Interactively exploring supramolecular assembly:
A molecular dynamics approach to virus construction

D. C. Rapaport
Physics Department, Bar-Ilan University, Israel

e-mail: rapaport@mail.biu.ac.il
www: http://www.ph.biu.ac.il/~rapaport
Viruses for pedestrians

- Virus: genetic material (RNA/DNA) packaged inside protein shell (simplest kinds, additional features optional) —

- Many possible geometrical forms for shell; nature tends to use two basic shapes - helical tubes and icosahedral shells; very similar structures occurring in animal/plant/bacterial viruses of different types hints at common design principles.

- Highly symmetric shapes minimize genetic (‘engineering’ design) information required for shell construction, since the same structural unit(s) can be used repeatedly.
• Near-spherical shape provides maximal volume for given surface area.
• Examples: rhinovirus (T=1) with 60 units; cowpea virus (T=3) with 180 units; images by J. Sgro using x-ray crystallography, cryo-electron microscopy —
Another example: adenovirus (T=25) with 252 units (240 × 6-valent, 12 × 5-valent), schematic and electron microscopy —

- Icosahedral symmetry (60-fold, excluding reflection) ubiquitous; each triangular face of the icosahedron consists of 3 smaller asymmetric protein subunits; this corresponds to tiling sphere with 60 identical triangles (cf. fullerenes).
Self-assembly

- Virus shell assembly provides example of the complex processes occurring in the very simplest of (biological) organisms.
- Process lies on the border of chemistry and biology.
- Similar to crystallization: both are governed by thermodynamics and driven by bond formation, but virus growth is self-limiting; goal (?) is minimum free-energy state (reminiscent of protein folding).
- Reversible shell formation is observed in vitro, even without genetic material; protein-protein interactions stabilize structure.
- Assembled shell structures extensively studied; little known about pathways and assembly mechanisms because partial intermediates are highly transient.
Plan

- **Importance:** Virus shells – knowledge of structure, assembly and interactions with cell essential for therapeutic reasons.

- **Assumption:** Given that the protein shells surrounding ‘spherical’ viruses tend to have certain common geometrical features, postulate general assembly mechanism for the ‘capsid’ (the protein coat surrounding the genetic material).

- **Simplification:** Use MD to study self-assembly of simple geometrical shapes (model ‘capsomers’) subject to specified interactions and association rules.

- **Motivation for MD approach:** Since virus assembly pathways difficult to access experimentally, important to have simulation testbed on which alternative construction scenarios can be examined.

- **Extensions:** Possibility of eventual quantitative estimates, but present work considers only simplified ‘geometrical’ models.
Quasi-equivalence

- Internal volume of minimal (60 subunit) shell holds very little DNA/RNA (simple volume calculation ⇒ bound on amount of genetic material within).
- ‘Quasi-equivalence’ principle (explains many observed structures) with > 60 capsomers, the maximum that can have complete equivalent neighborhoods.
- Larger shells consist of $60 \times T$ units, with observed $T \leq 25$ (shell sizes typically in range 20-80 nm).
- Unique structure with 60-fold symmetry made from 12 pentagons and certain numbers of hexagons.
- Inspired by geodesic domes (Buckminster Fuller); such shells have minimal range of variation of bond lengths and angles.
- Cannot deduce assembly rules from final structure; although protein subunits are identical, differentiation (‘autostery’) occurs so they can occupy distinct quasi-equivalent positions (e.g., valence 5 and 6).
- Soccer ball: has 12 pentagons + 20 hexagons; triangulate faces ⇒ 180 facets ⇒ $T=3$.
- Can make paper figures from planar triangular nets using suitable cuts and folds.
- Schematic example: triangulation numbers $T=3$ and $T=1$. 
**Molecular dynamics simulation – some interactive examples**

- Simple fluid . . .
  - MD: classical atomistic approach to studying bulk matter; Newton’s second law governs motion.
  - Monatomic system – short-range Lennard-Jones interaction
    
    \[
    u(r_{ij}) = 4\varepsilon \left[ (\sigma/r_{ij})^{12} - (\sigma/r_{ij})^{6} + \frac{1}{4} \right], \quad r_{ij} \leq r_c = 2^{1/6}\sigma
    \]

  - Efficient computation based on neighbor lists reduces \(O(N^2)\) problem to \(O(N)\).
  - Leapfrog integration of equations of motion.
  - Other details: boundary conditions, initial state, measurements, . . .
  - Study trajectories, temperature and density dependence, fluid-solid transition, etc.
• Rigid-body fluid …
  ◦ Rotational dynamics described by Euler equation.
  ◦ Linear and tetrahedral molecules.

• Flexible chains …
  ◦ Link spherical atoms with rigid bonds, optional attractive forces; can examine helix-coil transition and folding.

• Micelle formation …
  ◦ Simple LJ solvent and solute, short amphiphilic surfactant chains.
  ◦ At low solute concentration observe micelle formation.

• Clusters of tapered rods …
  ◦ Selective pair interactions to ensure alignment.
  ◦ Typical cluster size determined by taper.
• Self-assembly of ‘Soma’ cube with selective interactions; does if happen? …
MD methodology for capsids

- Robustness of self-assembly suggests MD simulations using low resolution models (‘shape’-based caricatures) might capture essence of process.

- First study of this kind (natural extension of old mechanical models for studying structure); provides simulated *in vitro* environment.

- Initial effort focused on assembly of pentagonal subunits into dodecahedra (also tetrahedra and octahedra – start small).
Subunits consist of rigid planar assembly of closely-spaced spheres; vertices are ‘bonding sites’ that interact via LJ potential.

Usually only repulsion (excluded-volume), but attraction allowed between prospective bonding sites – depending on assembly ‘rules’ and subunit status (other bonds, etc.).

Bond formation: when attracting spheres first approach to within bonding range interaction is switched to steep and narrow potential well (unbreakable bond).

Rules determine which spheres can attract at any instant; if all vertex spheres attract ⇒ amorphous globule, but if subunit can bond with just one partner along each edge ⇒ multiply-connected network; rules ensure correctly built structures.

Use of rules eventually becomes unwieldy as shells get larger; increasing computer power allows larger monomers, with multiple interaction sites per face, and so help resolve problems that rules were designed to overcome.
• Techniques (standard MD):
  ◦ Neighbor lists for efficient force computation.
  ◦ Rigid-body dynamics.
  ◦ Leapfrog numerical integration.
  ◦ Since visualization important – both for development and for monitoring behavior – initially used hard (elastically reflecting) wall boundary conditions (periodic boundaries visually confusing).
  ◦ Gradual heating due to exothermal bond formation is avoided by adding a small velocity-dependent damping force to each bond.

• Omissions:
  ◦ No solvent (heavy computational cost).
  ◦ Poor biochemistry: missing ions, genetic material, ‘scaffold’ proteins.
  ◦ No constraints to enforce structural rigidity, or angle-dependent forces to help alignment.
  ◦ Rely on thermal motion to ensure that subunits eventually align, but ‘floppiness’ limits structure size and reduces assembly rate.
- System of 125 pentagons: early random state, and final state with 6 dodecahedra (everything else omitted).
Improved model

- Would like to construct larger polyhedra (e.g., pentakisdodecahedron with 60 triangular faces, and T=1 shell with 60 trapezoidal units):

- Previous model lacks rigidity; also complex *ad hoc* assembly rules (not biologically implausible but preferably avoided).
• New model
  - Several attraction sites located on lateral ‘faces’ of particle (capsomer).
  - Faces are beveled to give correct dihedral angles for closed shell.
  - Example (component balls and effective shape) —

• Torsional forces act during bonding to accelerate process; damping forces remove excess bond vibration energy.
• Attractive forces (simple and effective) —

\[
u(r) = \begin{cases} 
  e \left(1/r_a^2 + r^2/r_h^4 - 2/r_h^2\right) & r < r_h \\
  e \left(1/r_a^2 - 1/r^2\right) & r_h \leq r < r_a
\end{cases}
\]

with (typically) \( e = 0.1, r_h = 0.3, \) and cutoff \( r_a = 2; \) only harmonic part applies for permanent bond.

• Rigidity
  
  ◦ When faces fully bonded, angular fluctuation between neighbors is \( \approx 1^\circ. \)
  
  ◦ Fluctuations in complete assembly are even smaller.

• Assembly pathway
  
  ◦ Assumption: growth of subassembly occurs by adding elements (monomers, or previously formed dimers or trimers); no combining of subassemblies.
  
  ◦ Reduce incorrect assembly by only allowing one element at a time to bond to a growing subassembly, but no imposed assembly sequence.
  
  ◦ Restrict allowed number of subassemblies that can nucleate to ensure adequate number of particles for growth.
Visualization: interactive computer graphics essential – color and various display options for greater information content …
Exploratory results

- Examine variety of different models, parameters and assembly details, until behavior appears promising.

- System size: e.g., 1000 monomers; allow formation of up to 13 polyhedra; use either triangular or trapezoidal ‘capsomers’.

- Complete assembly of all the 60-hedra (in \( \approx 4 \times 10^5 \) timesteps).

- ‘Mutant’ structures rarely form since bond tolerance is tight.

- Color images show development of system (overleaf).

- Can make ‘movies’ to show dynamical history (show recordings); use for quantitative and visual analysis . . .

- Qualitative observation (tentative): partial assemblies tend to be compact rather than ramified (experimental relevance?).

- Growth of shells with 60 triangular and trapezoidal monomers —
• Mutants rarely occur; example shows growth of 2nd shell (101 subunits, but could continue forever) —

• Ability to encourage minor structural changes has potential medical value.
Bigger shells

- Shell of size 180 (T=3) with 3 slightly different monomer types, to mimic effect of quasi-equivalence (without ‘autostery’) —

- Need to be more specific as to which faces can bond —
- Growth pathway can involve small intermediate structural elements, e.g., dimers or trimers —

- System of 4096 monomers $\Rightarrow$ 10 T=3 shells via dimer pathway . . .
• Quantitative aspects: measure number of complete shells vs time (permanent bonding); observe slowdown near completion (‘starvation’) —
Reversible growth

- Discard irreversible bonding; allow bonds to form and break reversibly without restriction (incorrect, energetically-unfavorable partial structures self-disassemble).
- Larger systems and longer runs needed for adequate yield.
- Destroy all shells below a certain size at regular intervals to provide monomers so larger shells can continue to grow.
- Monomer - details (note ‘thickness’ of unit) —

For T=1 shells: 4096 monomers ⇒ 45 complete shells over $10^7$ timesteps; interactions tailored to encourage initial trimer formation . . .
- Can use periodic boundaries to avoid wall bunching ...

- Measure number of complete T=1 shells vs time (reversible bonding) and capsomer fraction with different bond counts (≤ 5) vs size —
• Measure cluster-size distribution vs time (trimer-weighted growth) —
Other kinds of self-assembly

- Simple lattice elements and cylindrical objects...
**Future work**

- Study populations and lifetimes of intermediate states; perhaps relate to scattering data.
- Sensitivity to particular subunit representation and choice of interactions?
- Coexistence of multiple pathways, with competition between different routes?
- Use of ‘continuous feed’ approach, with removal of completed products.
- Use of experimental / computed bonding energies to emulate real virus capsomers.
- Mechanism for quasi-equivalence to appear ‘naturally’.

**References**