Journal of Structural Biology xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Protein tentacles

Stephen C. Harrison

Harvard Medical School, Boston Children's Hospital, and Howard Hughes Medical Institute, 250 Longwood Ave., Boston MA 02115, United States

ARTICLE INFO

Article history:
Received 4 May 2017
Accepted 26 May 2017
Available online xxxx

Keywords:
Protein interactions
Peptide-surface association
Regulated assembly
Virus structure

ABSTRACT

Virus structures were among the earliest illustrations of how regulated protein assembly can proceed by folding of polypeptide-chain segments into complementary sites on partner proteins. I draw on Caspar's image of protein "tentacles" and his metaphor of SV40 pentamers as five-legged, aquatic organisms ("pentopuses") to suggest a helpful vocabulary. "Tentacular interactions" among component subunits organize most subcellular molecular machines. Their selective advantages include facile regulation of both assembly and disassembly by modifying enzymes and by folding chaperones.

Biomolecular self assembly requires specificity. A large-scale structure cannot evolve without precisely defined interactions between assembling units. Caspar and Klug developed the notion of quasi-equivalence to reconcile this requirement for specificity with the observation that the capsids of many icosahedral viruses assemble from defined multiples of sixty protein subunits and hence that these subunits must have alternative contact geometries (Caspar and Klug, 1962). With only the high-resolution structure of myoglobin from which to derive empirical principles in 1962, the quasi-equivalence notion was necessarily vague. It left open two obvious questions. First, if there are alternative contact geometrics, how different can they be? Second, what determines the correct alternative at each position in a closed shell? For small capsid sizes, Caspar and Klug proposed (at the time, somewhat indefinitely) that some degree of compliance at the contact between two well-folded building blocks and adjustment of the final, closed structure to a minimum-energy configuration might be the respective answers.

The high resolution structure of tomato bushy stunt virus (TBSV) (Harrison et al., 1978) showed substantially more "sophisticated" solutions to both these puzzles than Caspar and Klug had suggested explicitly in 1962 (Fig. 1). The most striking compliance is within the subunits, which have two well-folded domains with an intervening hinge. The Ca^{2+} stabilized contacts around the local threefold axis are certainly "quasi-equivalent" – indeed, nearly equivalent in geometry as well as in chemistry – but a framework of N-terminal arms resolves, by a switching mechanism, any ambiguity in choice of an interface alternative at the other contacts. Although the alternating "direct" and "divided" contacts around a threefold are, as Caspar and Klug wrote, "deformed in slightly different ways" (with some reservations about "slightly"), the intervention of the folded arms also generates an all-or-none distinction between the "direct" contact (which is the default option, as it is also present around the fivefold and, in the absence of arms, in T = 1 "small particles" (Harrison and Jack, 1975)) and the "divided" contact (Harrison, 1980). Whether the switch violates the spirit of the original quasi-equivalence notion is the sort of distinction best left for post-prandial verbal debates: Caspar and Klug presciently recognized that non-rigidity would be a critical characteristic of even very specifically folded proteins. Switches essentially identical to those in TBSV, created by an underlying T = 1 framework, are present in nearly all T = 3 viral capsids (the RNA phage being an evident exception) (Abad-Zapatero et al., 1980; Prasad et al., 1999; Valegard et al., 1990).

The all-pentamer polyoma and SV40 structures famously violated the T = 7 prediction of hexamers and pentamers (Rayment et al., 1982). They have pentamers at all the T = 7 lattice points – both five- and six-coordinated (Fig. 2) (Liddington et al., 1991). Extended protein arms come to the rescue of specificity in this case, by docking equivalently into their target subunits. Indeed, nearly all the inter-subunit contacts are identical, as the variability is largely accommodated by alternative directions adopted by the arms as they emerge from their subunit of origin. In a minireview of the SV40 crystal structure, Don Caspar waxed both lyrical and punningly allusive (Caspar, 1992). "In the cytoplasm, or isolated in vitro, an individual SV40 pentamer will behave like an animate creature (dubbed here a 'pentopus'), erratically flexing its donor organ near the end of each tentacle and grasping with its acceptor organ near the base of each face of its five-sided head. If swarming pentopi could be seen, their chaotic movements in search of each

E-mail address: harrison@crystal.harvard.edu

http://dx.doi.org/10.1016/j.jsb.2017.05.012
1047-8477/

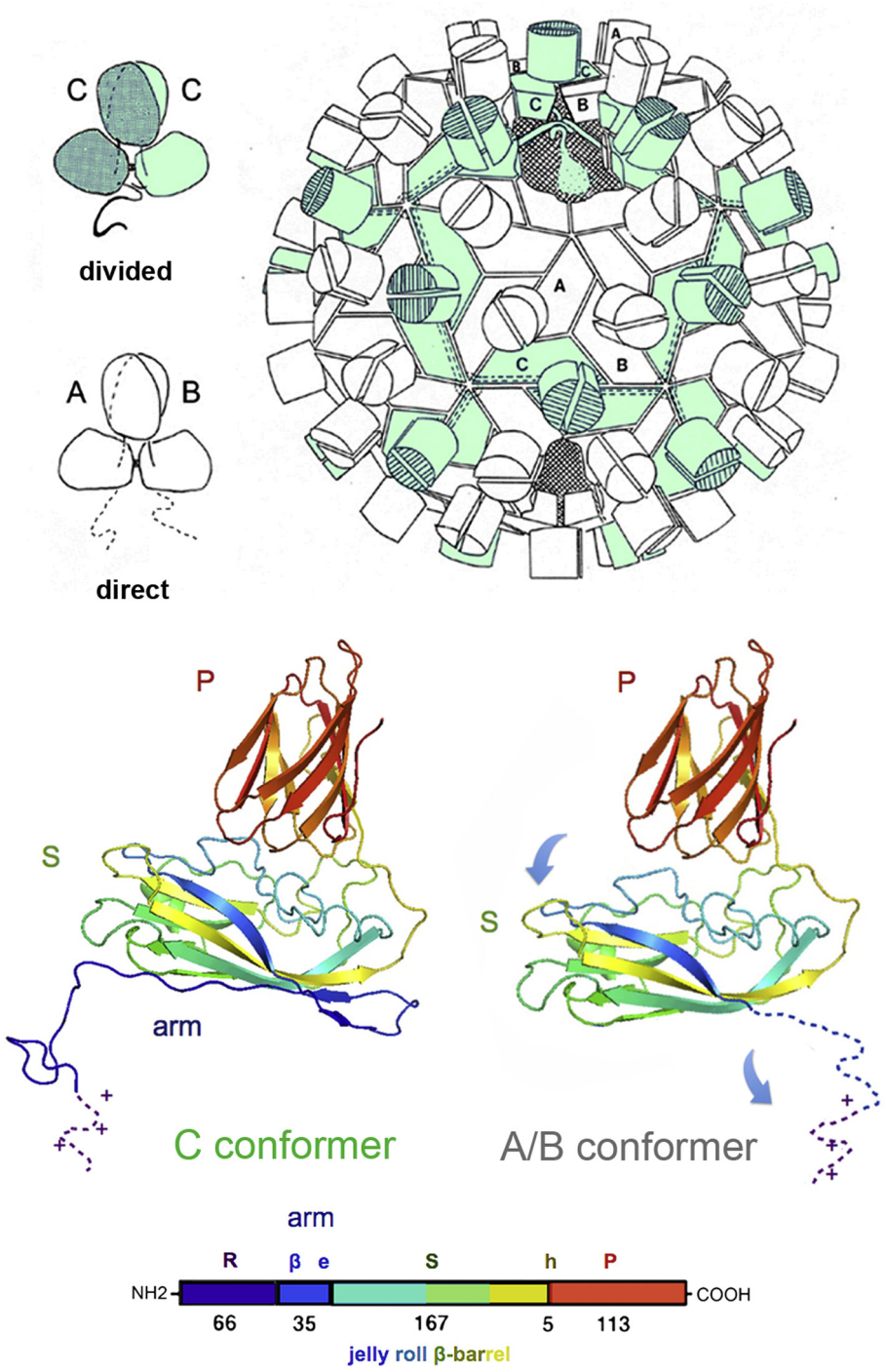

Fig. 1. TBSV. The 180-subunit (T = 3) shell assembles from dimers of a 40 kDa protein subunit (bottom). Subunits at the three quasi-equivalent positions in the capsid, designated A, B and C, cluster as 60 A:B dimers and 30 C:C dimers. The A and B conformations are nearly identical, but differ from the C conformation by the hinge angle between the two globular domains ("shell", S, and "projection", P) and unfolding of the arm (middle). The positively charged, 66-residue R segment at the N-terminus interacts with the viral RNA; segments on subunits that nucleate assembly may co-fold with RNA packaging signals. The folded arms of C:C dimers interact with arms from two other C:C dimers, creating an inner scaffold (top right). The C:C contact (and the C:B contact) along one side of the A:B:C triangle is "divided" by the scaffold of arms; the A:B contact (and the A:A contact) along the two other sides of the A:B:C triangle is "direct" (top left). The interaction between S domains in the two dimers is essentially a rotation around a fixed fulcrum, conserving many of the side-chain contacts; the P-domain contacts are invariant.

other might seem a mad pursuit. When guided to a conducive environment in the infected cell, the mutual attractions of 72 VP1 pentamers inexorably lead to their intricate frozen embrace; either to form a vacant vessel by themselves or, conjointly with the minor protein go-betweens, to envelop the irregularly compacted viral minichromosome in the precisely fashioned protein coat."

The relevance of protein tentacles extends well beyond an opportunity for colorful metaphor. Although apposition of prefolded subunit interfaces, with TMV as a precedent, may have dominated considerations behind the original Caspar-Klug descriptions, "peptide-surface association" – docking of a flexible extension of one subunit into a specific receiving site on another

Please cite this article in press as: Harrison, S.C. Protein tentacles. J. Struct. Biol. (2017), http://dx.doi.org/10.1016/j.jsb.2017.05.012