Published in final edited form as:

J Struct Biol. 2017 December ; 200(3): 244–247. doi:10.1016/j.jsb.2017.05.012.

Protein tentacles

Stephen C. Harrison

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

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.

Keywords

protein interactions; peptide-surface association; regulated assembly; virus structure

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 geometries, 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 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 pre-folded 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 (Harrison, 1996) -- is an equally fundamental principle of sub-cellular macromolecular assembly. In such cases, assembly is simply an extension of protein folding, with part of one protein attaining its selected folded structure only when associated with its partner. The same machinery that assists protein folding, by stimulating unfolding to allow the protein to "try again", can then assist protein assembly, in a similar correction mode, by mediating local disassembly. Hsp70-type folding chaperones can be large-scale disassemblases, as the mechanism of clathrin uncoating illustrates, precisely by trapping transient instabilities (Bocking et al., 2011). SV40 and polyomavirus also lock in the docking of arms with disulfide bonds, which presumably form when the assembled particle

experiences an oxidizing environment -- e.g., upon lytic release (Liddington et al., 1991; Stehle et al., 1994). During entry, protein disulfide isomerases unlock the arms (Schelhaas et al., 2007; Walczak and Tsai, 2011) Any role for Hsp70-type chaperones in subsequent uncoating -- e.g., by capture of unlocked arms that fluctuate out of their docking site -- remains to be determined.

Chaperones are not the only enzymes that can take advantage of a transient fluctuation. Most post-translational modifications require that an extended segment of polypeptide chain fit into a specificity groove -- e.g., the substrate site in a protein kinase (Zheng et al., 1993). As the buried surface area in a peptide-recognition interaction is smaller than it is in most cases of surface-surface docking, the peptide partner will "breathe" out of its pocket, allowing capture by a suitable localized modifying enzyme. Thus, a kinase, phosphatase, acetylase, deacetylase, or similar modifier can also take apart an assembly, if the modification prevents reinsertion of the peptide into its target. A recent example from our own work is phosphoregulation of kinetochore assembly (Zheng et al., 1993), but there are many precedents (e.g., histone modification).

Tentacular interactions are ubiquitous features of regulated protein assemblies. Caspar's metaphor deserves to be revived, as there are monopuses and dipuses and polypuses of all kinds. Protein tentacles, for which the inappropriate designation "intrinsically unstructured" seems to have gained currency, are unsatisfied valences awaiting a relevant complement. They are the opposite of "intrinsically unstructured"; like polypeptide chains of folded domains as they emerge from a ribosome, tentacles are "intrinsically structured" (where "intrinsic" in any biological context needs to mean "selected by evolution") but temporarily unsatisfied. As Caspar implied, their fluctuations are in no sense a mad pursuit.

Acknowledgments

I am indebted to Don Caspar, whose sometimes cryptic wisdom has always been worth deciphering. Our studies of virus structure have been possible through long-term support from NIH grant CA-13202 and from the Howard Hughes Medical Institute.

References

- Abad-Zapatero C, Abdel-Meguid SS, Johnson JE, Leslie AGW, Rayment I, Rossmann MG, Suck D, Tsukihara T. Structure of southern bean mosaic virus at 2.8 Å resolution. *Nature*. 1980; 286:33–39. [PubMed: 19711553]
- Bocking T, Aguet F, Harrison SC, Kirchhausen T. Single-molecule analysis of a molecular disassemblase reveals the mechanism of Hsc70-driven clathrin uncoating. *Nat Struct Mol Biol*. 2011; 18:295–301. [PubMed: 21278753]
- Caspar DL. Virus structure puzzle solved. *Curr Biol*. 1992; 2:169–171. [PubMed: 15335969]
- Caspar DLD, Klug A. Physical principles in the construction of regular viruses. *Cold Spr. Harb. Symp. Quant. Biol*. 1962; 27:1–24.
- Harrison SC. Protein interfaces and intersubunit bonding. The case of tomato bushy stunt virus. *Biophys J*. 1980; 32:139–153. [PubMed: 7248446]
- Harrison SC. Peptide-surface association: the case of PDZ and PTB domains. *Cell*. 1996; 86:341–343. [PubMed: 8756715]
- Harrison SC, Jack A. Structure of tomato bushy stunt virus. Three-dimensional x-ray diffraction analysis at 16 Å resolution. *J Mol Biol*. 1975; 97:173–191. [PubMed: 1177320]

- Harrison SC, Olson A, Schutt CE, Winkler FK, Bricogne G. Tomato bushy stunt virus at 2.9 Å resolution. *Nature (London)*. 1978; 276:368–373. [PubMed: 19711552]
- Liddington RC, Yan Y, Zhao HC, Sahli R, Benjamin TL, Harrison SC. Structure of simian virus 40 at 3.8 Å resolution. *Nature*. 1991; 354:278–284. [PubMed: 1659663]
- Prasad BV, Hardy ME, Dokland T, Bell J, Rossmann MG, Estes MK. X-ray crystallographic structure of the Norwalk virus capsid. *Science*. 1999; 286:287–290. [PubMed: 10514371]
- Rayment I, Baker TS, Caspar DLD, Murakami WT. Polyomavirus capsid structure at 22.5 Å resolution. *Nature*. 1982; 295:110–115. [PubMed: 6276752]
- Schelhaas M, Malmstrom J, Pelkmans L, Haugstetter J, Ellgaard L, Grunewald K, Helenius A. Simian Virus 40 depends on ER protein folding and quality control factors for entry into host cells. *Cell*. 2007; 131:516–529. [PubMed: 17981119]
- Stehle T, Yan Y, Benjamin TL, Harrison SC. Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature*. 1994; 369:160–163. [PubMed: 8177322]
- Valegard K, Liljas L, Fridborg K, Unge T. The three-dimensional structure of the bacterial virus MS2. *Nature*. 1990; 345:36–41. [PubMed: 2330049]
- Walczak CP, Tsai B. A PDI family network acts distinctly and coordinately with ERp29 to facilitate polyomavirus infection. *J Virol*. 2011; 85:2386–2396. [PubMed: 21159867]
- Zheng J, Knighton DR, ten Eyck LF, Karlsson R, Xuong N, Taylor SS, Sowadski JM. Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. *Biochemistry*. 1993; 32:2154–2161. [PubMed: 8443157]

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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 2.

SV40 and murine polyoma virus. The 72 VP1 pentamers (top right) interact with each other almost exclusively through their C-terminal "tentacles", projecting from a pentamer of jelly-roll β -barrel domains (top left). The tentacles insert into "slots" in neighboring pentamers. Disulfide bonds and divalent cation bridges then lock the tentacles in place (top right). Bottom: Don Caspar's representation: a pentopus (left) and its "frozen embrace" with neighbors (right). (From (Caspar, 1992))