Commentary

A groovy new structure

Eva J. Neer*† and Temple F. Smith‡

*Cardiovascular Division, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA 02115; and ‡Boston University, Biomolecular Engineering Research Center, Boston, MA 02215

Recently ter Haar et al. (1) showed that the N-terminal domain of clathrin, the major protein component of clathrin-coated vesicles, is made up of the increasingly commonly recognized β propeller fold. In this issue of PNAS, ter Haar et al. (2) now show that peptides from two clathrin adapters, β-arrestin 2 and the β subunit of AP-3, bind to a groove on the β propeller surface. This appears to be a novel method of protein–protein binding for β propeller structures that adds a new facet to the ingenious ways that these structures have devised for binding protein ligands, sugars, prosthetic groups, and ions.

Clathrin-coated vesicles are one class of transport vesicles that have a cage of proteins covering the cytosolic surface. Clathrin is a protein complex that is made up of three heavy and three light chains. Together, these proteins form a three-legged structure called a triskelion. The triskelions assemble to form a network of hexagons and pentagons that make up the coated pits on the cytoplasmic surface of the membrane. The amino terminal portions of each heavy chain turn inward and form a secondary shell inside the cage-like clathrin structure. As the clathrin-coated vesicles assemble, they recruit membrane-anchored proteins to the cage-like structure through adapter proteins that interact with the N-terminal, β propeller domain. As ter Haar et al. (2) show, two adapter proteins, β-arrestin and AP-3, bind in part through interaction of short peptides with a groove between the first two propeller structural repeats called blades. It is likely that two other adapter proteins, AP-1 and AP-2, also bind to the groove between blades 1 and 2 (2). It is interesting that both peptides bind to the same site on the seven-bladed clathrin propeller because earlier work suggested that β-arrestin and AP-2 would bind to different sites (3). However, the holoprotein may bind to additional regions besides those identified for these peptides. Although the proteins can, in principle, compete for the same binding sites, this is not likely to be important in vivo because there is more clathrin than either β-arrestin or the AP proteins. They are not, therefore, likely to saturate the clathrin binding sites. Some clathrin ligands can interact with each other and, the authors suggest, might cooperate in binding to the coated vesicles.

β propellers turn out to be a common fold that has been highly exploited over the past three billion years of evolution. These proteins have a symmetric architecture made up of 4–8 structural repeats. Each of these repeats or propeller blades is made up of four antiparallel β strands radiating outward from a central axis. All of the proteins with this fold have a high percentage of large and/or aromatic hydrophobic residues that form the contact surface between the blades. The toroidal structure has a narrower end (often called the top) and a wider end (the bottom). In addition, there is a central tunnel that varies in shape and diameter. The mean diameter of the tunnel increases as the number of blades increases. Its degree of cylindrical symmetry is directly related to the degree of structural similarity among the blades.

The β propeller proteins can be divided into several families. The most diverse and currently largest single family is characterized by the so-called WD sequence repeat. The only WD repeat protein whose structure has been determined is the β subunit of heterotrimeric G proteins (5–7). However, the highly conserved sequence of the WD repeat proteins makes it very likely that they all have a propeller structure, but with different numbers of blades. The WD repeat consists of a nearly constant length region defined by a pattern of amino acid residues typically bracketed by the dipeptides GH and WD (4). This sequence pattern has a one-to-one correspondence to the inner three strands of the structural repeat. The outermost strand (d strand) is encoded in a variable length region with sequences specific to the different WD repeat functional subfamilies. The most conserved feature of the WD repeat proteins is an aspartic acid residue six positions N-terminal to the WD that is present in 86% of the WD repeats (8). In the β subunit and with the heterotrimeric G proteins, it forms the tight turn between the b and c strands of the propeller blade. The innermost strand is the a strand and the outermost is d. Clathrin is without a recognizable sequence repeat and does not contain the typical WD repeat pattern or the highly conserved aspartic acid on the upper surface of the ring. However, given the apparent universal need for at least one very tight turn either between strands a and b or b and c of all known β propeller proteins, the high frequency of small residues in the inner strands of the propeller blades and hydrophobic aromatics in the remaining strands, it is not surprising that all β propeller proteins show some weak sequence similarity.

More than 160 different WD repeat proteins have been identified so far, the overwhelming majority of which are found in eukaryotes (4). The occasional WD repeat proteins found in prokaryotes probably arose by horizontal gene transfer, but that family has not expanded there to the extent that it has in eukaryotes (for example, there are more than 50 WD repeat proteins in the yeast Saccharomyces cerevisiae). In contrast, the non-WD repeat β propellers frequently are found in prokaryotes. It is not known how many non-WD repeat propellers exist because they do not have a signature sequence that allows for their easy identification. However, more than 15 have been crystallized so far, approximately half from prokaryotes. Some proteins are made up of more than one propeller. For example, hemopexin has two four-bladed propellers per chain and dimerizes to form a symmetrical structure of four four-bladed propellers (9). The individual blades of these propellers can be superimposed on the blades of the Gβ subunit and with other non-WD repeat propellers with a fit of 0.9–2.2 Å rms deviation (10, 11), but the overall structure and degree of symmetry varies among the proteins. For example, the central tunnel can be very circular when the proteins are highly symmetrical or oval, as in clathrin and neuraminidase, when the structure is less symmetrical (2, 9). The reduced symmetry appears to arise most often from the lack of a fourth strand, which sometimes is replaced by a short helix, or from the fact

See companion article on page 1096.

†To whom reprint requests should be addressed: Cardiology Division, Brigham and Women’s Hospital, 75 Francis Street, Boston, MA 02115. E-mail: neer@calvin.bwh.harvard.edu.
that some of the c and d strands extend far above the upper surface. The latter is seen in blades 1 and 4 of the clathrin structure (Fig. 1).

In contrast to the WD repeat propellers, none of which have enzymatic activity in the propeller domain, non-WD repeat β propellers are frequently enzymes. The unifying function of the WD repeat β propellers appears to be the organized binding of many proteins, either simultaneously or sequentially. Interestingly, this also appears to be the function of some non-WD repeat propellers, such as the clathrin N-terminal domains (2) and hemopexin that binds heme and several proteins that cause release of heme (9).

The β propellers have developed a large variety of ways to use the loops and turns that protrude from their surfaces, the outermost strands and the central tunnel to bind ligands and to create catalytically active sites. In cytochrome cd 1/nitrate reductase, a heme group is coordinated at the opening of the tunnel (11). The top of the central tunnel in G6 coordinates a histidine from its partner G protein α subunit (5, 6, 10). Methylamine dehydrogenase, a seven-bladed propeller, and methanol dehydrogenase, an eight-bladed propeller, both use a quinone as a prosthetic group but in quite different ways (12–15). Both proteins have an H3L2 structure in which the heavy chain (H) is a β propeller and the light chain (L) is either an extended structure that runs across the surface of the heavy chain (methanol dehydrogenase) or is a globular protein with extensive contacts to the H subunit (methylamine dehydrogenase). In methanol dehydrogenase, the noncovalently bound quinone and essential Ca2+ ion are coordinated at the narrow end of the central tunnel. In contrast, in methylamine dehydrogenase, the active site is in a cleft between the H and L chains and the quinone is covalently bound to the L chain. Galactose oxidase is probably the most creative at forming an active site. Two β strands from the C-terminal domain enter the tunnel and form a binding site for a copper ion required for catalysis (16).

Nonenzymatic β propellers also vary in the way they bind ligands. The newly described clathrin structure reveals that peptides from two clathrin binding proteins, β-aretin and AP-3, bind into a groove formed in part by the d strand of blade 1 and the upper surface residues of blade 2. ter Haar et al. (2) propose that other β propellers also may avail themselves of this mode of binding. Two other β propeller proteins use a variation of this way of binding, albeit not for proteins but for sugars and heme. Very recently, hemopexin was shown to bind heme to a similar groove between blades 3 and 4 (17). Tachylectin 2, a protein that is part of the innate immunity system of the horse-shoe crab, is a five-bladed non-WD repeat β propeller with five identical binding sites for N-acetylgalactosamine and N-acetylgalactosamine located between the β-sheets (18).

The G6 subunit is a seven-bladed propeller that has a G7 subunit tightly bound to the bottom surface. It interacts with more than 20 proteins (19). The structure of G6y subunit bound to the G protein α subunit and to the regulatory protein, phosducin, have been solved (5, 6, 20). In each case, the proteins bind to the top surface of Gβ near the central tunnel, as well as to the side of the torus. The α subunit extends a helix along the side of blade one. Its contact is with a groove created by the protrusion of a loop from blade 2, but not with a groove between the blades. Phosducin also binds to the side of the torus at blades 6 and 7 but, again, not into a groove in the basic propeller structure. [Note that the Sigler group (6, 7, 20) uses a different convention for numbering blades from the Sprang group (5). We have used the convention of Sprang and called the first blade that contains the full WD repeat, blade 1.]

Finally, intersubunit contacts in multimeric β propellers reveal other potential modes of protein–protein interactions. Viral neuraminidase is tetramer of six-bladed propellers with all enzymatic and antigenic activity located in the propellers. The outer β strand of one neuraminidase monomer contacts its counterpart in another monomer so that there is a continuous β-sheet across the subunit interface (21).

All of the above functions require a very stable core that can be customized by changes in its surface residues. Many proteins can modify a relatively small part of their surface to change the characteristics of binding or catalysis. However, it may be that the β propeller fold creates a core structure whose stability is largely independent of its surface residues, producing a uniquely stable multisided platform capable of simultaneously accommodating many surface modifications. Clathrin appears to use the highly adaptive surfaces of the β propeller to bind adapter proteins. In clathrin, the basic propeller fold and symmetry remain even with major surface modifications on multiple sides, including those that are directly linked to the propeller or are extensions of the core secondary structure (Fig. 1). Thus, we expect to find this domain fold exploited in many more functional families, both ancient and more recently arisen.

This work was supported by National Institutes of Health Grant GM36259 to E.J.N. and National Human Genome Research Institute Grant T32 HG0041 to T.F.S.

Fig. 1. Side view of the clathrin β propeller. Strands extending above the torus in blades 1 and 4 with a β hairpin turn are shown in red. At the bottom of blade 5, two loops that extend below the torus are shown in blue. The figure was made from coordinates kindly provided by T. Kirchhausen (2) and plotted by using KASMOLE.