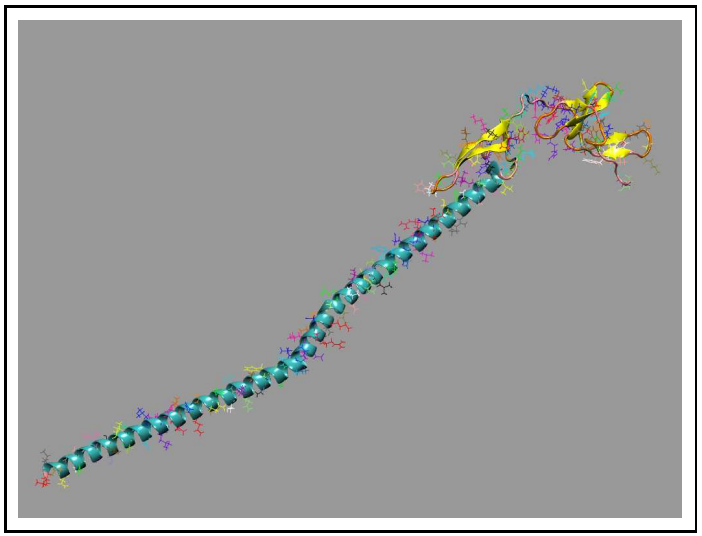
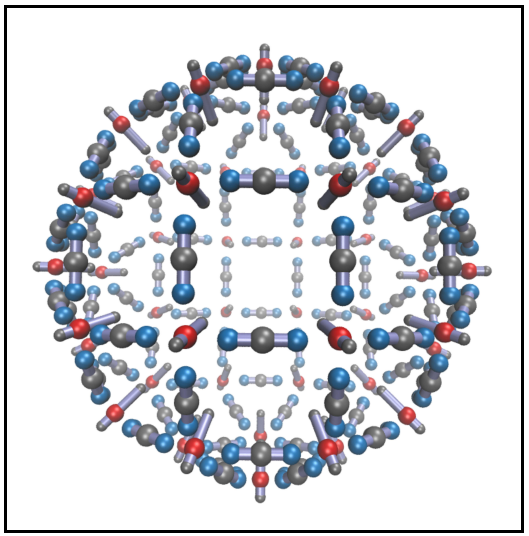
Energy Landscapes for Self-Assembly

This conceptual and computational framework is based on stationary points (minima and transition states) of the potential energy surface.

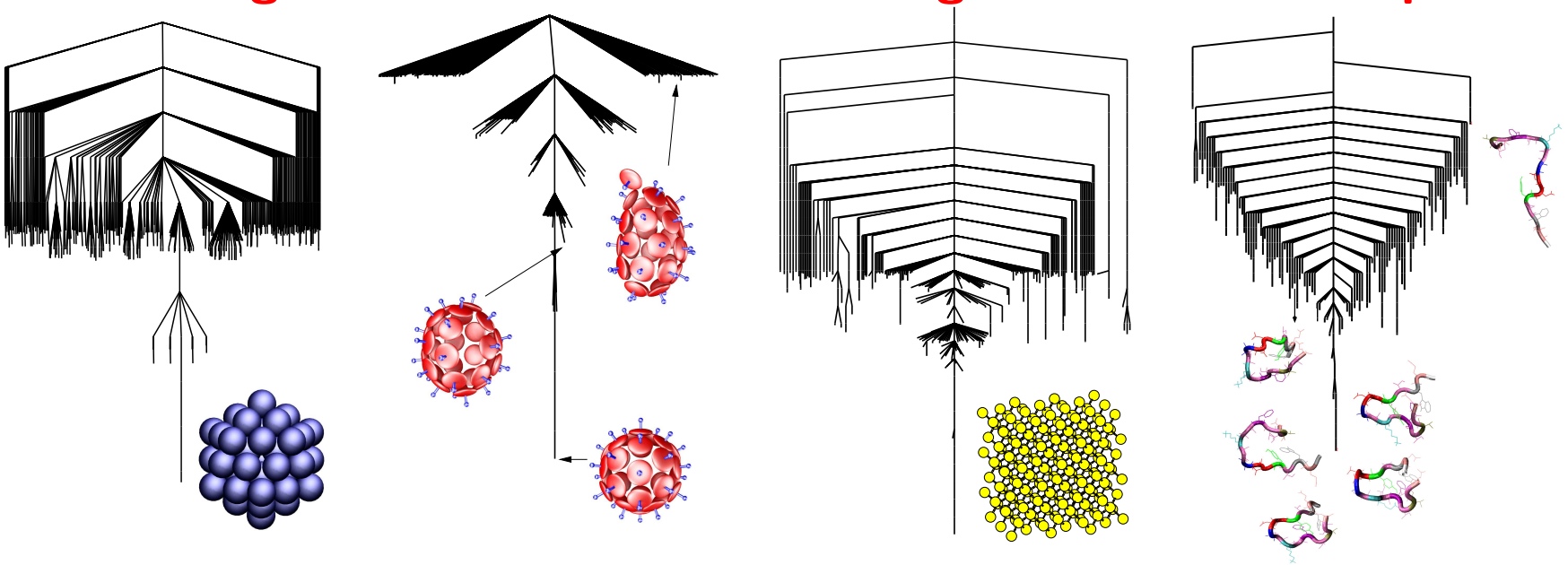
- Basin-hopping for global optimisation (*J. Phys. Chem. A*, **101**, 5111 1997)
- Basin-sampling for global thermodynamics (*J. Chem. Phys.*, **124**, 044102, 2006)
- Discrete path sampling for global kinetics (*Mol. Phys.*, **100**, 3285, 2002)

For small molecules all the relevant stationary points and pathways can be located. Larger systems require appropriate sampling.



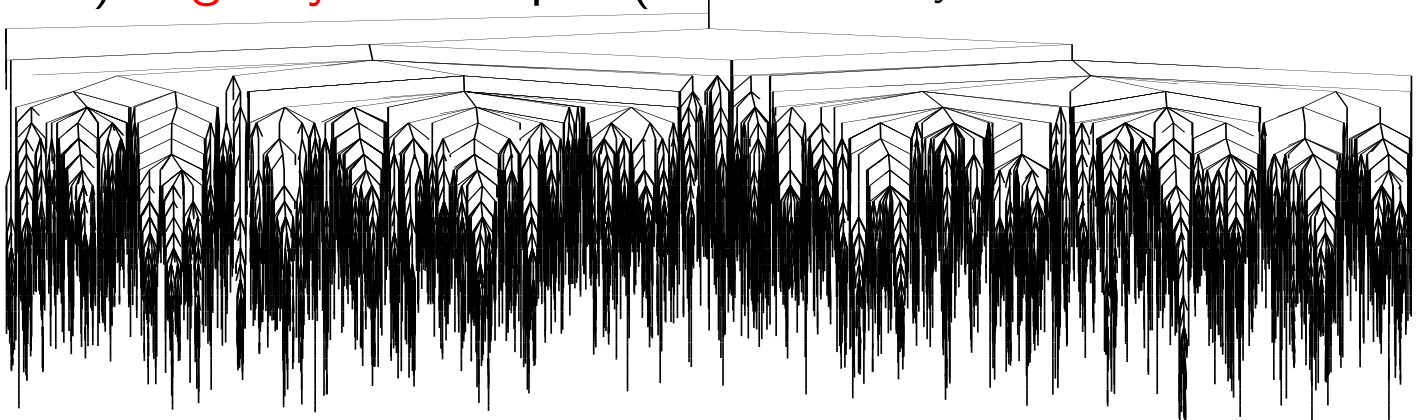


Self-Organisation is Encoded in Single Funnel Landscapes



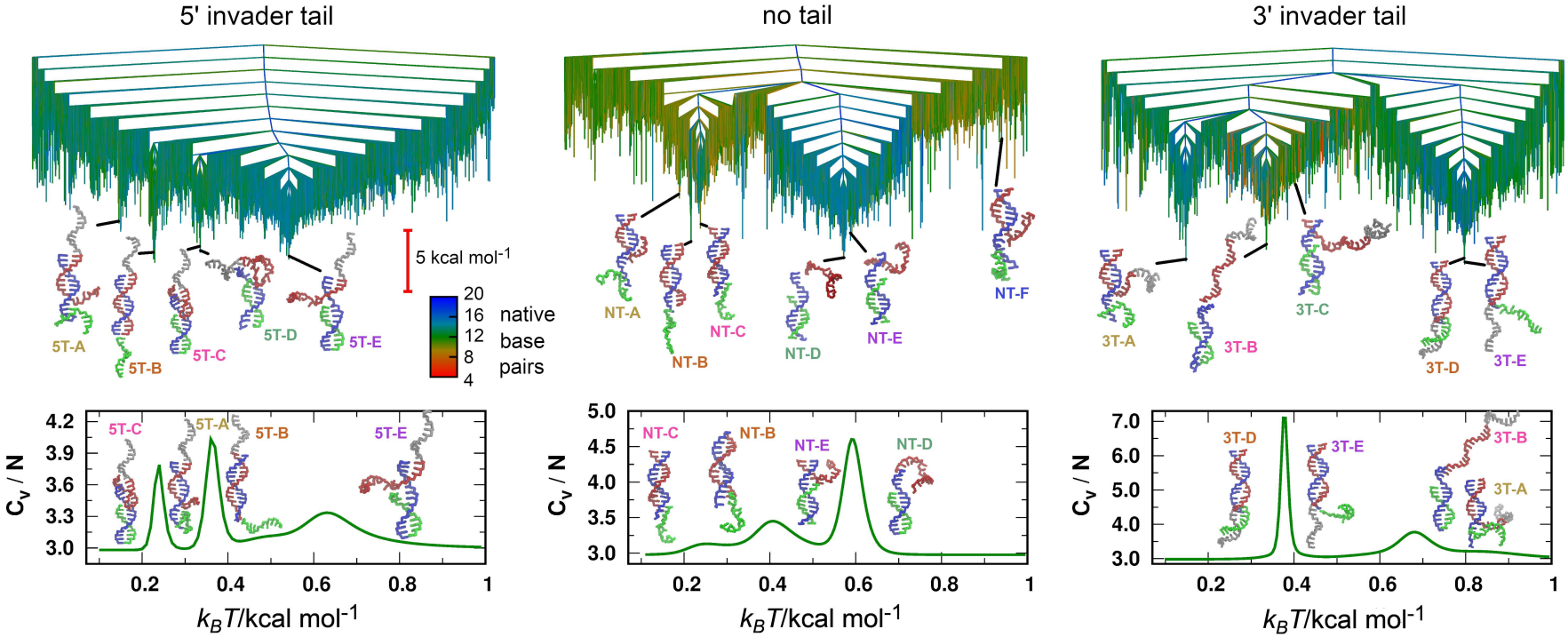
(Above) Landscapes for self-organising systems (*Nature*, **394**, 758, 1998).

(Below) A glassy landscape. (*Phil. Trans. Roy. Soc. A*, **363**, 357, 2005).



Molecular switches and intrinsically multifunctional molecules correspond to multifunnel landscapes (*Adv. Theory and Simulations*, **2**, 1800175, 2019).

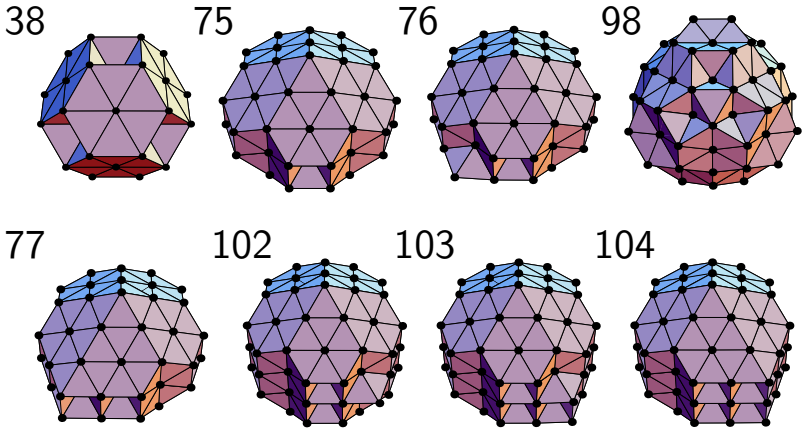
Dynamic DNA Nanotechnology: Toehold Strand Displacement



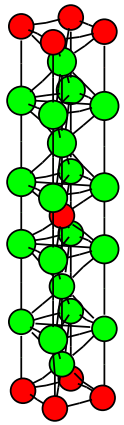
Chemical reaction networks can be programmed using DNA strand displacement reactions. An invading strand displaces the incumbent strand from substrate via a sticky toehold domain.

5′ and 3′ tails on the invader accelerate and retard the displacement, respectively, allowing kinetic modulation of a catalytic circuit for signal amplification or pulse generation.

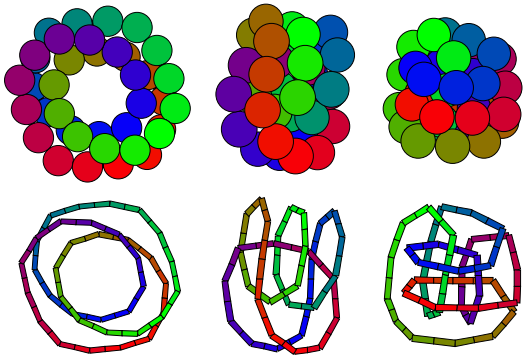
Basin-Hopping Global Optimisation (*J. Phys. Chem. A*, **101**, 5111, 1997)



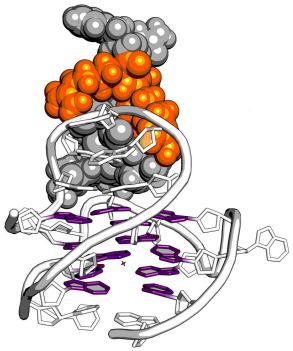
non-icosahedral Lennard-Jones clusters



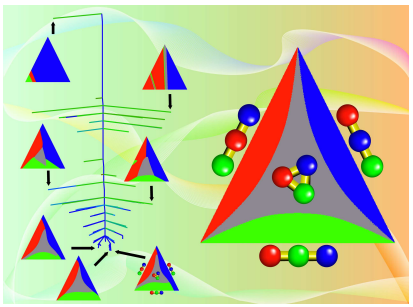
binary LJ crystal



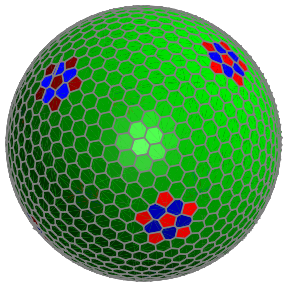
Stockmayer particles



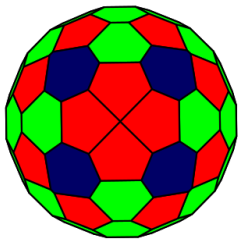
bicyclic ligands for G-quadruplexes



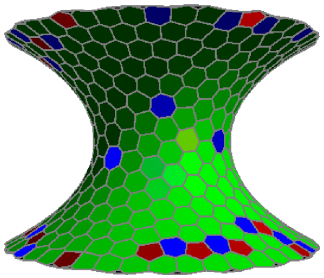
neural network



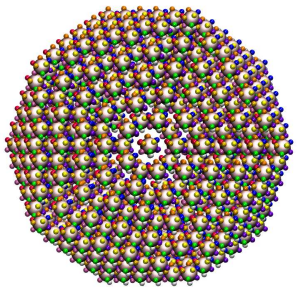
Thomson problem



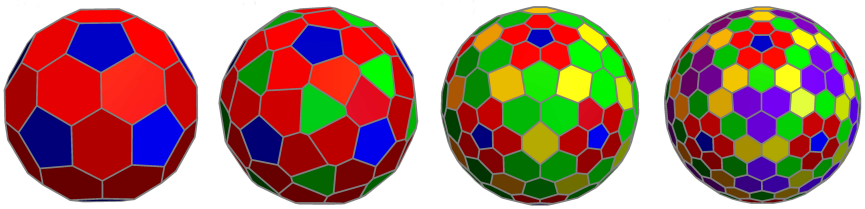
bidisperse



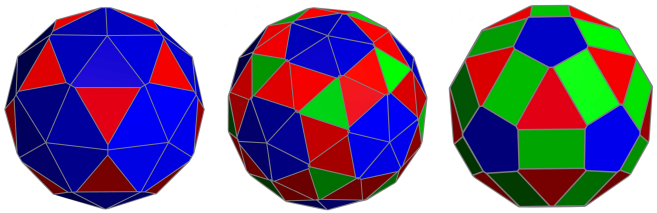
unduloid capillary bridge

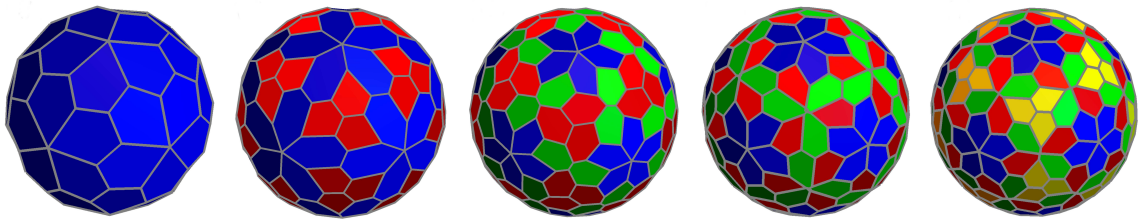


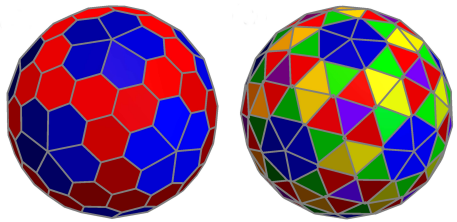
addressable



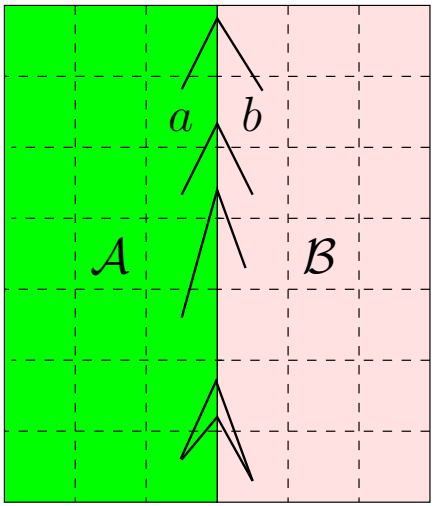
icosahedral shells



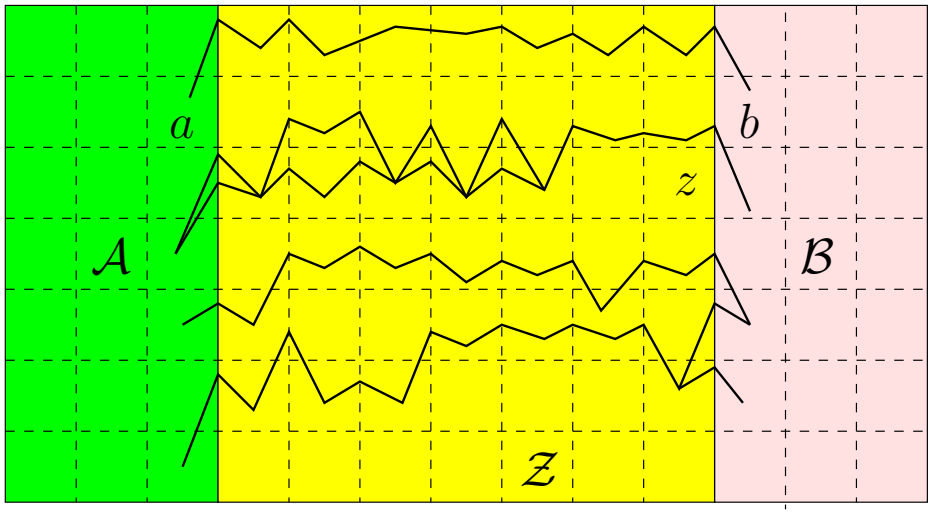




Discrete Path Sampling (*Mol. Phys.*, **100**, 3285, 2002; **102**, 891, 2004).



no intervening minima

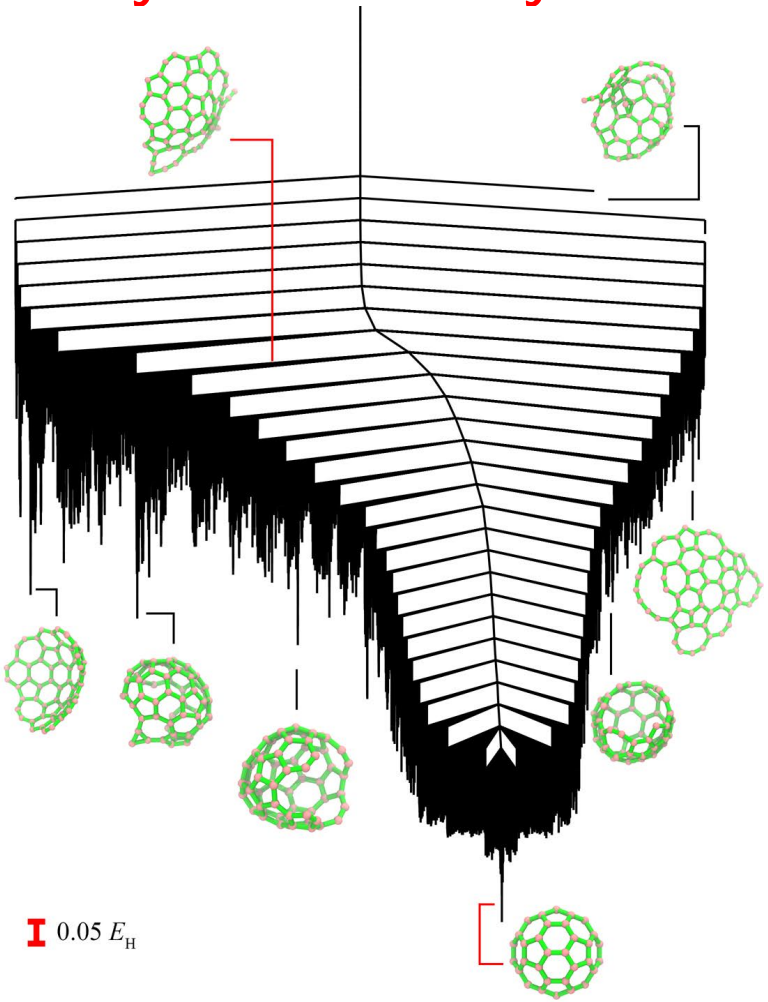


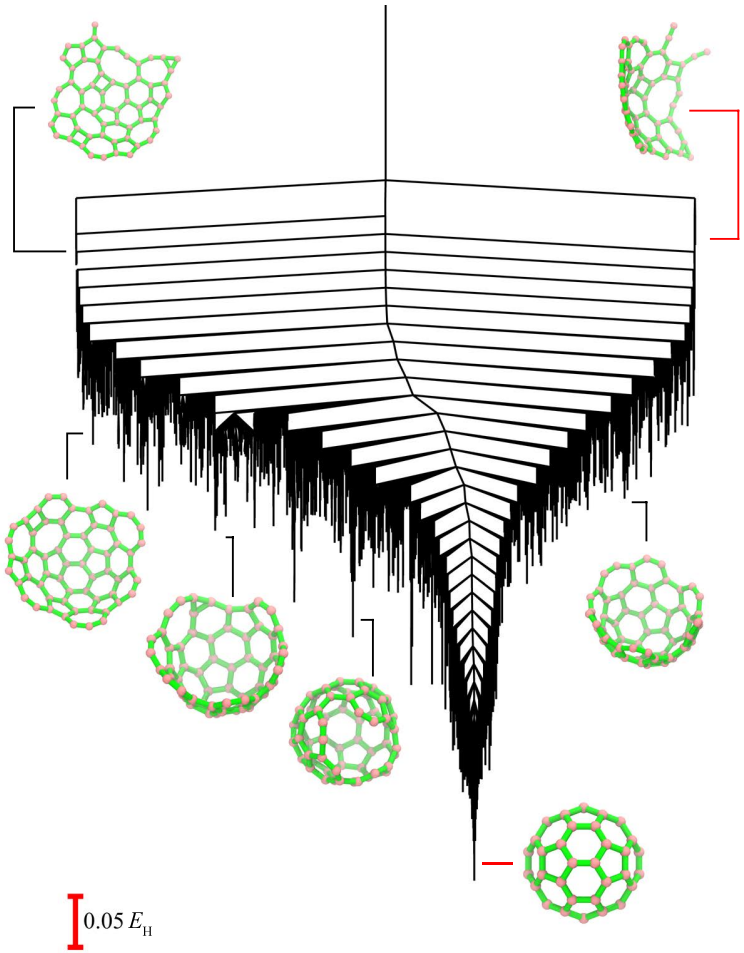
products intervening minima reactants

Sample kinetic transition networks (local minima and transition states) using geometry optimisation. No reaction coordinate and no projection.

Mean first passage time from $\mathcal{S} = \mathcal{Z} \cup \mathcal{B}$ to $\mathcal{A}$ is $\mathcal{T}_{\mathcal{AS}} = \boldsymbol{\tau}_{\mathcal{S}} \mathbf{G}_{\mathcal{S}} \mathbf{P}_{\mathcal{S}}(0)$, with waiting times $\tau_s = 1/\sum_\alpha K_{\alpha s}$, fundamental matrix $\mathbf{G}_{\mathcal{S}} = [\mathbb{I}_{\mathcal{S}} - \mathbf{B}_{\mathcal{SS}}]^{-1}$, Markov rate matrix $\mathbf{K}$, initial occupation probabilities $P_s(0)$, identity matrix $\mathbb{I}_{\mathcal{S}}$, and branching probability $B_{s's} = K_{s's}\tau_s$.

Buckybowl to Buckyball Pathways (*J. Phys. Chem. A*, **126**, 2342, 2022)





We compared two quantum-based potentials: SCC-DFTB and GFN2-xTB.

The highest energy structures are buckybowls with dangling chains.

The fastest paths between them each have around 50 transition states.

Decoding Chemical Kinetics (*JPCL*, **13**, 6349, 2022)

The first passage time distribution $p(t)$ can be written as

$$p(t) = \sum_\ell \lambda_\ell e^{-\lambda_\ell t} A_\ell, \quad \text{and for } y = \ln t, \quad \mathcal{P}(y) = \sum_\ell \lambda_\ell e^{y - \lambda_\ell \exp(y)} A_\ell,$$

where $-\lambda_\ell < 0$ are the eigenvalues of the matrix defining the master equation dynamics for absorbing products, and the A_ℓ are amplitudes, which depend on the eigenvectors and the initial condition.

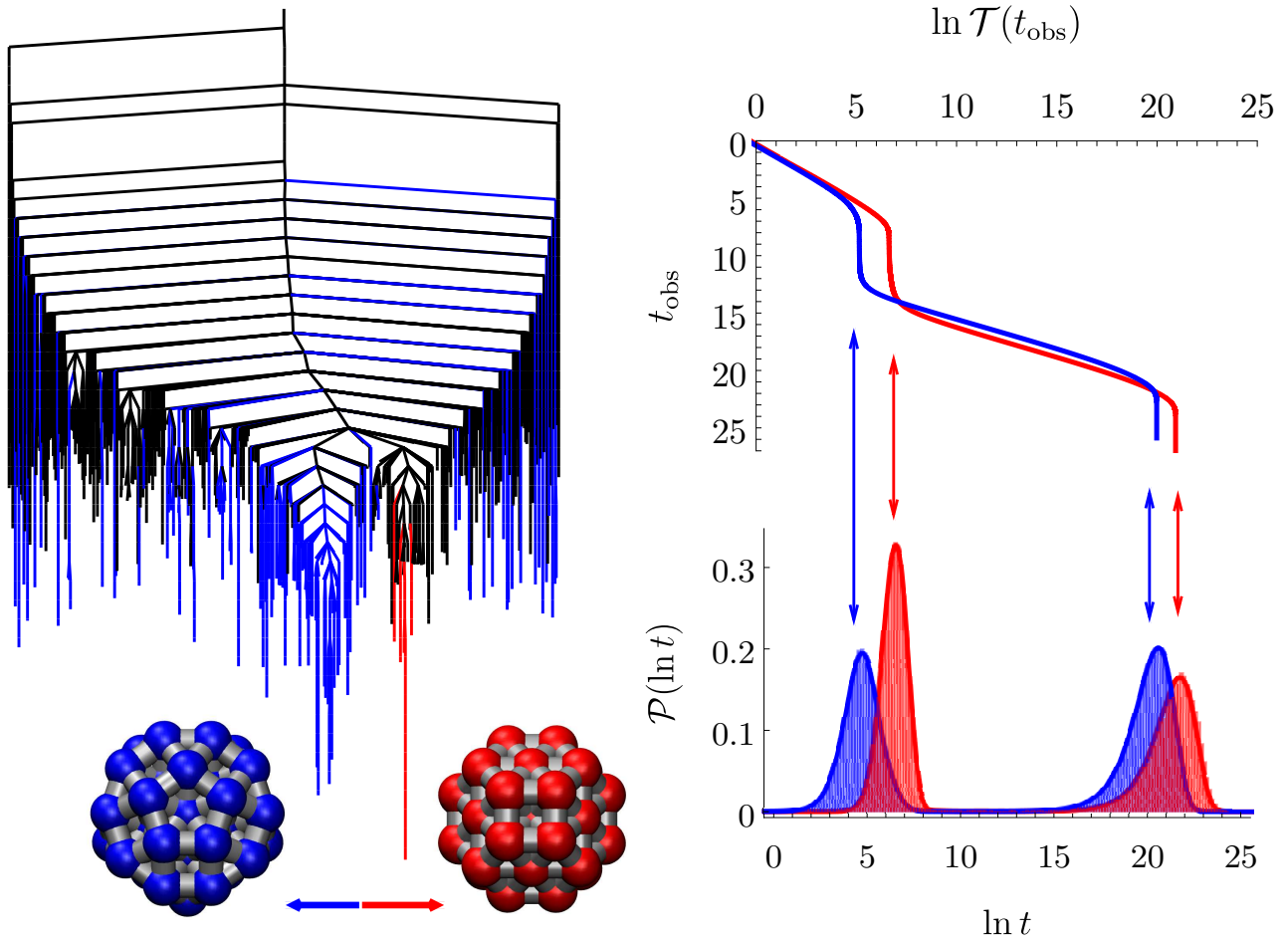
The mean first passage time can be defined for observation timescale $t_{\rm obs}$:

$$\mathcal{T}(t_{\rm obs}) = \sum_\ell \frac{A_\ell}{\lambda_\ell z(t_{\rm obs})} \left[1 - e^{-\lambda_\ell t_{\rm obs}}(1 + \lambda_\ell t_{\rm obs})\right],$$

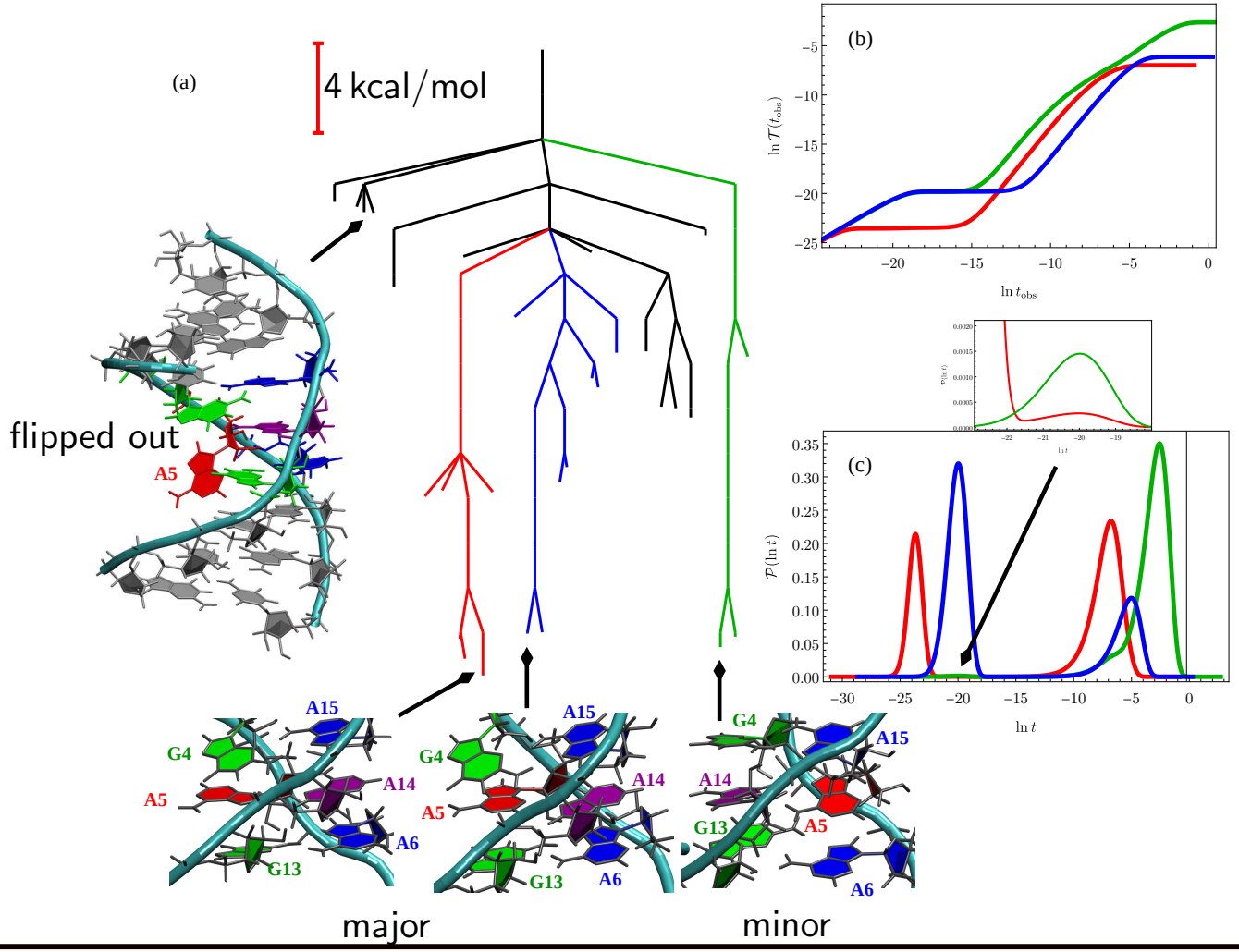
where $z(t_{\rm obs})$ is the normalisation for the restricted distribution.

$$\lim_{t_{\rm obs}\to\infty} \mathcal{T}(t_{\rm obs}) = \sum_{\ell=1}^{\ell_{\rm max}} \frac{A_\ell}{\lambda_\ell} \Big/ \sum_{\ell=1}^{\ell_{\rm max}} A_\ell.$$

$\mathcal{T}(t_{\rm obs})$ exhibits steps corresponding to escape from kinetic traps where we sum up to $\ell_{\rm max}$, $\ell_{\rm max} - 1$, $\ell_{\rm max} - 2 \ldots$, with $\ell_{\rm max}$ the slowest relaxation.

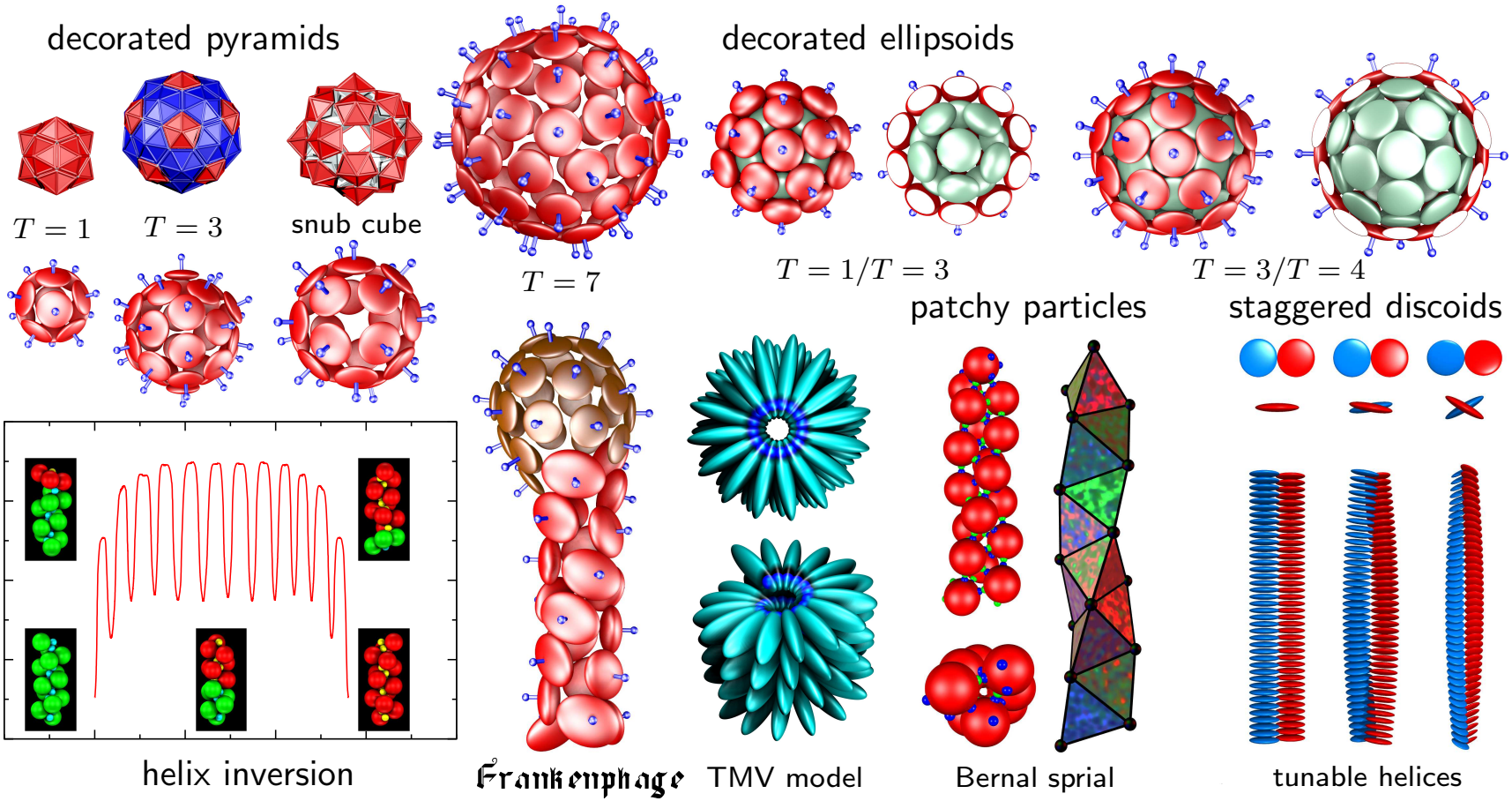


In the double-funnel LJ_{38} cluster two timescales appear in $\mathcal{P}(\ln t)$ and $\mathcal{T}(t_{\rm obs})$ for relaxation from a high energy minimum to the competing close-packed and icosahedral structures. The longer time scale corresponds to switching between morphologies in each case.



The landscape for a nine base pair RNA duplex with a non-canonical adenine-adenine contact exhibits major and minor forms with A14 or A5 stacked between A6 and A15. Three distinct time scales are discernible for relaxation to the minor form, with corresponding steps in $\ln \mathcal{T}(t_{\rm obs})$.

Coarse-Grained Models (*PCCP*, **11**, 1970, 2009; *ACS Nano*, **4**, 219, 2010)



The angle-axis formulation provides a compact description for mesoscopic modelling of decorated rigid bodies and ellipsoids.

All the angle-axis coordinates are obtained by the action of a rotation matrix and its derivatives, which are system-independent.

1st derivatives: $\mathbf{R}_k \equiv \dfrac{\partial \mathbf{R}}{\partial p_k} = \dfrac{p_k \sin\theta}{\theta}\,\widetilde{\mathbf{p}}^2 + (1-\cos\theta)(\widetilde{\mathbf{p}}_k\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_k) + \dfrac{p_k\cos\theta}{\theta}\widetilde{\mathbf{p}} + \sin\theta\,\widetilde{\mathbf{p}}_k$, with $\widetilde{\mathbf{p}}_1 = \dfrac{1}{\theta^3}\begin{pmatrix} 0 & p_1p_3 & -p_1p_2 \\ -p_1p_3 & 0 & p_1^2-\theta^2 \\ p_1p_2 & \theta^2-p_1^2 & 0 \end{pmatrix}$

2nd derivatives : $\mathbf{R}_{kk} \equiv \dfrac{\partial^2\mathbf{R}}{\partial p_k^2} = \dfrac{2p_k\sin\theta}{\theta}(\widetilde{\mathbf{p}}_k\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_k) + \left(\dfrac{p_k^2\cos\theta}{\theta^2} - \dfrac{p_k^2\sin\theta}{\theta^3} + \dfrac{\sin\theta}{\theta}\right)\widetilde{\mathbf{p}}^2$

$+\ (1-\cos\theta)(2\widetilde{\mathbf{p}}_k^2 + \widetilde{\mathbf{p}}_{kk}\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_{kk}) + (-\dfrac{p_k^2\sin\theta}{\theta^2} - \dfrac{p_k^2\cos\theta}{\theta^3} + \dfrac{\cos\theta}{\theta})\widetilde{\mathbf{p}} + \dfrac{2p_k\cos\theta}{\theta}\widetilde{\mathbf{p}}_k + \sin\theta\,\widetilde{\mathbf{p}}_{kk}$,

and $\mathbf{R}_{kl} \equiv \dfrac{\partial^2\mathbf{R}}{\partial p_k p_l} = \dfrac{p_k\sin\theta}{\theta}(\widetilde{\mathbf{p}}_l\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_l) + (\dfrac{p_kp_l\cos\theta}{\theta^2} - \dfrac{p_kp_l\sin\theta}{\theta^3})\widetilde{\mathbf{p}}^2 + \dfrac{p_l\sin\theta}{\theta}(\widetilde{\mathbf{p}}_k\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_k)$

$+\ (1-\cos\theta)(\widetilde{\mathbf{p}}_{kl}\widetilde{\mathbf{p}} + \widetilde{\mathbf{p}}_k\widetilde{\mathbf{p}}_l + \widetilde{\mathbf{p}}_l\widetilde{\mathbf{p}}_k + \widetilde{\mathbf{p}}\widetilde{\mathbf{p}}_{kl}) - (\dfrac{p_kp_l\sin\theta}{\theta^2} + \dfrac{p_kp_l\cos\theta}{\theta^3})\widetilde{\mathbf{p}} + \dfrac{p_k\cos\theta}{\theta}\widetilde{\mathbf{p}}_l + \dfrac{p_l\cos\theta}{\theta}\widetilde{\mathbf{p}}_k + \sin\theta\,\widetilde{\mathbf{p}}_{kl}$.

Denote positions in the body-fixed frame by superscript 0. For rigid bodies I and J with sites i and j defining site-site isotropic potentials U_{ij}^{IJ} the potential energy is

$$U = \sum_I \sum_{J<I} \sum_{i\in I} \sum_{j\in J} f_{ij}(r_{ij}), \qquad \text{where} \qquad r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_i - \mathbf{r}_j| \qquad \text{and} \qquad f_{ij} \equiv U_{ij}^{IJ} \qquad \text{so that}$$

$$\frac{\partial U}{\partial \zeta} = \sum_{J\neq I} \sum_{i\in I} \sum_{j\in J} f'_{ij}(r_{ij})\frac{\partial r_{ij}}{\partial \zeta}, \quad \text{where} \quad f'_{ij} = \frac{df_{ij}(r_{ij})}{dr_{ij}}, \quad \frac{\partial r_{ij}}{\partial \mathbf{r}^I} = \hat{\mathbf{r}}_{ij}, \quad \frac{\partial r_{ij}}{\partial p_k^I} = \hat{\mathbf{r}}_{ij}\cdot\frac{\partial \mathbf{r}_{ij}}{\partial p_k^I} = \hat{\mathbf{r}}_{ij}\cdot(\mathbf{R}_k^I\mathbf{r}_i^0), \quad \mathbf{r}_{ij} = \mathbf{r}^I + \mathbf{R}^I\mathbf{r}_i^0 - \mathbf{r}^J - \mathbf{R}^J\mathbf{r}_j^0.$$

$$\frac{\partial^2 U_{ij}^{IJ}}{\partial r_k^I \partial r_l^J} = f_2(r_{ij})r_{ij,k}r_{ij,l}\epsilon_{IJ} + f_1(r_{ij})\delta_{kl}\epsilon_{IJ},$$

$$\frac{\partial^2 U_{ij}^{IJ}}{\partial p_k^I \partial p_l^J} = f_2(r_{ij})(\mathbf{r}_{ij}\cdot\mathbf{R}_k^I\mathbf{r}_i^0)(\mathbf{r}_{ij}\cdot\mathbf{R}_l^I\mathbf{r}_i^0)\delta_{IJ} - f_2(r_{ij})(\mathbf{r}_{ij}\cdot\mathbf{R}_k^I\mathbf{r}_i^0)(\mathbf{r}_{ij}\cdot\mathbf{R}_l^J\mathbf{r}_j^0)(1-\delta_{IJ}) + f_1(r_{ij})(\mathbf{R}_k^I\mathbf{r}_i^0)\cdot(\mathbf{R}_l^I\mathbf{r}_i^0)\delta_{IJ}$$
$$-f_1(r_{ij})(\mathbf{R}_k^I\mathbf{r}_i^0)\cdot(\mathbf{R}_l^J\mathbf{r}_j^0)(1-\delta_{IJ}) + f_1(r_{ij})(\mathbf{r}_{ij}\cdot\mathbf{R}_{kl}^I\mathbf{r}_i^0)\delta_{IJ},$$

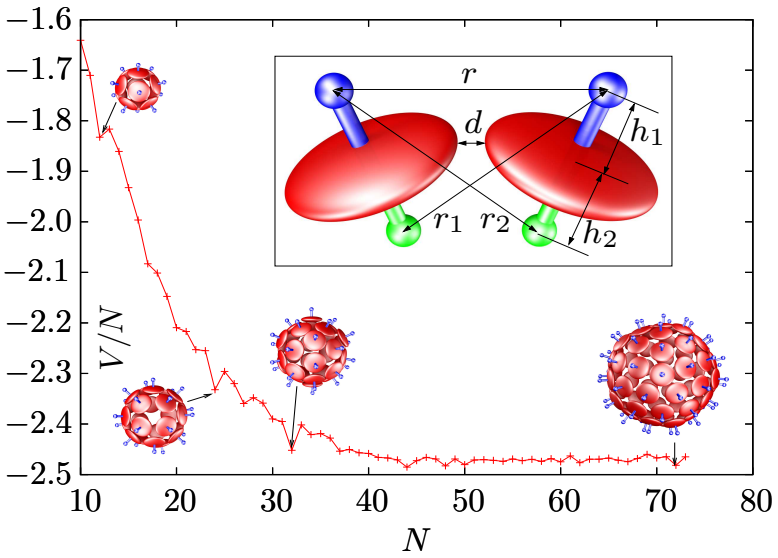
$$\frac{\partial^2 U_{ij}^{IJ}}{\partial r_k^I \partial p_l^J} = f_2(r_{ij})(\mathbf{r}_{ij}\cdot\mathbf{R}_l^I\mathbf{r}_i^0)r_{ij,k}\delta_{IJ} - f_2(r_{ij})(\mathbf{r}_{ij}\cdot\mathbf{R}_l^J\mathbf{r}_j^0)r_{ij,k}(1-\delta_{IJ}) + f_1(r_{ij})[\mathbf{R}_k^I\mathbf{r}_i^0]_l\delta_{IJ} - f_1(r_{ij})[\mathbf{R}_l^J\mathbf{r}_j^0]_l(1-\delta_{IJ}).$$

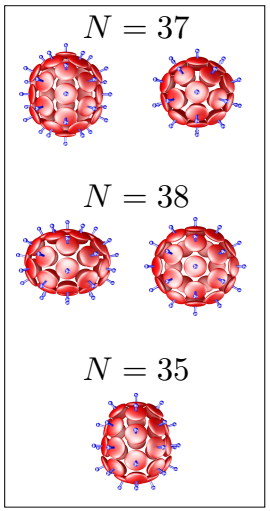
where $f_1(r_{ij}) = f'_{ij}(r_{ij})/r_{ij}$, $f_2(r_{ij}) = f'_1(r_{ij})/r_{ij}$, $\epsilon_{IJ} = 1$ for $I = J$ and $\epsilon_{IJ} = -1$ for $I \neq J$, and δ_{IJ} is the Kronecker delta.

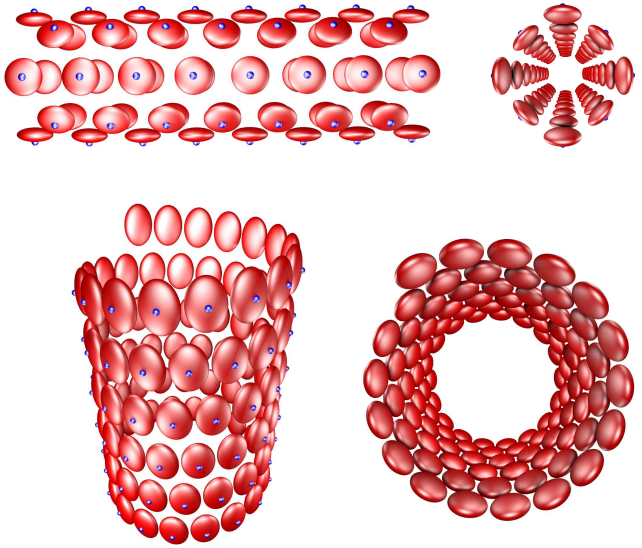
Emergent Behaviour from Simple Models (*ACS Nano*, **4**, 219, 2010)

Adding two repulsive axial Lennard-Jones sites to an ellipsoidal core produces versatile building blocks. Oblate ellipsoids favour shells; stronger repulsion for the longer semiaxis produces tubes and spirals.

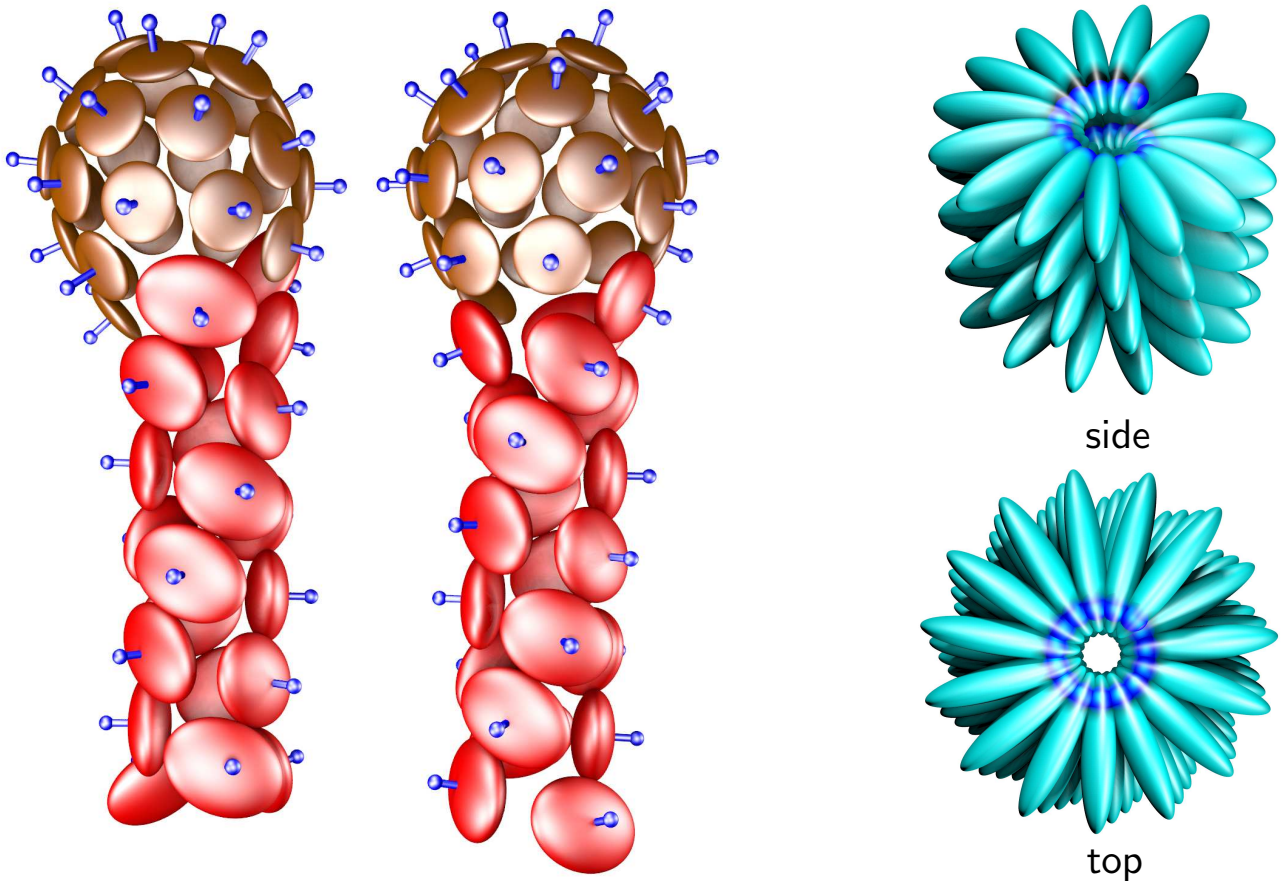
The global minima for oblate ellipsoids are icosahedra for $N = 12$, 32 and 72 ($T = 1, 3$ and 7), the snub cube observed for polyoma virus capsids at $N = 24$, and conical, biaxial, prolate, and oblate shells at other sizes.







Modelling Mesoscopic Structures (*ACS Nano*, **4**, 219, 2010)



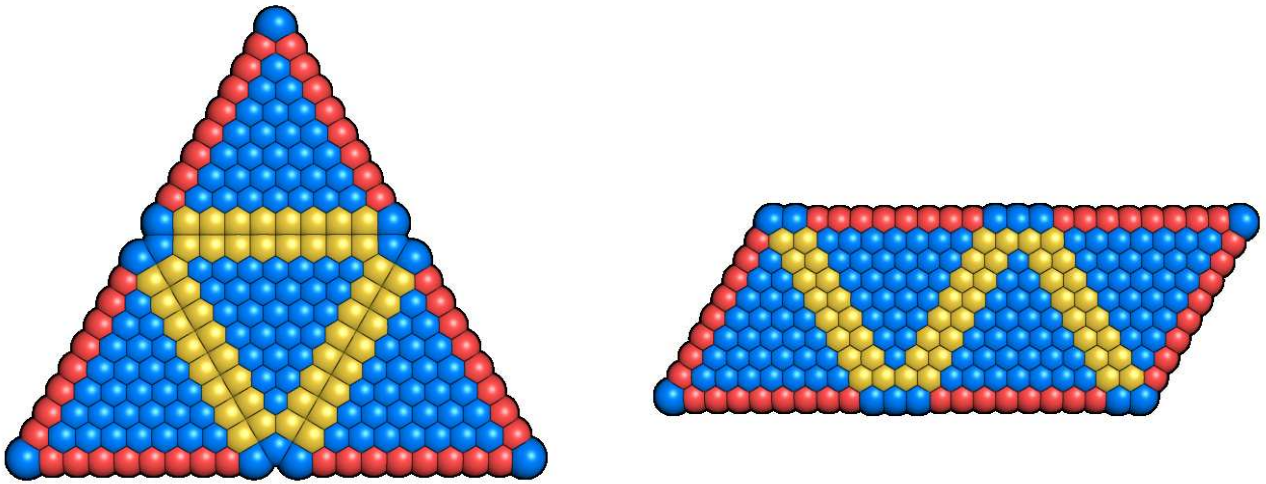
Mixing ellipsoidal building blocks that favour shells and tubes produces structures with distinct head and tail regions (left): the Frankenphage.

Particles with a Lennard-Jones site buried in the ellipsoid assemble into a spiral structure (right) that matches tobacco mosaic virus.

Polyhedral Nets

Successful self-assembly of 200 and 500 μm cubes and octahedra from tethered, multi-component nets was found to correlate with compactness.

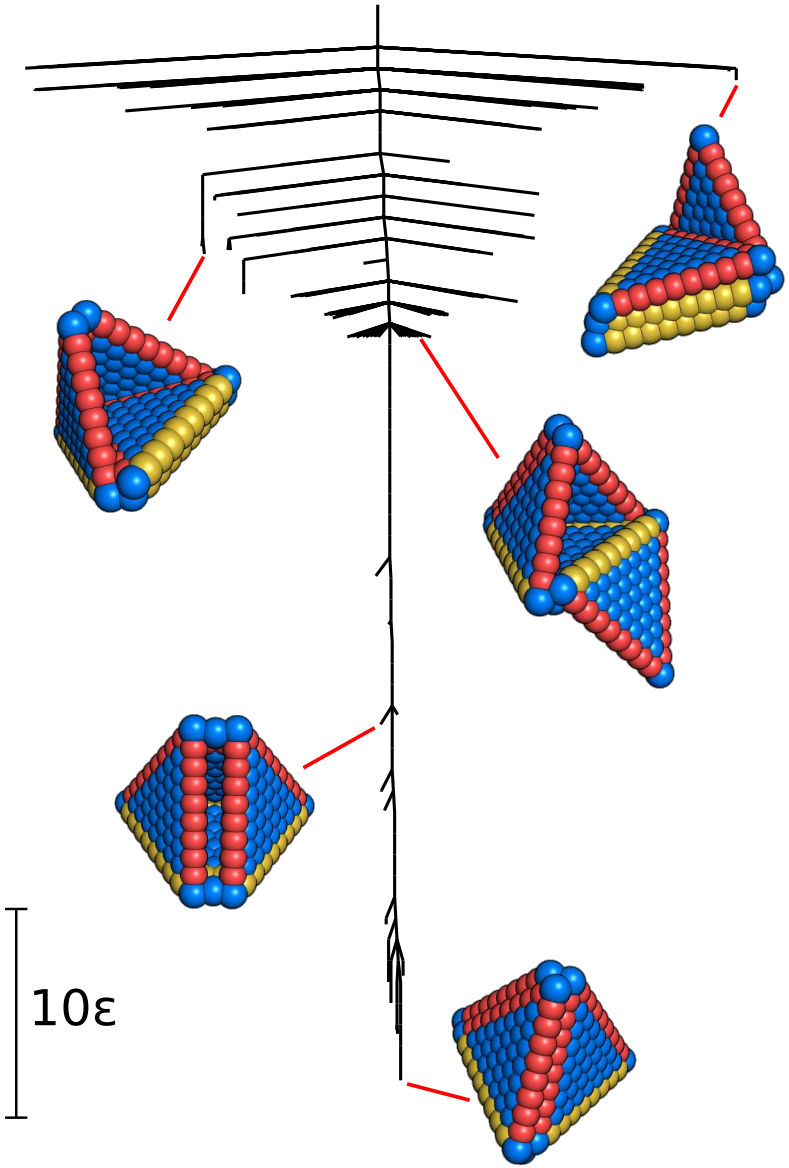
(Gracias and Menon, *Artificial Life*, **20**, 409, 2014; *PNAS*, **108**, 19885, 2011.)

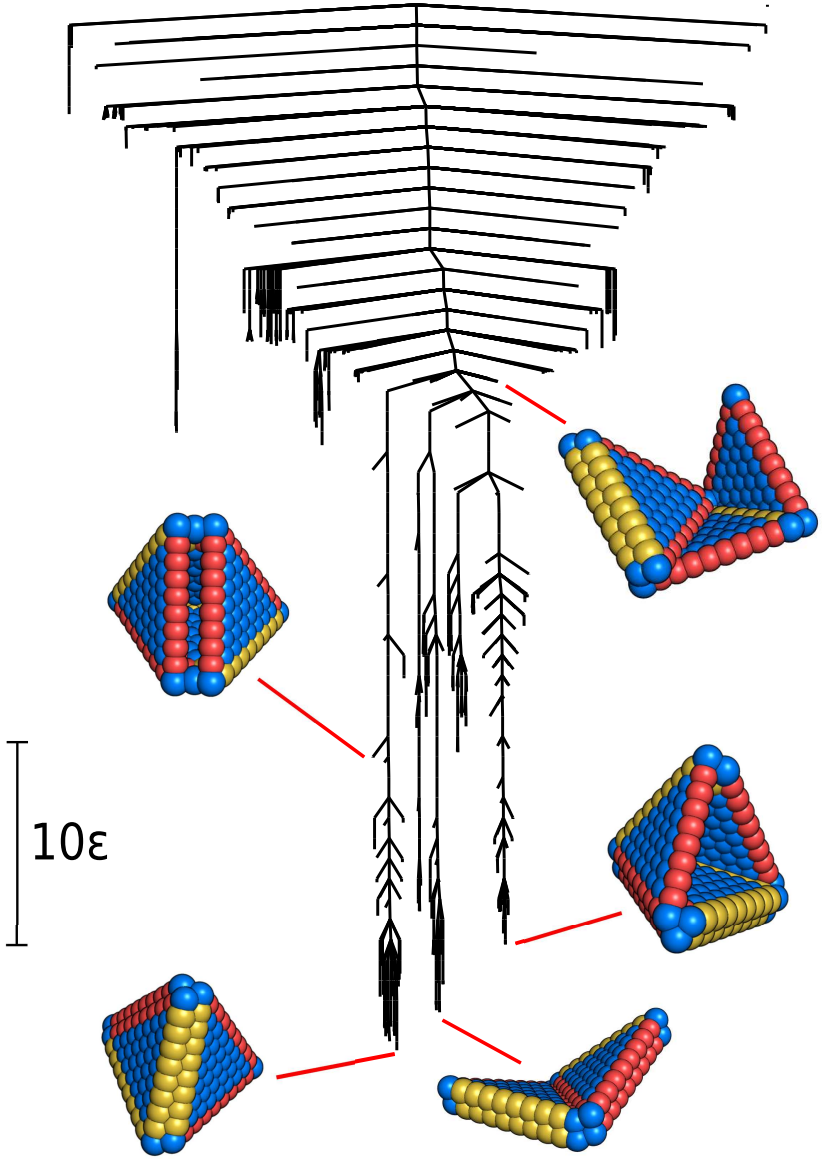


Model potential (Dodd, Damasceno and Glotzer, *PNAS*, **115**, E6690, 2018) is based on

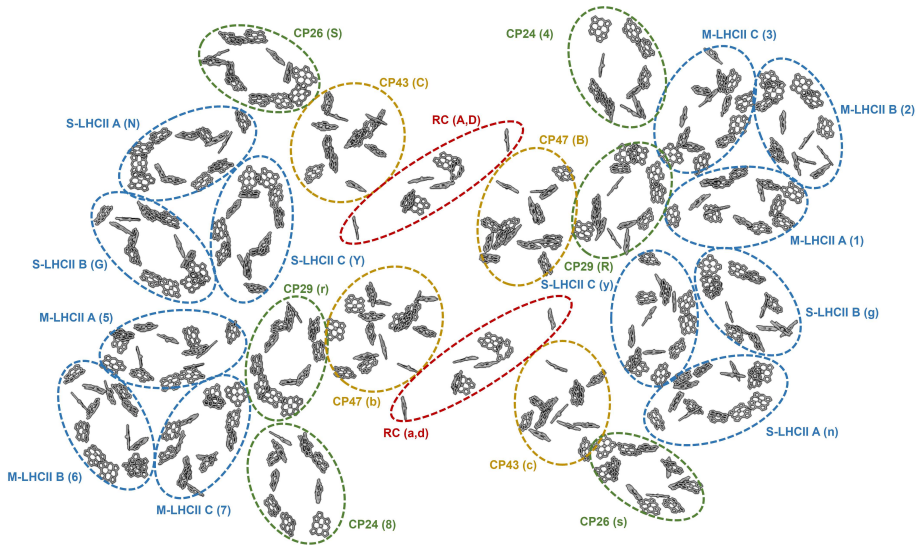
- Repulsive interior sites - Weeks-Chandler-Andersen potential
- Sticky sites along free edges (locking hinges) - Lennard-Jones potential
- Stiff harmonic springs along fixed edges (folding hinges)

Self-assembly of a tetrahedron is more efficient for the triangular net (left) than the frustrated linear net (right).



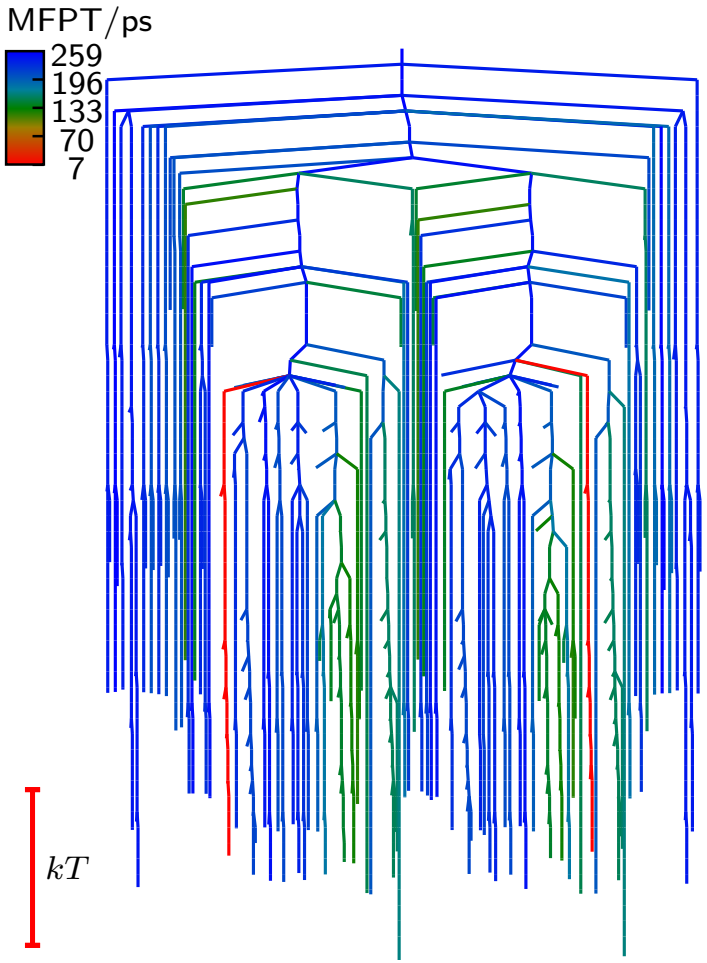


Light Harvesting in Photosystem II Supercomplexes



Photosynthesis begins with absorption of light by pigments bound to proteins of PSII coupled to major light harvesting complexes. The excitation energy is transferred to a reaction centre and converted to chemical energy via charge separation.

By analysing the kinetics in mutants we aim to understand the functional roles of individual proteins in energy transfer and photoprotection.

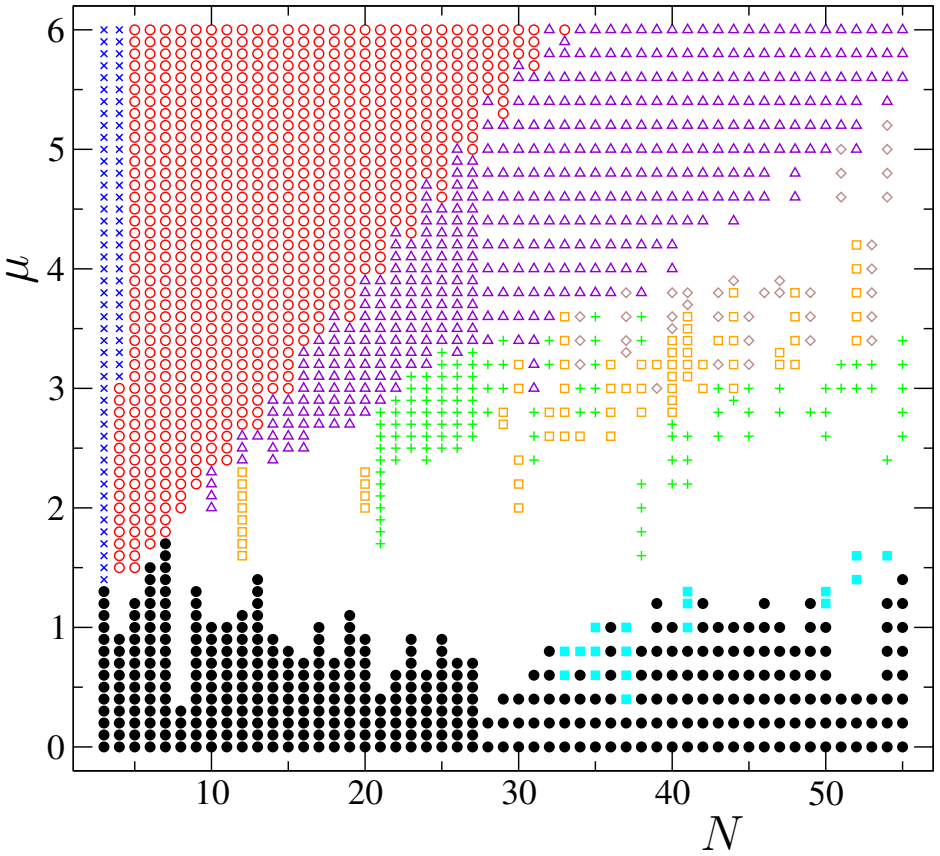


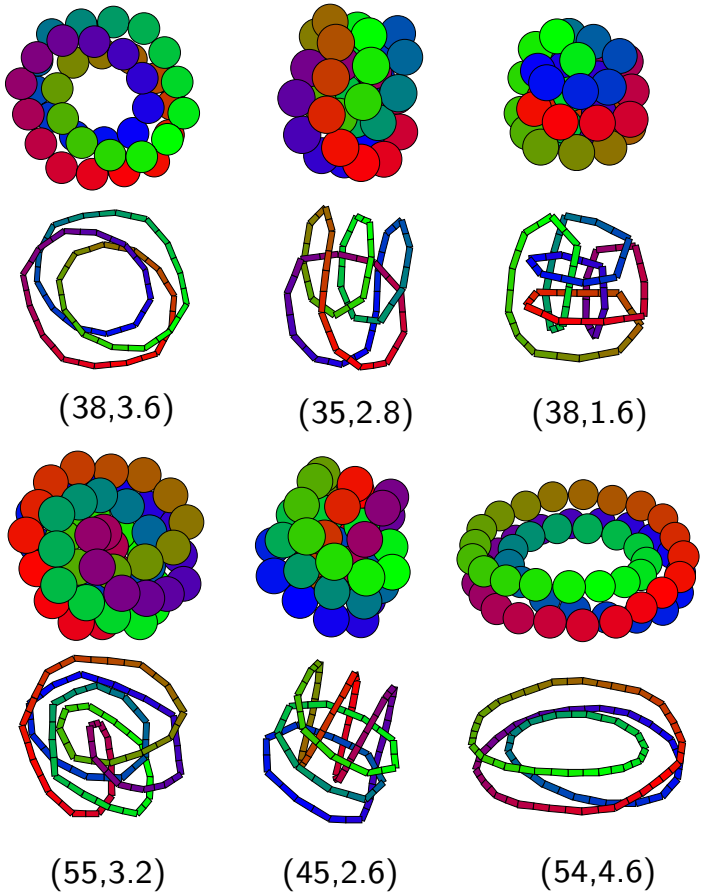
The free energy disconnectivity graph is constructed from rates calculated using time-dependent perturbative approaches: modified Redfield theory for intradomain electron-phonon interactions; generalized Förster theory for interdomain coulomb interactions between chlorophylls.

The landscape does not support kinetic traps, in contrast to systems featuring rare event kinetics. The mean first passage time for an excitation to reach a reaction centre is relatively homogeneous.

Predictions agree with two-dimensional electronic-vibrational spectroscopy.

Stockmayer Clusters (*J. Phys. Chem. B*, **109**, 23109, 2005.)

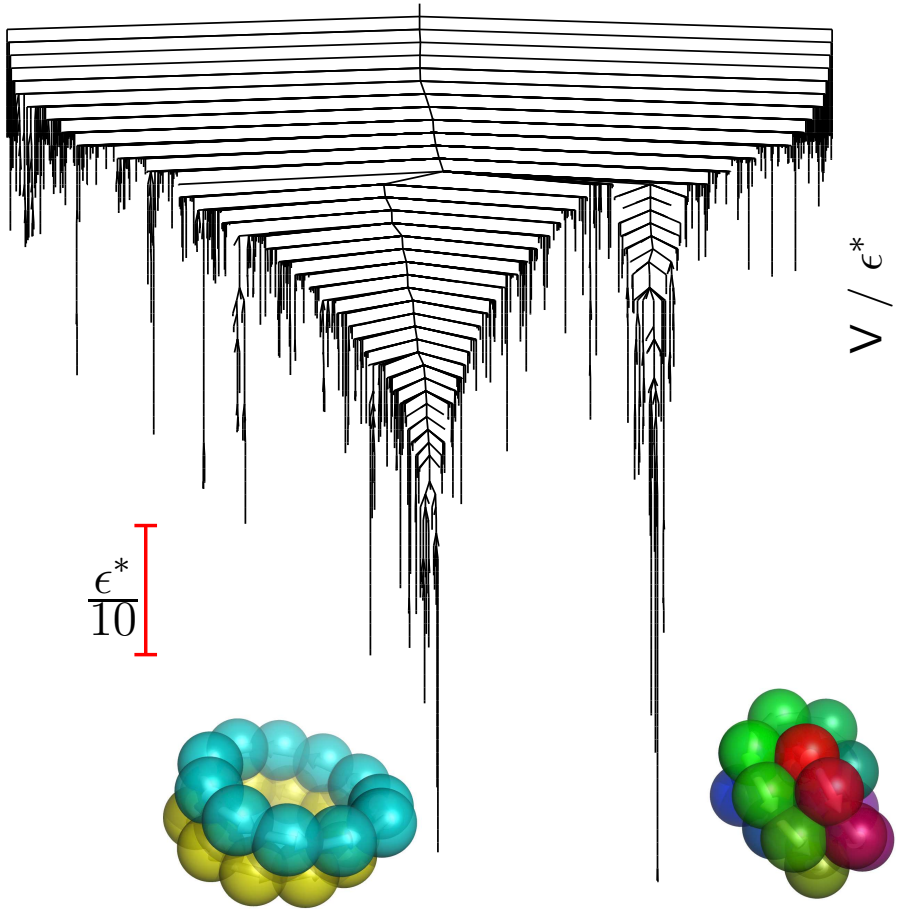


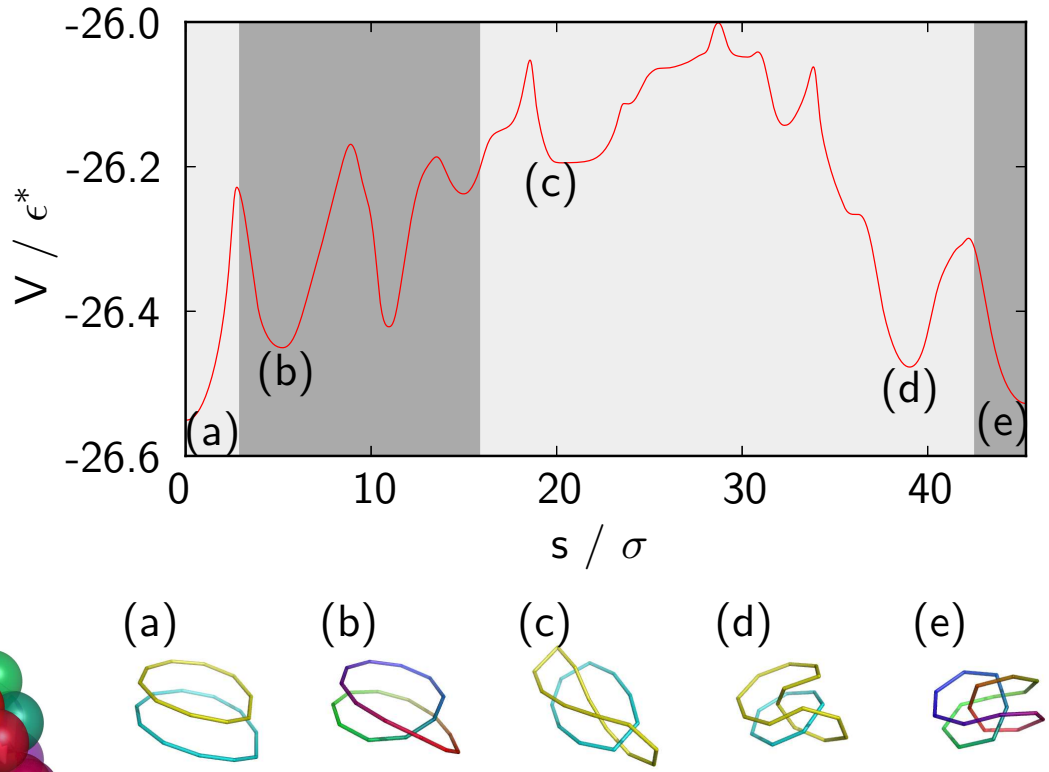


Left: structural map for Stockmayer global minima. Key: relaxed Lennard-Jones (black filled circle), decahedral (cyan filled square), linear (blue cross), ring (red open circle), two stacked rings (purple triangle), coil (brown diamond), link (orange open square), knot (green plus).

Right: global minima for selected (N, μ) values. Trefoil knot 3_1, knot 5_1, knot 8_{19}, knot 10_{124}, knot 10_{139}, coil with knot topology.

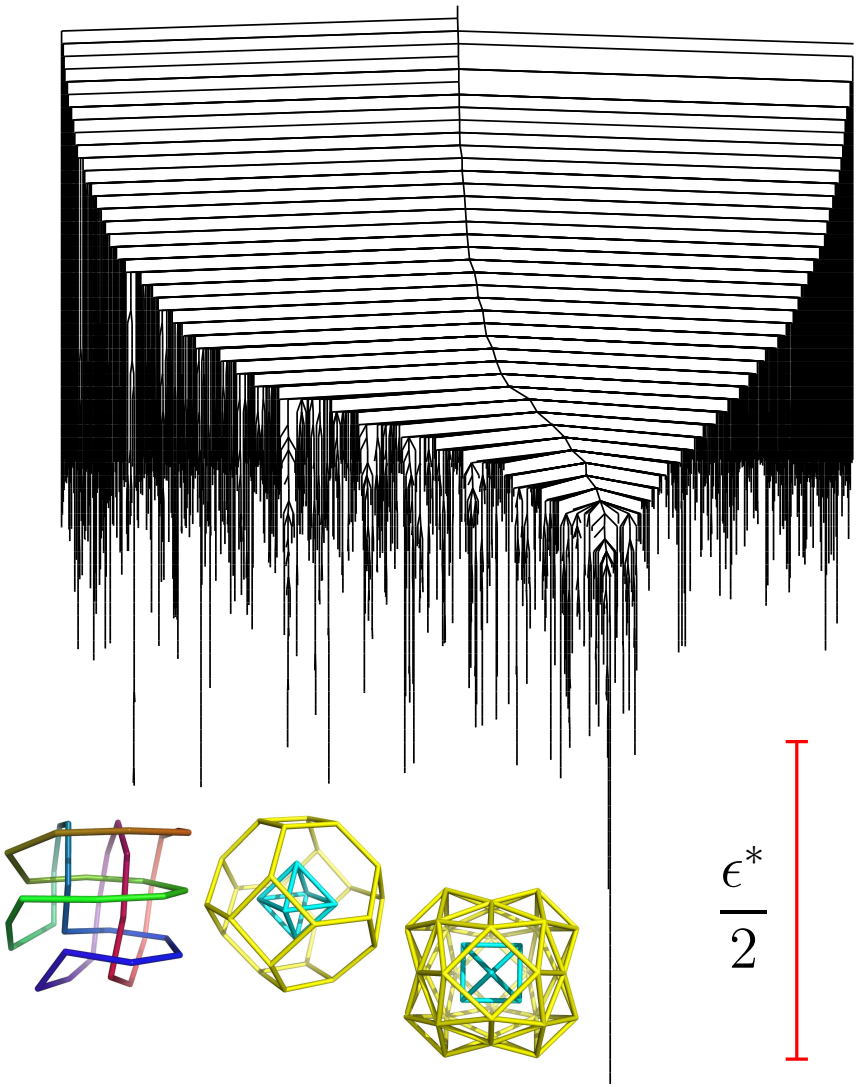
Pathways to Knotted Stockmayer Clusters (*Soft Matter*, **9**, 5407, 2013.)





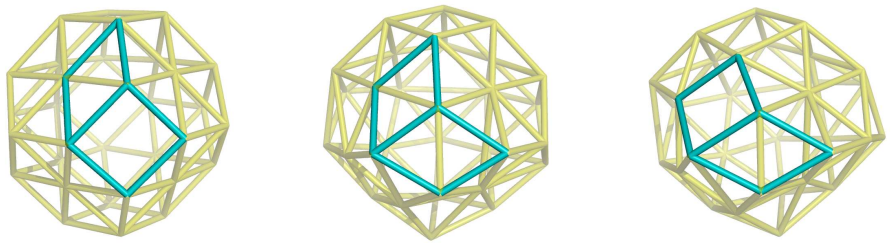
(Left) disconnectivity graph for St_{21} at $\mu = 2.9$.

(Right) pathway connecting the stacked ring and trefoil knot minima, containing 11 transition states. The unlink first rearranges to a coil, which contracts and then twists.

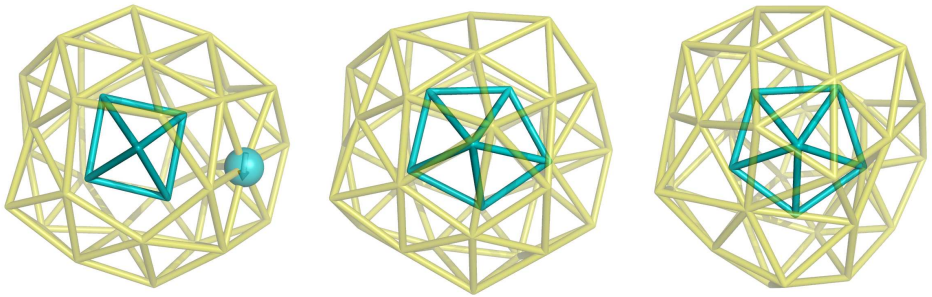


$\frac{\epsilon^*}{2}$

(a)



(b)



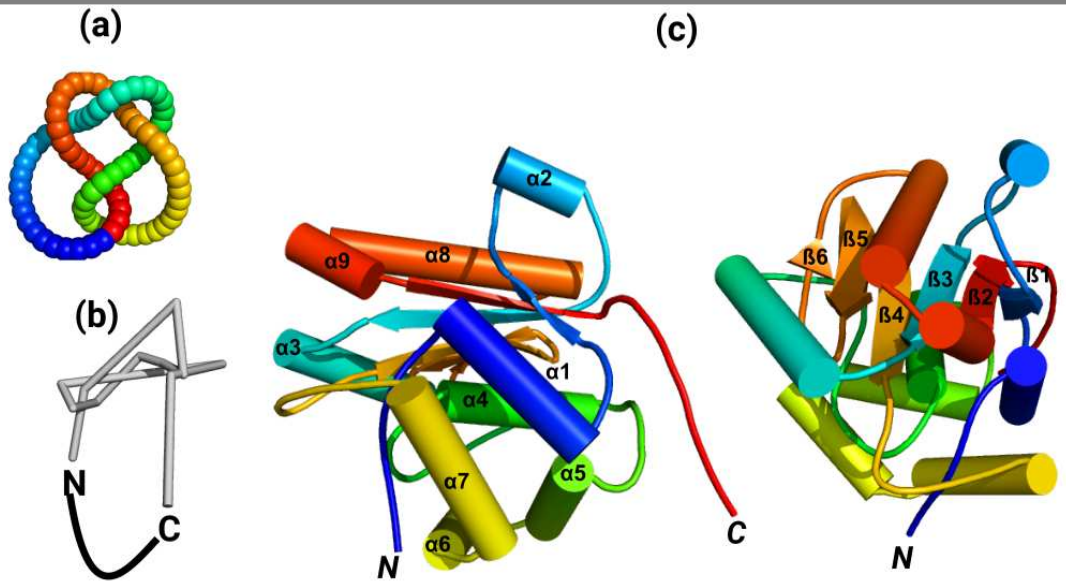
(Left) disconnectivity graph for St_{38} at $\mu = 1.22$. The global minimum has D_{4h} symmetry and the two other low-lying structures shown are O_h and a 8_{19} knot.

(Right) two of the three steps in the pathway from the O_h to the 8_{19} minimum. (a) A square-diamond–diamond-square rearrangement on the surface of the truncated octahedron. (b) Core-shell rearrangement in the final step.

Cutting a Gordian Knot

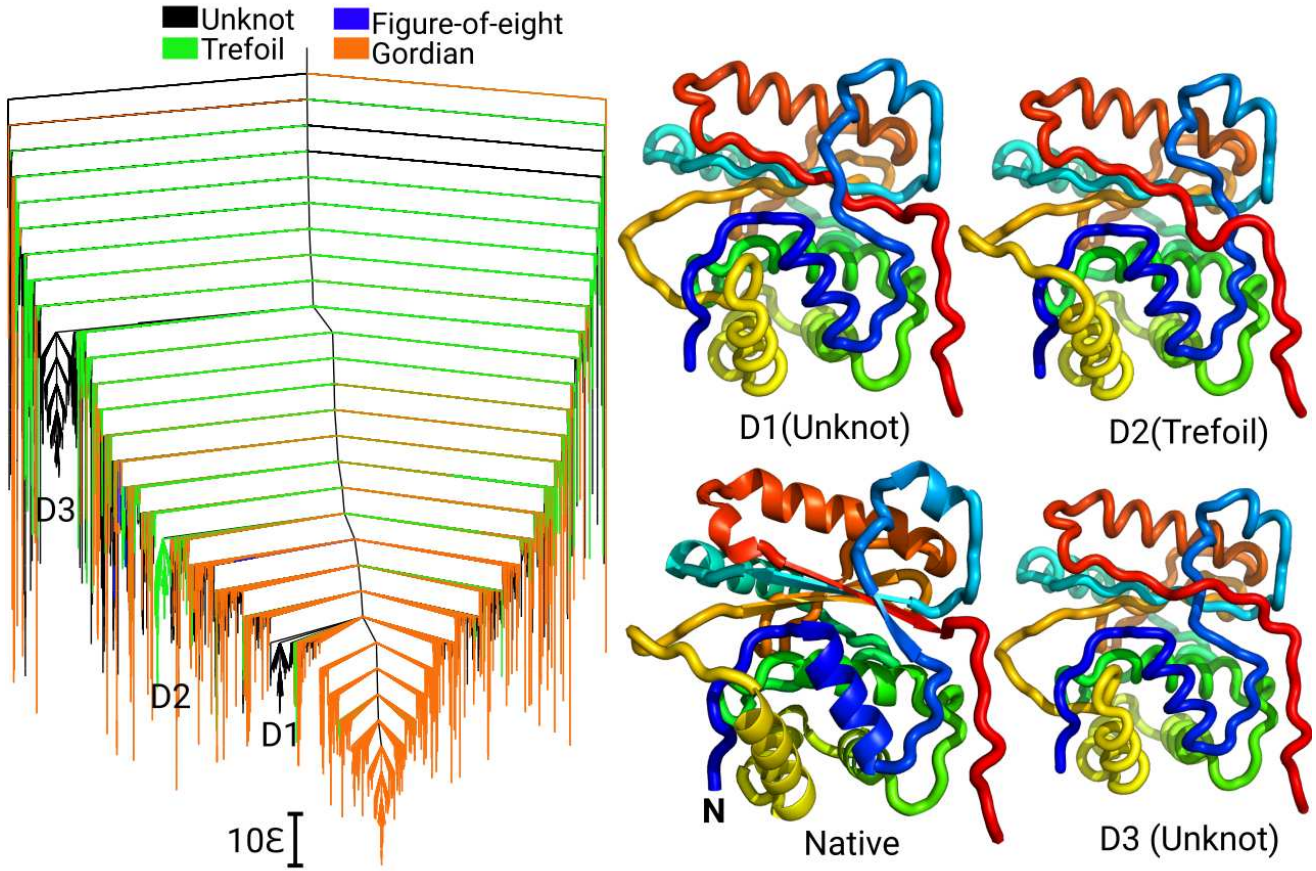
Ubiquitin C-terminal hydrolase isoenzyme L1 is a deubiquitinating enzyme linked to Parkinson's and Alzheimer's. The N- and C-termini have 5 and 10 residues, while 216 residues form a shallow Gordian knot.

The initial path required quasi-continuous interpolation, penalising internal minima for distances between discrete images (*JCTC*, **8**, 5020, 2012).



(a) The ideal 5_2 or Gordian knot with a closed loop coloured in the N- to C-direction from blue to red. (b) Sketch of backbone and loop-closure.

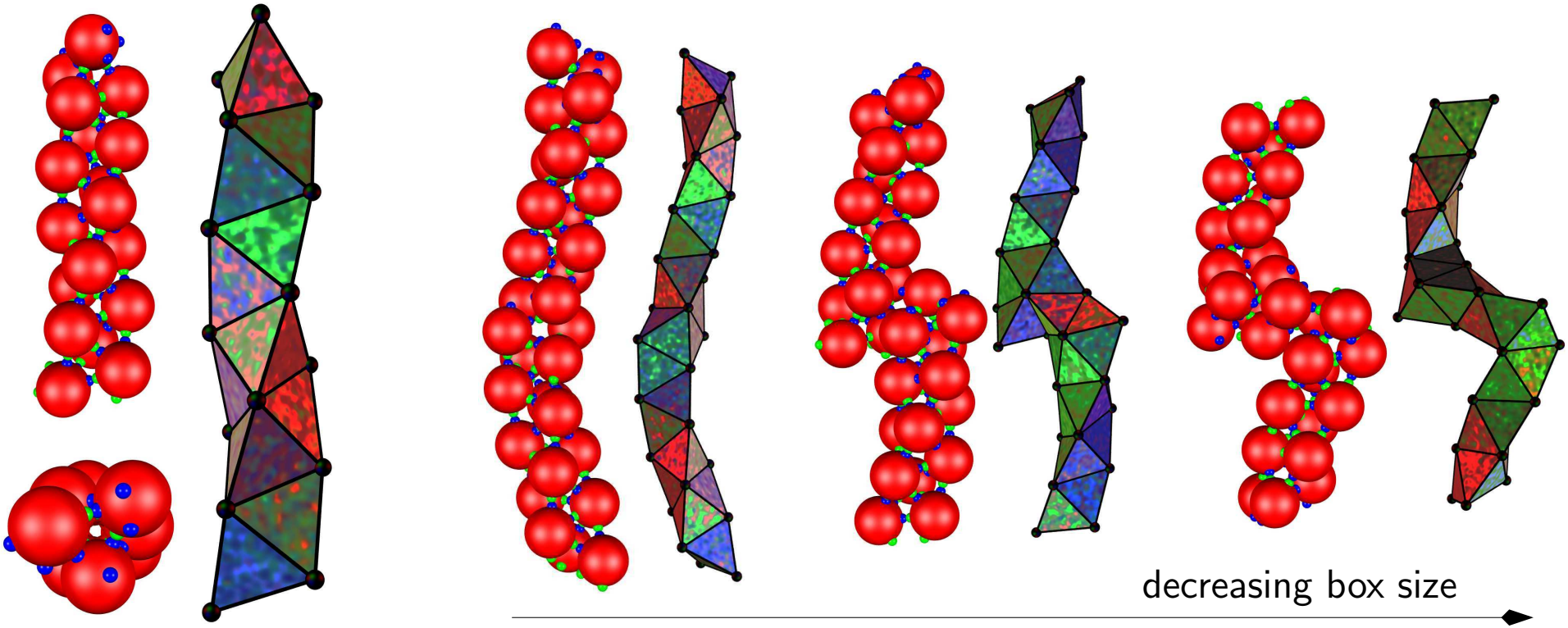
(c) Secondary structure of UCH-L1 with nine α-helices and six β-sheets.



The landscape features Gordian, figure-of-eight, and trefoil knot configurations, with regions corresponding to a jammed trefoil knot (D2) and to unknotted structures (D1 and D3).

Several distinct classes of folding pathway exist, including either trefoil or figure-of-eight intermediates, or a direct conversion from the unknot.

Designing a Bernal Spiral (*ACS Nano*, **7**, 1246, 2013)

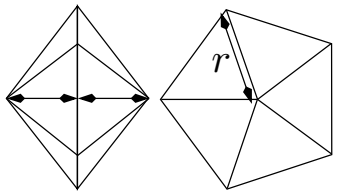


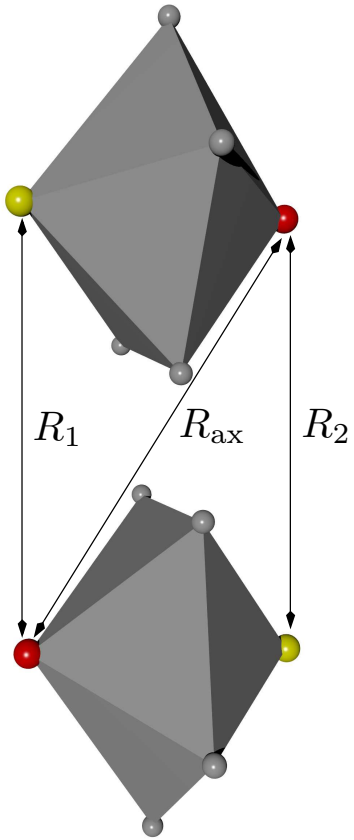
The simplest building blocks that support a Bernal spiral as the global minimum involve a single patch-antipatch pair offset by about 10° from linearity.

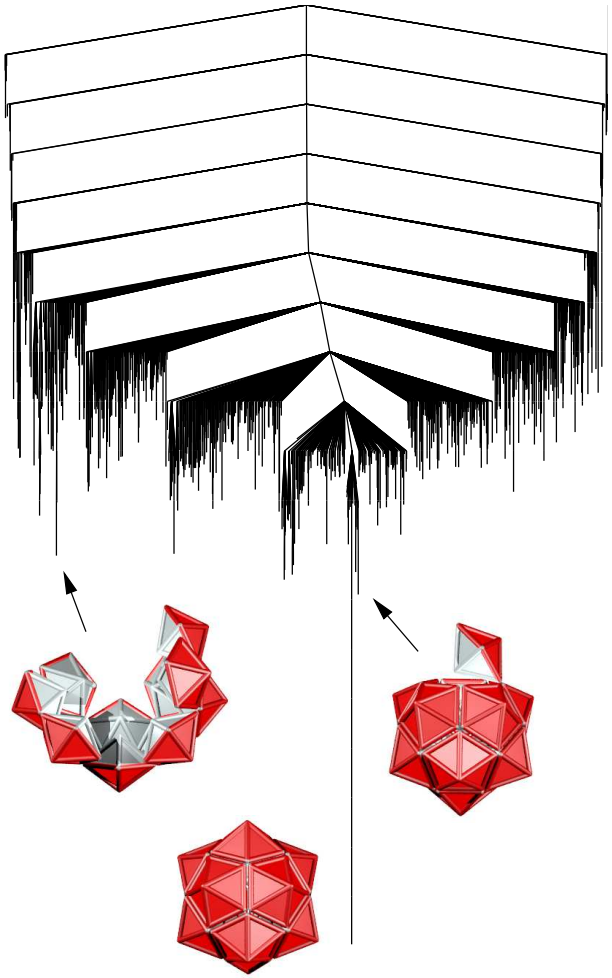
Left: Alternative views of a chiral Bernal spiral consisting of 18 particles.

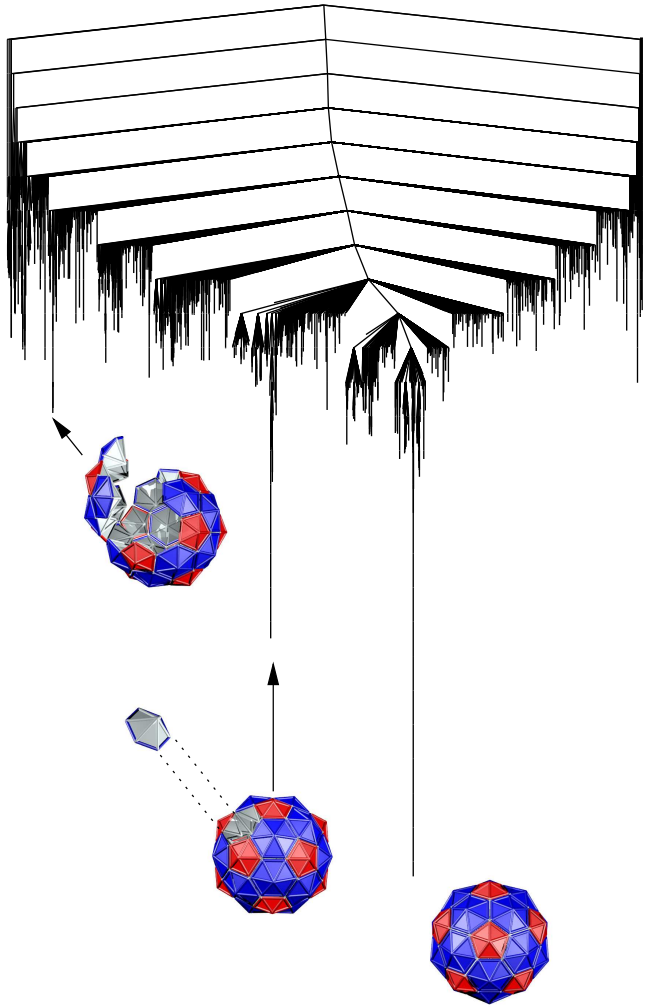
Right: compressed spirals (30 particles, periodic boundaries) exhibit supercoiling or breaks, resembling structures seen in confocal microscopy.

Self-Assembly of Icosahedral Shells (*PCCP*, **11**, 2098-2104, 2009)





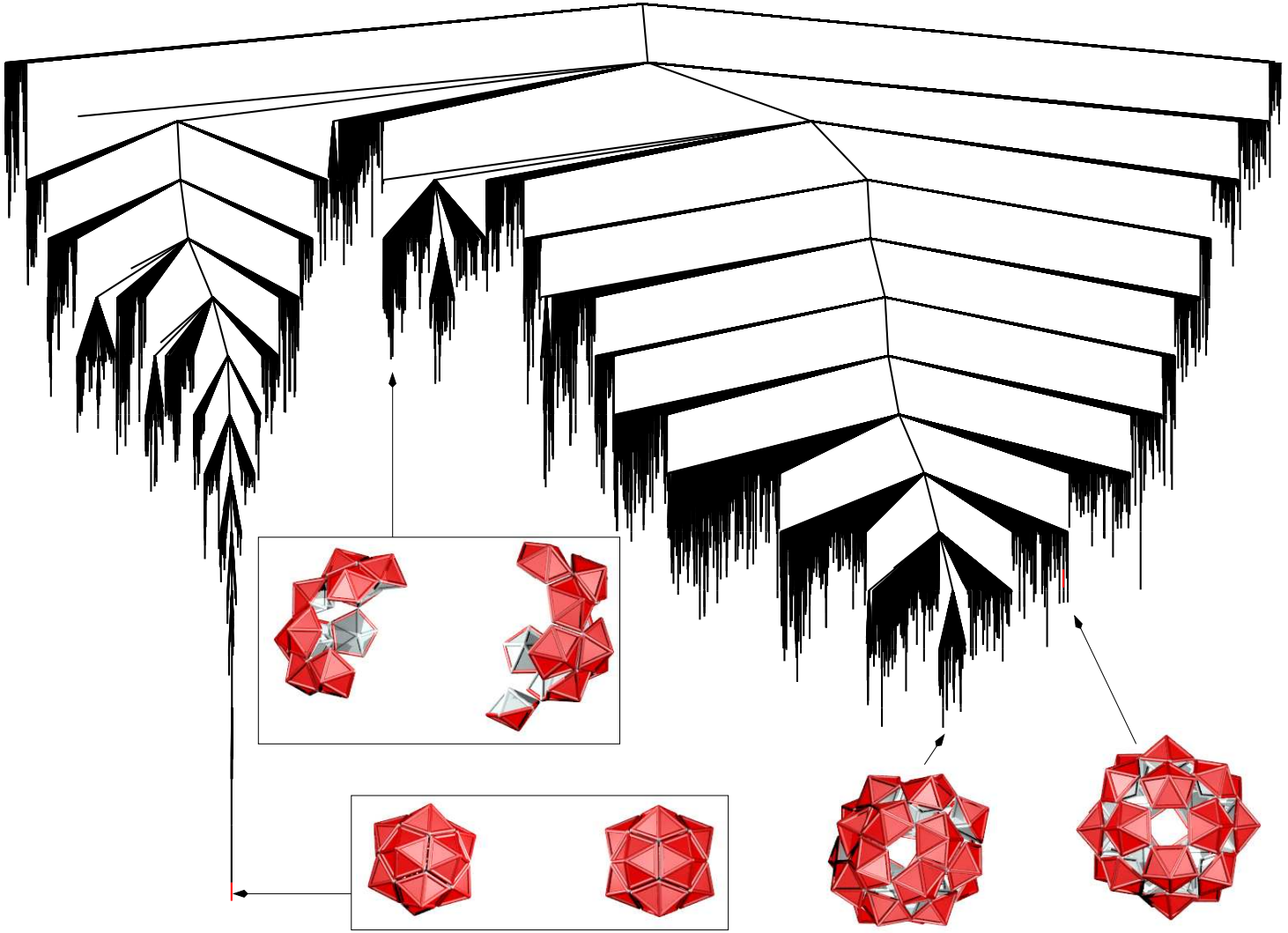




Palm tree disconnectivity graphs with I_h global minima are found for $T = 1$ and $T = 3$ shells constructed from pentagonal and hexagonal pyramids.

These landscapes correspond to efficient self-organisation.

24 Pentagonal Pyramids



For the same parameters two $T = 1$ icosahedra are similar in energy to a single shell based on a snub cube. *Polyoma virus* capsid protein VP_1 forms a left-handed snub cube from alkaline solution in the absence of the genome.