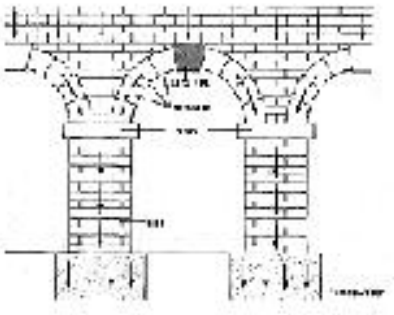
PROTEIN FOLDING, MODULARITY, SELF-ASSEMBLY, AND INTERCHANGEABLE PARTS

a. What is biological self-assembly?

A remarkable number of macromolecular structures can be assembled in a test tube starting with discrete components taken from the completed structure. This insight led to the concept of *self-assembly* ...

How are the voussoir supported?



b. Self-assembly is one of the great organizing principles in biology

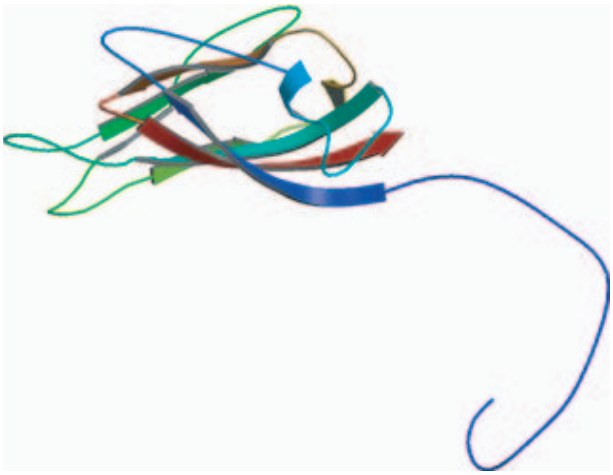
c. Templates and enzymes are rarely used in macromolecular assembly, the "big 3" notwithstanding.

Principles of Macromolecular Assembly

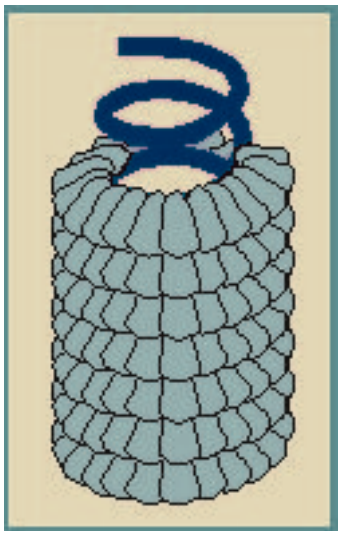
1. Large assemblies are built of subunits, frequently identical subunits, e.g. tobacco mosaic virus.

2. Specificity in macromolecular assemblies is achieved by multiple weak bonds on the surfaces of subunits with complementary topographies.

3. Many macromolecular assemblies are symmetric structures because the identical subunits make equivalent (or quasi-equivalent) contacts with their neighbors.



158 amino acids



2,130 identical subunits

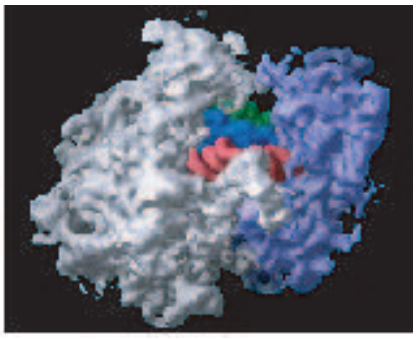
Principles of Macromolecular Assembly

4. **Multi-subunit structures assemble by defined sequential pathways along which new properties emerge as the structure and complexity grow.**

5. **Sequential pathways provide multiple opportunities for regulation.**

50S subunit: 23S RNA, 5S RNA, 31 Proteins

30S subunit: 16S RNA, 21 Proteins

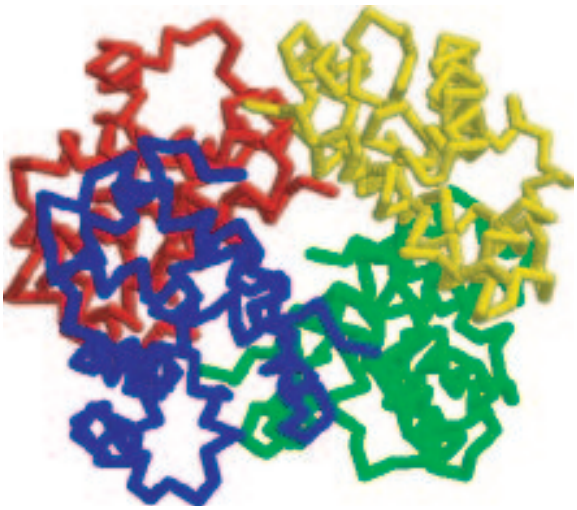


d. Top-down Hierarchy

Larger Assemblies ← ... ← Small Assemblies ← Multimers ← Monomers ← Domains

Complex structures are assembled spontaneously from smaller structures, which are, in turn, are assembled from yet smaller structures. What is the Aristotelean unmoved-mover in this process? Where does it stop? Remarkably, individual protein domains assemble themselves.

αɡ
βɔ



e. How do individual protein domains self-assemble?

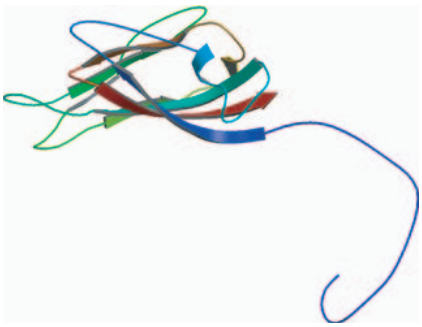
Under physiological conditions, globular proteins do, in fact, assemble themselves into a unique, three-dimensional, biologically relevant structure. The reaction has been given a special name: *protein folding*.

2. Why are macromolecular structures assembled from subunits?

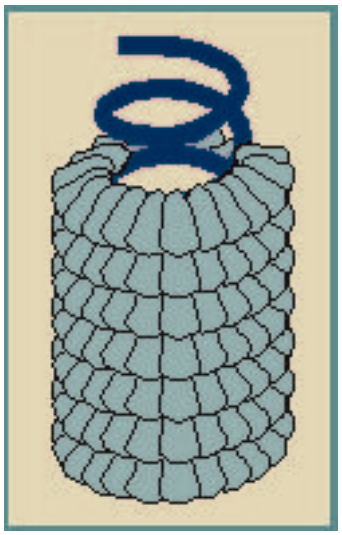
Viral genome = 6390 nucleotides

2130 identical subunits X 158 amino acids = 474 nucleotides (7.4% of genome)

2130 proteins X 158 amino acids = 1,009,620 nucleotides long (~160X genome!)



158 amino acids



2,130 identical subunits

b. Error rate: one error every 3000 residues.
Odds that any given residue is correct:

$$1 - \frac{1}{3000} = 0.99967$$

Probability that one large chain is error-free:
chain = 158 X 2130 = 336,540 residues
probability of error = $.99967^{336540} = 1.87 \times 10^{-49}$

Probability that a module is error-free:
$0.99967^{158} = .949$
~95% of all TMV coat proteins will be perfect.

c. Defective subunits eliminated during assembly
d. Subunits recycled

4. Protein folding

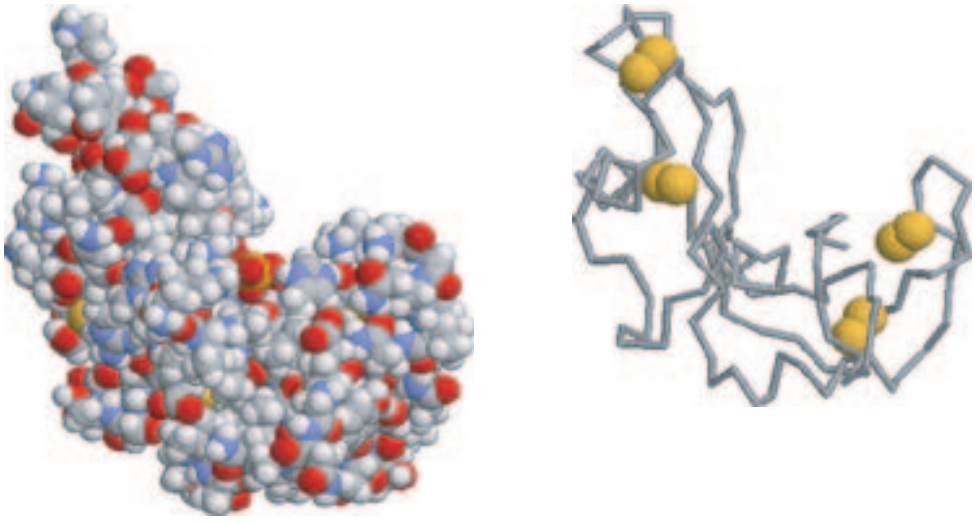
a. Proteins fold spontaneously – the Anfinsen experiment

Upon denaturation, native ribonuclease could spontaneously renature, with return of both structure and enzymatic activity.

124 residues, Mw = 14,000, 8 Cys residues, 4 S-S bridges.

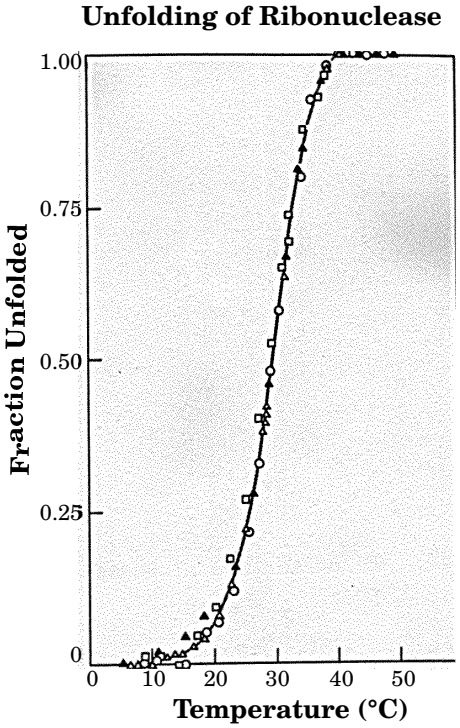
How many ways can 8 Cys residues be arranged into 4 S-S bridges? Ans: 7 x 5 x 3 = 105. Upon renaturation, only the native pattern is detected. Control: oxidation in the presence of 8M urea yields only1% of the native S-S pattern.

Conclusion: protein folding dictated by amino acid sequence alone.



b. Classical picture of protein folding/unfolding

Classic experiment by Ginsburg and Carroll. ribonuclease A unfolds (i.e. melts) as temperature increases. The different tic-marks represent differing probes used to assess the conformation. Each such probe falls on the same curve. This fact implies that there is a single unfolding transition.



... each molecule is either fully folded or fully unfolded and does not linger en route between one state and the other.

$$\text{unfolded} \rightleftarrows \text{folded}$$

$$K = \frac{[\text{folded}]}{[\text{unfolded}]}$$

$$\Delta G^0 = -RT \ln K$$

Typical values of ΔG^0 are quite small, lying in the range -5 to -15 kcal/mol.

Note that this difference in free energy between the folded and unfolded states is* very *small, the energetic equivalent of a few water:water hydrogen bonds.