nature structural biology

december 1997
vol. 4 no. 12

String and sealing wax

"I think it's still here." I followed David Davies down to the end of the basement corridor, Building 5, National Institutes of Health, and there, perched on a rather rickety trolley, was a dusty metal model of γ-chymotrypsin, suspended in a square metal frame. Faded labels hung from bits of cotton tied to residues scattered about the structure, apparently at random. Other bits of cotton had lost their labels. David frowned at the model. "Let me see, now…" Carefully, he inserted his hand into the nest of wire and proclaimed; "…ah, here's the active site!".

Modelling is — and always has been — very much at the core of structural biology. But the models, and the way they are generated, have changed somewhat over the years.

The double helix

"Jim was always clumsy with his hands. One had only to see him peel an orange…" These opening sentences to the book, tentatively entitled *The Loose Screw*, that would have recorded Francis Crick's memory of the building of the most famous model in structural biology, has, sadly, never been written[1]. Thus, we turn to Watson's account of that electrifying moment in the Cavendish Physics Laboratory, Cambridge, as reported in *The Double Helix*[2]:

"When I got to our still empty office the following morning, I quickly cleared away the papers from my desk top so that I would have a large flat surface on which to form pairs of bases held together by hydrogen bonds. Though I initially went back to my like-with-like prejudices, I saw all too well that they led nowhere. When Jerry [Donohue] came in I looked up, saw that it was not Francis [Crick], and began shifting the bases in and out of various other pairing possibilities. Suddenly I became aware that an adenine-thymine base pair held together by two hydrogen bonds was identical in shape to a guanine-cytosine base pair held together by at least two hydrogen bonds. All the hydrogen bonds seemed to form naturally; no fudging was required to make the two types of base pairs identical in shape."[2]

This final, critical step in the discovery of the structure of DNA, was achieved with cutouts of the *keto* forms of the four bases which Watson had crafted from cardboard the previous afternoon, in March of 1953: the metal bases they had ordered from the machine shop at the Cavendish "for systematically checking all the conceivable hydrogen-bonding possibilities"[2] had yet to arrive. They were only completed several days later — with some egging on from Watson to finish the final soldering — and immediately incorporated into the first complete model of the double helix[2].

The simplicity of those cutouts, and the solution to the problem of Chargaff's rules and how to fit the bases between the deoxyribophosphate backbone seems so improbably elegant that, even now, it is hard to believe. In Watson's portrayal of the discovery of the structure, the two men appear almost as a pair of bumbling magicians: the reader gasps as they stumble, and then is amazed when they pull the structure out of the air. As Aaron Klug noted, rather more dryly, "there was no inexorable logic on the part of any of the protagonists leading directly to the solution."[3]

The α-helix

The double helix had its roots in another model, of the α-helical polypeptide chain. The critical features of the α-helix had been realized some five years before, in 1948, by Linus Pauling (Fig. 1).
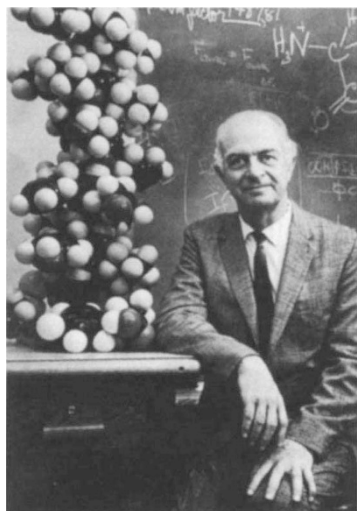
Fig. 1 Linus Pauling poses beside a model of the α-helix. In his attempt to solve the structure of DNA, in late 1952, Pauling modelled the molecule as a triple helix with the phosphate backbone on the inside and the bases on the outside[13]. To accomodate the hydrogen bonds between phosphates in the interior of the helix, Pauling assumed that they were not ionized. Watson was delighted by this blunder[2], but was petrified that Pauling, as soon as he realized his mistake, would try again.

editorial
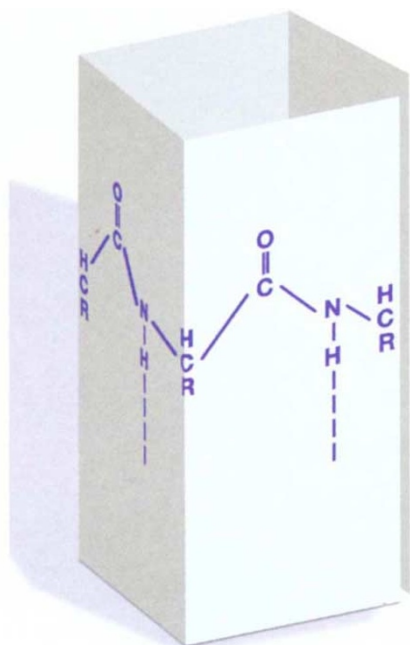
Fig. 2 Pauling's pencil and paper α-helix. Olby[4] summarizes Pauling' method as follows. "(i) Symmetry principals demand that the polypeptide chain when folded must assume a helical conformation. (ii) All peptide residues are assumed to be equivalent. (iii) The peptide C–N bond is planar, therefore there can be no rotation at either end of this bond. (iv) Rotation is achieved by the dihedral angle NαCC at the Cα atom. (v) The most stable, and therefore most probable, configuration is one in which hydrogen bonding takes place between residues along the chain. (vi) Satisfactory hydrogen bonding is achieved by adjusting the pitch of the helix." Figure adapted from ref. 4.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fig. 3 John Kendrew and the 'wire forest' model of myoglobin. The model was built in the Mathematics lab in Cambridge. The scale of the model was an impressive 5 cm to 1 Å. Kindly provided by David Davies.

Pauling, who at that time was the George Eastman Visiting Professor at Oxford, had been puzzling over the structure of the polypeptide chain in α-keratin for at least 10 years, but to little effect. Then, ill in bed with the 'flu, and bored of reading detective novels, he had modelled the structure of the α-helix using no more than a pencil, ruler and a sheet of paper (Fig. 2)[4].

Of the several criteria he used to derive the model, perhaps the critical point was the appreciation that the peptide bond was planar and that rotation could only occur about the alpha-carbon atoms of the peptide chain.

"Pauling simply drew the polypeptide chain across a sheet of paper, with the peptide bond in the plane of the paper and the Cα atom rotated so as to bring all the carbonyl groups on the same side of the chain. He then drew parallel lines through the Cα atoms at an arbitary angle to the chain and folded the paper along these lines through the dihedral angle (109°). This operation twisted the chain into a helix. It was then a matter of trial and error to find the orientation of the fold which brought the carbonyl and amino groups into line for acceptable hydrogen bonding."[4]

Although the early X-ray data on the structure of haemoglobin that Max Perutz had shared with Pauling were suggestive of the α-helix it was some time before Pauling was sufficiently confident of the model to publish it. Specifically, the 5.4 Å repeat predicted by Pauling's model was just not consistent with Astbury's measurements on α-keratin, which had suggested there should be a rise per turn of helix of 5.1 Å. Furthermore, the α-helix did not have an integral number of residues per turn, and yet the simple assumption by many at the time would be that they would[4].

It was work on synthetic polypeptides by scientists at the Courtauld's laboratory, in Maidenhead, Berkshire, UK, that would provide evidence of the parallel arrangement of the intramolecular hydrogen bonds, the helical arrangement of the polypeptide chain and the ~5.5 Å repeat that gave Pauling, Corey and Branson the courage to publish[4,5].

Pauling's was not the first helical model for the polypeptide backbone. Maurice Huggins had modelled a 3_{10}-helix as early as 1943, but had not incorporated the critical planar peptide bond[4,6]. This latter model had inspired efforts by Lawrence Bragg, John Kendrew and Max Perutz, at Cambridge, to tackle the problem. And they essentially solved it, too, in 1950, without realizing they had done so, only to reject the model because it, too, did not have the 5.1 Å repeat that Astbury's data suggested it should have.

Crick later provided the answer to the source of the 5.1 Å repeat which had so deceived first Pauling and then, at a critical juncture, the Cambridge group[7,8]. The biological structures that Astbury had been studying were coiled coils — α-helices that wrapped around one another to form a superhelix. The rise per helical turn was reduced because of this unusual quaternary arrangement.

As soon as Perutz saw Pauling's papers, he realized there was a simple test of the model. Perutz immediately set to work collecting X-ray data from a horse hair and a porcupine quill and came up with a crucial piece of supporting evidence for Pauling's α-helix, cementing the model's place in history[4].

Myoglobin

The α-helix and the double helix were models, in the sense that Pauling, Watson and Crick had interpolated the structures from a relatively limited set of data, which did not specifically define the positions of particular chemical groups within either of the helices. The determination of the 2 Å resolution structure of myoglobin, in 1959, also involved the interpretation of the data in terms of a model, but now the positions of the Cα backbone of the polypeptide chain and the amino acid side chains could be derived experimentally, once the phase had been determined, rather than being modelled, in the sense of the work on the α-helix and the double helix.

Modelling, in terms of generating a three-dimensional structure from the experimental data, would become more of a technical (although still non-trivial) skill, rather than the focus of the intellectual effort. Indeed, once the phase problem had been solved the structure solution would become essentially predictable, that is, deterministic (although not without its own problems).

The 2 Å electron density map for myoglobin was calculated in July, 1959, mainly by Kendrew, Richard Dickerson and Bror Strandberg. Paper tapes bearing the data and programs needed to analyze the data were taken from what Perutz fondly referred to as his 'Institute' (a prefabricated hut in the Cavendish known as 'the Hut') to the Mathematics lab, where they were fed into what was then probably the fastest electronic computer in the

Fig. 4 a, The original "Folly" outside Davids Phillips' office in the old Zoology building circa 1967. The model shows the structure of ribonuclease-S just then completed at Yale. The brightly lit 'real' model is shown below and above is its reflected image in the large half silvered mirror. Some sections of the stacked map can be seen behind the mirror. The vertically adjustable lights, used to illuminate a thick section of the model for actual use in fitting, can be seen on either side of the model. The vertical and horizontal calibrated bars above and beside the model were used to measure the *x, y, z,* coordinates of the atoms after completion of the model. **b,** Fred's folly in action. Stereo diagram showing the fitting of a model of ribonuclease-S to a 3 Å electron density map[11]. Reproduced with permission[11]. **c,** The amphitrite illusion. The viewer (left) believes he sees a woman (lower centre) rise from the sea and, without any visible means of support (image, upper right) revolve around and around while gracefully moving her arms and legs. The basis of the illusion is the half-silvered mirror (M–M'). Reproduced from Hopkins, A.A. *Magic: Stage Illusions and Scientific Diversions* (Benjamin Blom, Inc., New York; 1897) Images kindly provided by Fred Richards..

IMAGE
UNAVAILABLE FOR
COPYRIGHT REASONS

world, EDSACII (EDSACI being the first electronic computer in the world, but very much slower)[9]. Michael Rossmann, working with Perutz, was one of the people who developed the software needed to solve such a vast molecular structure. Dickerson recalls that Rossmann's "least squares and rotation translation programs became so complex that the Computing Centre developed the practice of using one of them for morning computer checks"[9].

The painstaking building of the model for myoglobin was started by Kendrew and Strandberg, with help from Davies, in the late summer of 1959. In the tradition of the small molecule X-ray crystallographers, the density was represented in horizontal sections on stacks of Plexiglass sheets. Here at last was the first experimental visualization of Pauling's α-helix, and the first 'high resolution' structure of a macromolecule. All the helices were right-handed, as theory had suggested they might be; right-handed helices being slightly more stable than left-handed helices. Dickerson recalls[9] that "[we] threw a cocktail party at dusk on the Peterhouse lawn to celebrate… Lawrence Bragg [was] taking the elbow of guests … and propelling them to the light box, pointing at an α-helix that ran obliquely through the map sections, and saying excitedly: "Look! See, it's hollow!"'

The lack of atomic resolution and the shear size of the protein suggested that stacks of Plexiglass sheets would be of limited use for obtaining the coordinates in three dimensions for the structure; a physical model was needed. Wooden boards, a drill press and ~2,500 rods, each 2 m in length, were ordered by Kendrew. Strandberg and Davies then set to work. Holes were drilled into the boards, at the positions of the grid points covering the unit cell, and the rods inserted, representing the vertical dimension of the volume. Coloured clips were attached to the rods to indicate regions of high, medium and low electron density. The brass-wire model was then built directly into the Fourier map by following the clips around the forest of rods (Fig. 3)[9,10].

Myoglobin was the only model built using such a 'wire forest'. Clearly, such a method was impractical for either looking at the protein (which was buried in the rods), or adjusting it (without being impaled). Nor was the later practice of transferring measurements from the stack of Lucite sheets to a 'rod-free' model (Fig. 5) much of an improvement: the transfer of information was error-prone, time-consuming and tedious.

"We used glass windows, rather than Plexiglass, which was expensive and difficult to come by," Rossmann recalls. "A structure could be represented on 50 sheets of glass. Each sheet had to be hand drawn (using India ink) and took two to three people an hour each. We employed people to do the drawing, including one of my daughters

editorial

Fig. 5 David Davies (left) and E. Padlan (right) view a model of the phosphocholine-binding FAB, McPC603. The scale of the model is 2 cm to 1 Å (the so-called Kendrew scale). Figure kindly provided by David Davies.

during her summer vacation. In the end we had stacks and stacks of these sheets; enough to build a greenhouse."

It was clear that the ability to superimpose the map of the electron density on the model of the protein provided the solution to the problem; but how could that be done with sheets of glass and brass wire? Impossible, unless magic was involved: as it turns out, it was.

The optical comparator

Fred Richards provided the solution. The optical comparator (also known as 'Fred's folly'), developed in 1968, employed a half-silvered (two-way) mirror that allowed the builder to see the model "suspended within the contour map so that the "fit" can be evaluated directly" (Fig. 4)[11]. Up to 8–12 Plexiglass sections of the map were viewed at any one time; the model was illuminated such that only that part which corresponded to the map was seen in the mirror (Fig. 4b).

Although Richards had drawn the idea for the comparator from telescope mirrors and microbalances, the real origin of the device came from a much earlier time, as such devices had commonly been used as part of 'smoke-and-mirrors' magic shows at the end of the previous century (Fig. 4c). Rossmann notes that "the Box was very good, and certainly a great improvement over the sheets and separately built model." But their history was, alas, short, as Richards notes:

"A number of these "instruments" were built by the number of increasingly successful protein structure laboratories around the world. This was a bit like a very small-scale version of the clipper ship era. In spite of the enormous literature describing the ships, that famous period lasted only about 10 years, after which the vessels powered by steam engines took over. In the same fashion the follies disappeared overnight when the computer graphics revolution got to the stage of ease of use and commercial availability"[12]

Frodo

Although it would be fair to say that the optical comparator was baroque in construction, it should also be noted that the computers of the early 1970's were not the convenient little plastic boxes we are familiar with today. "The first *mini*-computer, the PDP11," Alwyn Jones point out, "weighed in at half a ton. And the main-frame [the Siemens 4004 at the Max Planck Institute for Biochemistry, Munich] was the biggest computer I had ever seen, being roughly the size of a tennis court". Even so, they were more convenient and easy to use than the follies, although only marginally quicker and more accurate. Furthermore, the stacks of sheets, and the model itself, were labile to sag and bend under gravity and other physical abuses.

During this period, a number of attempts were made to make an electronic version of the optical comparator; some used 'home made' hardware, some used commercial systems that were just beginning to be produced. It was Alwyn Jones who came up with the computer graphics 'killer ap' for structural modelling, though. Having stopped in Munich while on his way to Chicago to start a postdoctoral fellowship in Paul Sigler's laboratory, he ended up staying on at the Max Planck Institute for Biochemistry to work in Johan Gassmann's computer centre (Sigler, it is rumored, has never forgiven him). Jones remembers that "they had just taken delivery of the first Vector General 3400 computer graphics system, and needed some programs to run on it".

It was here that Jones, left in peace and quiet and relative solitude, wrote Frodo. The program allowed the crystallographer to rip the molecule apart and to fit it into the density drawn on the computer screen, much like bending the metal bonds of the first wire models. In particular, it made the crystallographic refinement of macromolecular structures much easier. Jones further assured Frodo's success by adapting the program to each new computer platform that came along. Who, then, needed to use physical models anymore?

Postscript

They're still around, though. Most labs, like David Davies', have one tucked away somewhere, gathering dust and occasionally being wheeled out for teaching purposes. And the smaller ones make great structural biologist executive toys. It is still possible to buy modelling kits, though, or even submit coordinates to the telemanufacturing facility of the San Diego Supercomputer Center where a three-dimensional surface representation of the structure is sculpted by laser from blocks of laminated paper. But there is something quite grand about those old brass wire models. Perhaps because the information they conveyed was all so new.

1. Crick, F.H.C. *Nature* **248**, 766–771 (1974).
2. Watson, J.D. *The Double Helix* (First edition, 1968; Norton Critical Edition, New York; 1980).
3. Klug, A. *Nature* **219**, 808–810 (1968).
4. Olby, R. *The Path to the Double Helix* (University of Washington Press, Seattle; 1974).
5. Pauling, L, Corey R.B. & Branson, *Proc. Natl. Acad. Sci. USA* **37**, 205–211 (1951).
6. Huggins, M. *Chem. Rev.* **32**, 195–218 (1943).
7. Crick, F.H.C. *Acta Crystallogr.* **6**, 685–689 (1953).
8. Crick, F.H.C. *Acta Crystallogr.* **6**, 689–697 (1953).
9. Dickerson, R. *Prot. Sci.* **1**, 182–186 (1992).
10. Kendrew, J.C. *et al. Nature* **185**, 422–427 (1960).
11. Richards, F.M. *J. Mol. Biol.* **37**, 225–230 (1968).
12. Richards, F.M. *Ann. Rev. Biophys. Biomol. Struct.* **26**, 1–25 (1997).
13. Pauling, L. & Corey, R.B. *Proc. Natl. Acad. Sci. USA* **39**, 84–97 (1953).