Protein design

[Kuhlman, B., Dantas, G., Ireton, G.C., Varani, G., Stoddard, B.L. and Baker, D. (2003) Design of a novel globular protein fold with atomic-level accuracy. Science, 302, 1364-1368]

What about folds that are **not** seen in SCOP or CATH?

Some are:

- physically impossible;
- not yet sampled by evolution;
- not observed by a structural biologist.

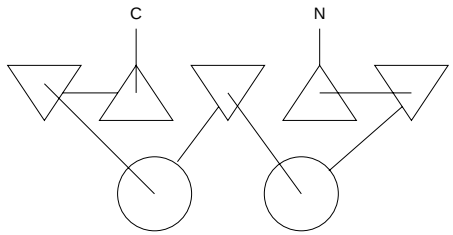
Goal was to achieve a highly stable protein with a new fold.

Approach to designing Top7 sequence

```
for i = 1 to 172 {
    generate starting structure;
    for j = 1 to 5 {
        for k = 1 to 15 {
            optimise sequence for fixed backbone;
            optimise backbone coordinates for fixed sequence;
        }
    }
}
```

Starting models are generated using a *de novo* approach ("Rosetta").
Assemble fragments taken from known structures.
Scoring function includes distance constraints from 2-D diagram.

TOPS cartoon of Top7



Optimise sequence

polar amino acid at the 22 surface β-sheet positions
(=> 75 rotamers per position)

any amino acid (except Cys) at the other 71 positions
(=> 110 rotamers per position)

Find combination of rotamers (and hence the sequence) with the lowest energy, using Monte Carlo search.

Optimise structure (1)

Measure energy

(i) Perturb backbone

a) choose between 1 and 5 residues at random and make small random adjustments to their main-chain torsion angles (ϕ, ψ),

or

b) replace the backbone of 1, 2 or 3 consecutive residues with a randomly selected fragment from the PDB, and adjust torsions of neighbouring residues to minimise the displacement of the downstream part of the chain.

Optimise structure (2)

(ii) Optimise side-chain structure

for those positions with higher energy after (i), replace current side-chain conformation with lowest energy rotamer.

(iii) Optimise backbone structure

optimise ϕ and ψ again in a 10-residue window around the perturbation site.

Measure energy again, and use Metropolis criterion to decide whether to accept or reject.

Steps (i), (ii) and (iii) are repeated several thousand times.

After every 20 such moves, a full combinatorial optimisation of side-chain rotamer conformations was carried out.