www-micro.msb.le.ac.uk/3035/3035Structure.html

- Virus Taxonomy, URL: http://www-micro.msb.le.ac.uk/3035/VirusGroups.html

- Walz, J., Tamura, T., Tamura, N., Grimm, N., Baumeister, W., & Koster, A. J., Tricorn Protease Exists as an Icosahedral Supermolecule In Vivo. Mol. Cell. 1(1) (1997), pp 59-65. URL: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6WSR-4194JBT-7&_coverDate=12%2F31%2F1997&_alid=414563778&_rdoc=1&_fmt=&_o

rig=search&_qd=1&_cdi=7053&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=411d694be6b06e28de417c3b54d1f8c9