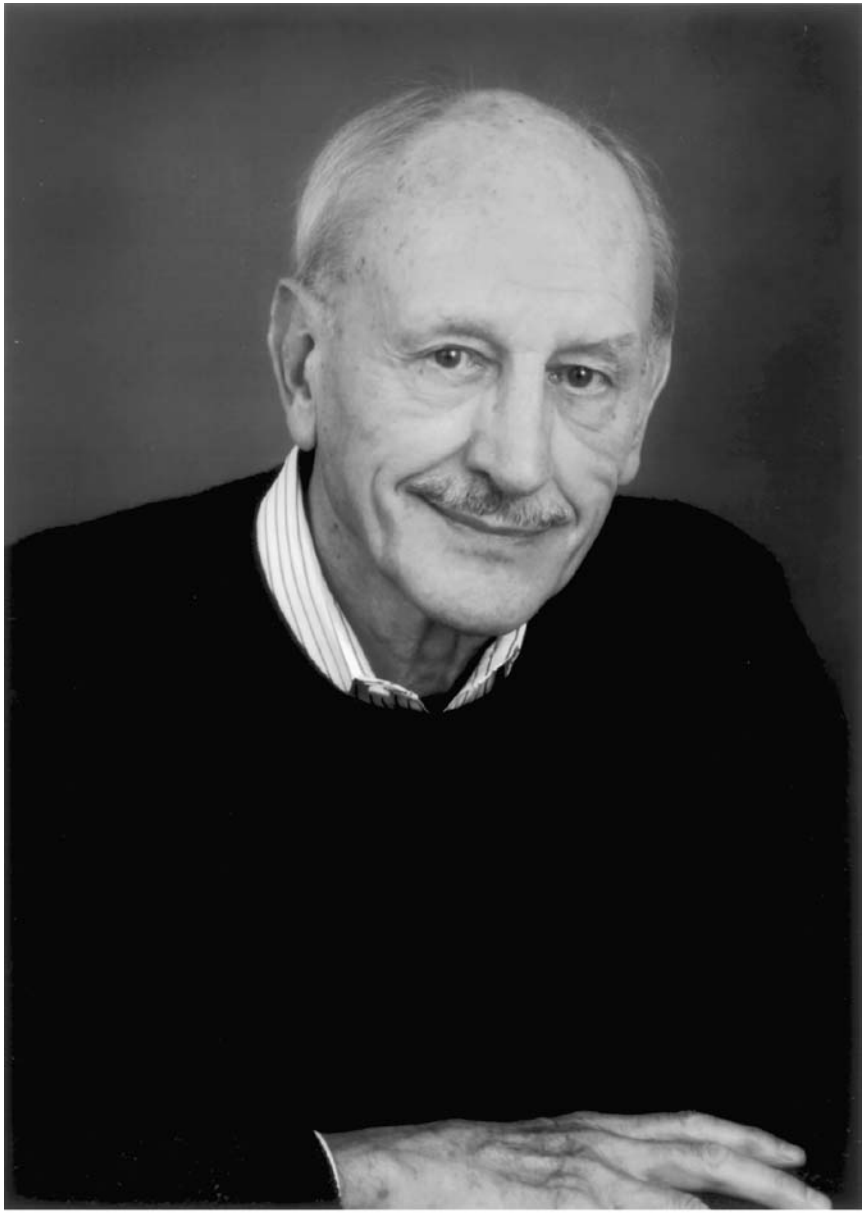

David Daiches

Annu. Rev. Biophys. Biomol. Struct. 2005. 34:1–20
doi: 10.1146/annurev.biophys.34.040204.144531

First published online as a Review in Advance on November 11, 2004

A QUIET LIFE WITH PROTEINS*

David Davies
Laboratory Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892; email: david.davies@nih.gov

Key Words crystallography, polynucleotides, structure, function

CONTENTS

PROLOGUE 1
INTRODUCTION 2
OXFORD 2
CALTECH 4
NIH: FIBER DIFFRACTION 6
Protein Crystals 9
The Laboratory of Molecular Biology 11
ANTIBODIES 12
Fab COMPLEXES WITH LYSOZYME 13
ASPARTYL PROTEASES 14
HIV INTEGRASE 14
TRYPTOPHAN SYNTHASE: TUNNELS AND CHANNELING 15
SOME RECENT STRUCTURES 15
CONCLUSION 17

PROLOGUE

I am fortunate to have been present during the enormous explosion in knowledge and understanding that has occurred in the biological sciences over the last half century, in what may well be regarded as the golden age of molecular biology. The contrast between our current knowledge and what was known in the early 1950s is so vast that it defies imagination. However, I have always been slow to rush into print, and for many years I have resisted suggestions to write this kind of perspective while nevertheless flattered to be invited. Finally, flattery, plus the knowledge that many of those involved in the early days are no longer with us, has led me to put down on paper some of the areas and events that have excited me in the past and continue to hold my interest.

*The U.S. Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

INTRODUCTION

In Wales, on the border between the counties of Glamorgan and Carmarthen, I grew up in a village called Pontardulais. Pontardulais had two main industries, a coal mine and a tin plating factory, now both closed. It was a relatively small village that was surrounded by beautiful hilly countryside. My parents, who were Welsh Baptists, encouraged me to attend services three times on Sundays. Because the services were in what my mother called "deep Welsh" (i.e., no English words thrown in), I had plenty of opportunity for theological contemplation and at the age of fourteen I abandoned all religious practice, although I hope the ethical and moral precepts stayed with me. I attended high school at the local area high school in Gowerton, about five or six miles away. The students represented the 10% or so of the school population who had managed to pass the now abandoned "11 plus" examination, in which at the age of eleven the students were selected to go on to grammar school. Gowerton School served a fairly wide area of both rural and industrial villages and was regarded as an excellent rugby school. Nevertheless, many of the teachers were outstanding and when I left I was well prepared in chemistry, physics, and math. It was wartime and we endured rationing of food and clothing, which was hard on our mothers, but we escaped the worst by living on the edge of a substantial farming community.

OXFORD

No one in my family had been to college, but I was encouraged to apply to Oxford University and to everyone's astonishment I was accepted by Magdalen College. Only one student from our school had previously gone to Oxford that I knew of, John Maddox (now Sir John), who had a distinguished career in the University and went on to become the longtime editor of *Nature*. In the fall of 1945 I took the train from South Wales to this archetype of the English academic establishment, this beautiful college on the Cherwell where I was to study physics. Magdalen had a long tradition of taking students from the aristocracy, the most notable having been the Prince of Wales. The majority of the student body at Magdalen came from the English public schools, with the remainder from the state-supported grammar schools, so I soon had to modify my heavily accented English in order to be understood. Many were returning, some badly injured, from World War II, where some of the physicists had acquired expert knowledge in fields such as electronics and radar. One of my contemporaries at Magdalen was a chemist, Leslie Orgel, who suggested we read quantum mechanics together but explained that he was rather a slow reader at about 40 pages per hour. Within a few years he was making major theoretical contributions to transition metal chemistry.

When I arrived at Magdalen I was given a suite of rooms in the New Buildings (300 years old) overlooking the deer park. Shortly after my arrival I was scheduled to interview with my tutor, Pat Johnson, who had just returned from the army, and I

was looking forward to a stimulating intellectual conversation. His first words were "Fix yourself a drink. The gin's in the bathroom," thus introducing me to my first taste of this beverage. He then produced a list of lectures, saying, "I don't suppose you'll want to go to many of these." So I spent the next three and a half years in such undergraduate pursuits as rugby, rowing, and rock climbing, although careful not to neglect my weekly essay. At the end of this period I was given the choice of serving in the army or studying to receive my D. Phil. After due deliberation I decided to do the latter and entered the mentorship of H.M. Powell, Reader in Chemical Crystallography. This was a field in which I could perhaps use my background in physics and also return to some of my earlier interests in chemistry.

The crystallography laboratory was located in the Gothic building that housed the Natural History Museum on Parks Road. The staff consisted of H.M. Powell and Dorothy Hodgkin. Powell occupied a small office in the south side of the building on the ground floor and the X-ray equipment was housed at the other end. All the graduate students and the postdoctoral fellows worked in one large open room on the second floor, also at the other end of the building. The advantage of this was that we students could learn almost everything from the postdocs who were much more accessible than some of our supervisors. The adjacent room of about the same size was occupied by Dorothy Hodgkin and we saw much of her.

Dorothy had quite a few postdocs and graduate students, including David Sayre, Jack Dunitz, and Pauline Harrison. Most of Powell's graduate students had been physics undergraduates. Sayre, although a graduate student, had already had extensive experience with Ray Pepinsky in the United States, working on a computer that would sum Fourier series and calculate electron density maps. His wife, Anne, was a writer and together they impressed us as very sophisticated. Dunitz had been a postdoc at Caltech, which had made a huge impression on him, and he would relay this to us in his strong Scottish accent.

"Tiny" Powell, as he was known to everyone, was a small-molecule crystallographer who had discovered clathrate compounds, cage-like crystal structures that could trap and retain other molecules. Dorothy, however, had always chosen to work at the cutting edge on molecules of biological interest, such as penicillin, vitamin B12, and, always lurking in the background, the protein insulin. She was a good friend of Max Perutz, who occasionally would visit from Cambridge. As a result we were kept aware of developments in the area of biological structures. In 1951 Linus Pauling, Robert Corey, and Herman Branson proposed the α-helix followed by the β-sheet structures to explain the diffraction patterns of many fibrous proteins obtained principally by W.T. Astbury at the University of Leeds. These models were convincingly based on dimensions obtained from precise crystal structures of amino acids and peptides that had been determined at Caltech.

Shortly afterward there was a meeting in Cambridge to discuss the status of crystalline protein structure determination. Several of us drove from Oxford in David Sayre's old Austin 7 to what was probably the first protein crystallography meeting. All the British protein crystallographers were there: Lawrence Bragg, the Cavendish Professor of Physics, Astbury, J.D. Bernal, John Kendrew, Perutz,

and a new emerging star in this field, Francis Crick. To us, as small-molecule crystallographers, the aspirations of these Cambridge crystallographers were astonishing. There were no computers and we knew the effort required to determine the structures of small molecules, where even a 10-atom structure could take a year to solve. Perutz had seen rods in his Patterson function projection of hemoglobin and he proposed a five- or six-helix bundle structure for the protein. Crick, who had recently joined Perutz's group as his graduate student, then presented data showing, from an examination of the Patterson at 5.4 Å radius, that this structure could not be correct because parallel helices would have resulted in a large peak that was not observed. It was an exciting meeting: Not only were helices in the air, but there was increasing support for the idea that protein structures would one day be solved.

Meanwhile, in the small-molecule field Sayre made an important contribution by developing what became known as the Sayre Equation. This provided a relationship between the X-ray intensities and their phases that could lead directly to a crystal structure and was used by him to produce a convincing electron density map of hydroxyproline. This paper (26a) was accompanied by papers by Hauptman & Karle (12a), Cochran (2a), and Zachariasen (37) along similar lines. The writing was on the wall and it was clear that as computers developed, these direct methods would likely replace the traditional methods for solving organic crystal structures.

CALTECH

We assembled Fourier maps with the aid of Beevers and Lipson strips, which were thin paper strips containing values of sines and cosines for given amplitudes and frequencies sampled at appropriate intervals, that we assembled and then summed to give the electron density projection. After a few years of this I completed the requirements for my D. Phil. and, thanks to Dunitz's suggestion, I went to Caltech to postdoc with Pauling and Corey. This was a mind-boggling experience on many counts. When we arrived in Pasadena, after flying low into the Grand Canyon, we stayed at the Athenaeum, the Caltech Faculty Club. The first morning, I opened the blinds and was astonished to see the 6000-foot facade of Mount Wilson. Dunitz had never mentioned this proximity of mountains almost twice as high as Snowdon, the tallest mountain in Wales. Then the smog descended and it was two months before we saw Mount Wilson again. To us, coming from an England that was just emerging from WWII rationing, California was a land overflowing with milk and honey. Then there was the weather, where months could pass without rain and seasonal differences were small. The Caltech social atmosphere was open and friendly, and there was a degree of socializing between faculty and students that would have been unthinkable at Oxford. In addition to the ocean we were within striking distance of the desert and the mountains, and many weekends were spent exploring the Sierra Nevada mountains and Death Valley.

However, the political climate in California was chillingly conservative. Richard Nixon had been elected vice president to Eisenhower shortly after we arrived. He was universally disliked by the postdocs and graduate students for the way he had campaigned against Jerry Voorhis and Helen Gahagan Douglas. However, the most striking manifestation of the general support for a reflex anticommunism lay in the phenomenon of Senator Joe McCarthy, whose rise and fall occurred during our stay in California. Pauling, who was outspoken on political matters, particularly pertaining to the control of nuclear weapons, was himself a victim of this atmosphere and was at one point denied a passport to visit a Royal Society meeting in London.

Caltech, however, was a wonderful place and we were most impressed with the caliber of the graduate students and postdocs, most of whom went on to distinguished careers in scientific research. I was partnered with Joe Blum, a physiologist who had come to Caltech to learn crystallography so that he could eventually work on the structure of muscle. There was still at Caltech an active program in crystallography aimed at accurate structures of small peptide-like molecules to improve the precision of the database that had been assembled for the construction of the α-helix and the β-sheets. I was asked to join this program, and Joe and I determined the structure of an eight-atom structure called parabanic acid. We collected a set of films containing the three-dimensional X-ray data, and we each independently measured the intensities visually, by comparison with an intensity scale that we constructed. The Caltech Chemistry Department was at the time probably the best place anywhere to practice crystallography. There was an array of IBM punched card computers that we could use, and the procedures for using them had been well established. Each machine could be "programmed" by wiring a plugboard, and they ranged from a simple sorter to the IBM 604, a computer that had a thousand vacuum tubes and could carry out sixty simple operations for the passage of each card. When we calculated a final difference map for the parabanic acid crystal, we found unexpected peaks that indicated that the molecule had anisotropic thermal vibration. So with the advice of John Rollett, who had come from Leeds, Joe and I developed a least squares procedure for refining the individual atomic anisotropic thermal vibration parameters using the three-dimensional data. This removed the unwanted peaks in the difference map, gave a suitably low R factor, and provided a remarkably precise set of atomic dimensions. In my second year I repeated this procedure with succinamide and was even able to see what looked like bonding electrons. With modern computers this procedure is now routinely applied to every crystal structure determination, but at the time, doing it for the first time was exciting (6, 7).

In the fall of 1953 there was a meeting at Caltech that was, I believe, originally intended to gather support for the Pauling and Corey structures. The α-helix model predicted a strong X-ray reflection close to the meridian at a spacing of 5.4 Å, whereas Astbury's fiber patterns showed a 5.1 Å reflection and this presented a problem for the structure that had caused Pauling to delay publication for some time. However, in the year preceding the conference there had been

several extraordinary developments. First, Crick had provided a clever and plausible explanation for the difference between the observed and calculated spacings by proposing a coiled-coil structure that would result in a 5.1 Å reflection. Second, Max Perutz and his colleagues David Green and Vernon Ingram had just achieved the first successful isomorphous replacement in hemoglobin crystals using a mercury derivative that enabled them to calculate the phases for a projection of the molecule, thus demonstrating conclusively that the determination of protein structures was possible. Third, Watson and Crick had just determined the structure of DNA. These events combined to make this the most remarkable structural biology meeting I have ever attended and made even a committed small-molecule crystallographer appreciate the treasures that were to be discovered from a study of biological macromolecules.

However, it was not to be immediately and I returned to England to an industrial job in Oldbury, on the west of Birmingham on the edge of the "Black Country." I worked in the research department of a company that manufactured phosphorus fine chemicals. It was my first introduction to industrial England, where the canal was green and greasy and the fogs impenetrable. However, the research department had a small but effective X-ray laboratory and I was able to determine the crystal structure of sodium triphosphate.

NIH: FIBER DIFFRACTION

After one year in Oldbury I was ripe for other possibilities, and when I received a letter from Alex Rich inviting me to join him at the National Institutes of Health (NIH), I replied by cable that I was on my way. At Caltech we had known Alex and his wife, Jane, well and had traveled on many of our trips with them. Alex, an M.D., had been at Caltech as a member of the Public Health Service, recruited by Seymour Kety, the then Scientific Director of the Mental Health Institute, to work eventually in the about-to-be-opened Clinical Center of the NIH. Alex had also recruited Gary Felsenfeld, who had been at Caltech as Pauling's graduate student and was also a close friend of ours. He had just spent a year in Oxford with C.A. Coulson. So, with my small family I crossed the Atlantic in December to arrive in Bethesda only to find that Alex was still in Cambridge, where he and Crick were working on a structure for collagen. However, several months later he turned up and we began a research program directed at elucidating the structure of RNA.

The structure of DNA had been so successful and its impact so dramatic, it seemed to Alex that a structure for RNA might have a comparable impact. So we set about organizing the lab to investigate RNA by solution and fiber diffraction methods. This was my introduction to the structure of biological macromolecules and it was quite exciting. The first thing I did was develop a computer program for calculating helical transforms. There had been a seminal paper by Cochran, Crick, and Vand (2b) in which they derived the mathematical formulation for

helical diffraction patterns. So I developed this algorithm and wired the boards for the punched card machines we had available at NIH, which were used mostly at that time to calculate the payroll. Alex was able to obtain some homopolyribonucleotides from Severo Ochoa and we looked at the experimental fiber diffraction patterns that they gave. In March 1956 Robert Warner published the first account of the optical properties of some polynucleotides that indicated an interaction between polyA and polyU, showing a drop in the absorption at 259 nm when the two polymers were mixed in solution. Fiber diffraction patterns of the poly(A + U) mixture showed a diffraction pattern resembling that of the B form of DNA, implying that the two strands were interacting in a hydrogen bonded manner similar to the adenine and thymine bases of DNA (22).

In the meantime Gary Felsenfeld had started to investigate the interaction in solution using the continuous variation method. I quote from Gary: "I do remember starting the optical titration studies, and that Günther Eichhorn suggested I might like to use the continuous variation method as a good way to analyze data, rather than the standard method I had devised. All went well at the beginning, but one day while I was analyzing the data and trying to get the lines to intersect at 1:1, I realized that there was a consistent bulge on the U side of the plot, and that the results were inescapably looking as though they indicated a 2:1 complex. At that moment David was sitting at his desk with his back to me, and I asked him if there was any way to add a second strand of U to the structure. He replied 'Let's see,' and we went into the lab room next door where the big metal model of poly (A + U) was standing. He picked up a U and tried fitting it into the structure in various ways; after a couple of tries he found the 'Hoogsteen' structure and that was that. It was a very exhilarating experience for us and of course in the end for the entire group."

This three-stranded structure, Figure 1, was the first clue that with these polynucleotides we might discover hitherto unsuspected non-Watson and Crick interactions between the strands (9), although it would be several decades before we would come to appreciate the true complexity of RNA structures. Several years later Hoogsteen solved a crystal structure of 7-methyl adenine and 1-methyl thymine and observed the same interaction that we had proposed for the third strand of our model.

Much more recently our colleague Todd Miles produced a DNA construct containing a portion of triplex structure. Sangkee Rhee crystallized this and determined the structure (23). It showed quite clearly that the triplet uracil attached to the AT pair in just the manner that we had observed in the model some 40 years previously. It also showed a protonated cytosine attached to the GC pair in a manner predicted by Michael Chamberlain.

The RNA polynucleotides offered a field ripe for investigation and we assigned structures to poly I + poly C (5) (DNA-like) and to poly A (34), which had a twofold axis relating the two adenines (Figure 1). A little later Marty Gellert, Marie Lipsett, and I observed the first G-quartet structure from fibers of GMP (10) (Figure 1). Polynucleotide phosphorylase could not make poly G and we had

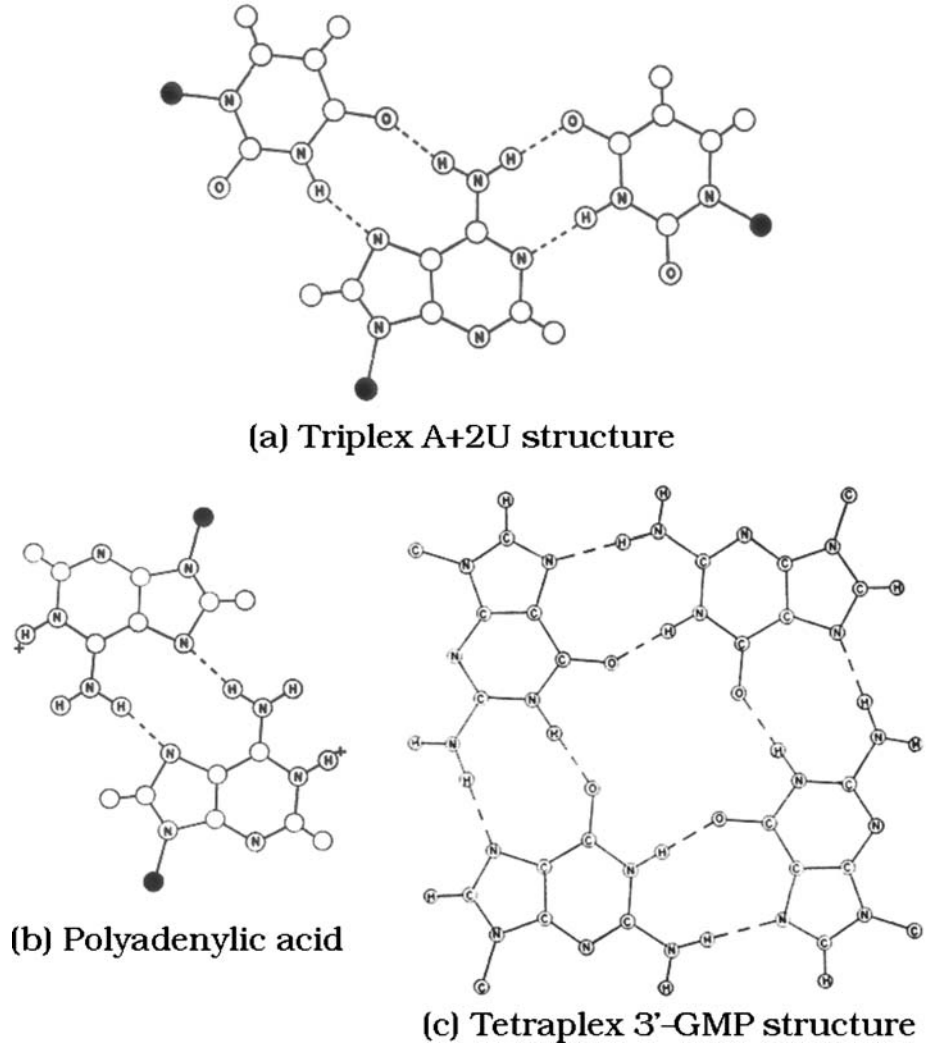

Figure 1 Base interactions for some polynucleotides. (*a*) The triplex structure of poly (A + 2U) (9). (*b*) The symmetric duplex structure of poly A (3). (*c*) The structure of a G tetraplex formed by 3′-GMP and by poly G (10).

always assumed that this was because it formed some sort of self-structure. Marie originally thought that she had been able to make poly G but was then disappointed to discover that what she had was unpolymerized GMP that was forming a viscous solution that looked just like DNA. As soon as she told me this I rushed over and pulled some fibers that gave diffraction patterns that could be explained by the formation of G-quartets. It always seemed to us that nature would have found a way

to use this remarkably stable structure, but it has only been relatively recently that G-quartets have been suggested to be an important motif in telomere structures.

Protein Crystals

In the meantime Alex and Gary had left the NIMH lab, and in 1959 I was becoming somewhat frustrated by the limitations imposed by these rather poorly oriented fiber patterns and was beginning to consider working with real crystals again, this time with proteins. Nucleic acids were my first true love but it was time to move on. There was a lot of discussion at that time about the relevance of protein structures to biology. Were the structures that would be observed in the crystal the same as those of the protein in solution? However, I was convinced by Crick's statement that it was much better to have a definite structure rather than no structure at all. So I went to Cambridge for a six-month stay with John Kendrew, originally intending to collect higher resolution data for myoglobin. When I arrived in the little hut in the Cavendish Courtyard I found that Bror Strandberg and Dick Dickerson were still in the process of finishing the 2.0 Å data collection, so we waited to see what that would produce. Soon the data were assembled onto huge bundles of paper tape and taken over to the Math Lab, where the Edsac2 electronic computer had been programmed by Michael Rossmann to sum Fourier series. Everyone who had any connection with the myoglobin project gathered there that evening in suspenseful anticipation, including David Phillips, who had measured some of the data at the Royal Institution in London. It took a long time for Dick and Bror to get the program to function properly. but eventually, late that night, they calculated the first section through the myoglobin electron density map. The section chosen was one that passed through the presumed position of the heme, based on the previous 6.0 Å map that had been calculated several years earlier. A loud cheer accompanied by a sigh of relief went up when it was observed that there was a large peak of density in the position that could only have been the iron atom.

At this point Dick returned to a new position in the United States and Bror went for a well-deserved vacation in Sweden. The completed three-dimensional map consisted of an α-numerical figure field. I was sharing space in Kendrew's tiny office and we discussed what to do next. One day we decided to calculate a cylindrical section in the region corresponding to one of the rods of the 6 Å map that had been assumed to be α-helical. We carefully interpolated the density onto two cylinders, one with the radius of the α-carbon atoms and the other with that of the β-carbon, which would give the hand of the amino acids. Finally the maps were complete and we drew in the contours and were delighted to see a clear outline of an α-helix (Figure 2). The density at the β-carbon positions showed that this was the correct enantiomorph and indicated that the helix was right handed. The clear density for the oxygen atoms of the backbone established the direction of the chain, and therefore by extension the N and C termini of the molecule. It was exciting because this was the first direct visualization of the α-helix that had previously existed as an interpretation of the fiber diffraction patterns. It also

Figure 2 The electron density of a cylindrical projection of one of the myoglobin rods at 2 Å resolution. The cylinder radius corresponds to that of an α-carbon in the α-helix (14).

confirmed that Perutz and Kendrew had been correct in believing that myoglobin and hemoglobin were mainly made up of α-helices (9). During this period the 5 Å map for hemoglobin was calculated, and Rossmann and Perutz demonstrated that hemoglobin contained four myoglobin-like molecules. It was the beginning of protein crystallography and I was fortunate to have been present when it happened.

The Laboratory of Molecular Biology

When I returned to NIH there were exciting new developments. Hans Stetten, the scientific director, had decided to start a new laboratory within the National Institute of Arthritis and Metabolic Diseases. In keeping with the times it was called the Laboratory of Molecular Biology and Gordon Tomkins was to be the chief of the laboratory. There were to be five sections, with Gordon in charge of one and Bruce Ames, Gary Felsenfeld, Harvey Itano, and myself as the other section heads. Because Gary, Harvey, and I had all worked with Linus Pauling and Bruce had been at Caltech during the same period, there was a very strong Caltech flavor to the Lab. We were a closely knit group and we had group seminars and journal clubs so that we were all familiar with the work in our colleagues' labs.

During the previous two years I had given some thought to our choice of protein for structure determination. I decided that a proteinase would offer simplicity of reaction that might be correlated with function, and I tried to crystallize several of them, some of which gave crystals readily. One was subtilisin, which gave large monoclinic crystals in a few days that diffracted to high resolution. However, Sid Bernhard, my lab chief at that time, dissuaded me from pursuing this further, saying that in his hands this protein was quite fragmented. This in retrospect turned out to be a big mistake and I should have trusted my judgment and pushed ahead with subtilisin regardless. However, I chose to work instead on the gamma form of chymotrypsin, which crystallized in the active pH range and was already known to give large tetragonal crystals.

One of the first postdoctoral fellows who joined me at that time was Paul Sigler. Paul was an M.D. who had just finished his residency at Columbia College of Physicians and Surgeons but was undeterred by the possibility that working with me would mean never seeing a patient again. His enthusiasm and broad chemical background enabled us to make significant progress in finding heavy-atom derivatives that labeled the active site of the enzyme. However, not knowing any crystallography, Paul initially worked on some polynucleotide exchange reactions. These required precise spectrophotometric measurements, and Paul tried to optimize the conditions by controlling the sample temperature with water from a water bath and flushing the sample chamber and the optical chamber with nitrogen. Unfortunately, a hose connection came off and water flooded the optical compartment, effectively fatally compromising the instrument. Fortunately, we were about to leave the Mental Health Lab for our new Laboratory of Molecular Biology in the Arthritis and Diabetes Institute, where Paul continued his experiments using a brand new spectrophotometer that had been purchased by Martin Gellert. This time he decided to flush the sample chamber with helium, which we had been using for the fiber diffraction cameras. To his horror the helium penetrated the windows of the photocells, putting them out of action. So we decided that the time was ripe for Paul to turn to protein crystallography. With John Kallos he obtained an isomorphous labeling of the active site using the pair of inhibitors, pipsyl and tosyl fluoride (30). He then left for Cambridge to get a PhD with David Blow, who had started to work on α-chymotrypsin.

It took us several years to acquire sufficiently sophisticated instrumentation for data collection and processing. When we finally determined the structure, I was chagrined to discover that, despite the difference in pH and despite the fact that γ-chymotrypsin crystals were active and α-chymotrypsin crystals were not, there was essentially no difference between this structure and that determined previously for α-CHT in David Blow's lab in Cambridge. However, we were able to bind oligopeptide inhibitors to the active site and David Segal was able to provide a reasonable structural basis for the extended subsite specificity (28).

In 1967 we were joined by Brian Matthews, who had worked with David Blow and Paul in Cambridge. Brian worked on several things when in the lab, but perhaps the most noteworthy was his paper on the assessment of the number of molecules within the asymmetric unit of a protein crystal. He used the volume per unit molecular weight of the protein, Vm, as a criterion and with typical thoroughness showed the range of values that Vm could adopt (19). This has become I believe the most-cited paper in protein crystallography, with Vm referred to as the Matthews coefficient. Many postdocs are quite surprised to learn that the "m" does not refer to Matthews.

ANTIBODIES

In the meantime we had become interested in the structure of antibodies. Bill Terry, who had taken a course I had given at the NIH on crystallography, came to my lab one day and asked me to look at precipitates of cryoglobulins. These were samples from patients with multiple myloma and cryoglobulinemia who made lots of circulating antibody that would precipitate in the cold. Many of the precipitates were microcrystalline, but our attempts to grow larger crystals by slow cooling were not successful. So Bill put out a request to laboratories working on cryoglobulins and obtained a response from one that described one protein that yielded crystals that grew like rock candy overnight. This protein was called Dob, presumably after the patient's name, and when we received a sample we found the size of the crystals had not been exaggerated. I mounted a crystal and with Brian Matthews we determined the crystal properties. Unfortunately, the crystals did not diffract beyond about 4.5 Å, but they were the first crystals of an intact antibody (34). With Ragupathy Sarma, a postdoc who had come to my lab, we determined a low-resolution structure that showed a T-shaped molecule. At the same time Louis Labaw, who worked in the same building, investigated the structure by electron microscopy. The Dob crystals had a relatively short axis and he obtained sections of the cross-linked crystal perpendicular to this direction. The electron micrographs looked like a poorly ploughed field, but by laboriously carrying out multiple superpositions in the two known crystal axis directions, he obtained a beautiful image of the crystalline array of T-shaped molecules (26).

The low resolution of the crystals prevented a complete analysis of the structure. So with David Segal, who had joined our group, we contacted Mike Potter at the

National Cancer Institute (NCI), who had been doing pioneer work on mouse antibodies generated by plasmacytomas in mice. With Stuart Rudikoff, Mike obtained pure concentrated Fab preparations that David then proceeded to crystallize. The first to crystallize was McPC603, a Fab that Potter had been able to show bound quite strongly to the hapten phosphocholine. Eduardo Padlan had joined our group, and with the aid of heavy-atom derivatives we determined the structure, showing the phosphocholine neatly trapped in the middle of the Fab-combining site (27). With the structure of a Bence-Jones protein by Allen B. Edmundson and Marianne Schiffer at Argonne National Laboratory and of a Fab by R.J. Poljak's group at Johns Hopkins, these results established the importance of the immunoglobulin domain structure and confirmed the proposals of Elvin A. Kabat and Tai Te Wu regarding the use of what were then called the hypervariable loops (now CDRs) in providing the basis for antibody specificity. We also determined the structure of a second Fab, J539, that had galactan binding specificity, but were unable to obtain crystals of the complex.

Fab COMPLEXES WITH LYSOZYME

As the production of monoclonal antibodies became established, it was possible to consider the investigation of Fab complexes with a protein antigen. Again we turned to Mike Potter's lab and this time it was Sandra Smith-Gill who produced several Fabs that bound quite tightly to lysozyme. We obtained crystals of two of these complexed with lysozyme, HyHEL-5 and HyHEL-10 (21, 29). The HyHEL-5 lysozyme complex crystallized in two forms that showed significantly different elbow-bend angles in the Fabs despite binding to identical epitopes, indicating that the elbow-bend angle was a manifestation of the flexibility of the linkage between the V and C regions of the Fab.

The HyHEL-5 and HyHEL-10 Fabs bound to different parts of the lysozyme surface and, with the D1.3 structure from the Poljak lab (1), formed a set of nonoverlapping epitopes (Figure 3*a*, see color insert), leading to the conclusion that any exposed part of a protein surface was potentially antigenic and that special sequences did not appear to be necessary for antibody recognition.

Perhaps the most striking feature of these antibody-lysozyme complexes, combined with Peter Colman's antineuraminidase structure, was the exceptional fit between the two surfaces, with good shape complementarity leaving few holes to be filled by water molecules (Figure 3*b*), although the periphery of the interface was well populated by solvent.

Since those early days this field has exploded: There are now more than 600 structures in the Protein Data Bank under the heading "antibody." These include many complexes with proteins, peptides, haptens, and nucleic acids. They include intact antibodies and single-chain Fvs. However, the general principles that were determined have still by and large prevailed. They would have been gratifying to Paul Ehrlich, who predicted many of them over 100 years ago.

ASPARTYL PROTEASES

In about 1971 Ian Swan joined our group and expressed a lack of interest in any of the current projects. So we looked around for some interesting proteins that would be structurally different from those being studied. This was not hard at that time, although it would be a much more difficult task these days when most of the shapes are rapidly being determined. Naturally, proteases won the day and Ian crystallized a protease from the fungus *Rhizopus chinensis* that was related to pepsin and was quite different from the serine proteases. We started work on this, joined by E. Subramanian, but were hampered by the fact that there was no sequence known for this protein (31). Two other aspartyl proteases, as they came to be known, were determined almost simultaneously in the laboratories of Michael James and Tom Blundell, and these three structures were similar and formed a superfamily like the serine proteases. Later, when the sequence was determined by Kenji Takahashi, Kaza Suguna refined the Rhizopus structure to a high resolution (32). A study of inhibitors bound to the protease led us to propose a mechanism of action for these proteases (33), which was interesting in itself but acquired more significance when it emerged that the HIV protease was also an aspartyl protease. At that time we tried hard to develop a project on HIV protease, but we were unable to find a suitable source of material. Some time later the structure was determined by the Navia group at Merck and later at higher resolution and with a correct chain connectivity by the Wlodawer group at the NCI. Both Manuel Navia and Alex Wlodawer had been overlapping postdoctoral fellows in our group and I was delighted with their success.

HIV INTEGRASE

Frustrated by our inability to work on the protease, we turned to Bob Craigie, who was working at the other end of our building on the third enzyme of HIV, the integrase. This was a most intractable protein and for a while defied all our efforts to obtain crystals. We focused on the core domain, which contained the catalytic residues, but even this suffered from problems of insolubility. Eventually, Craigie identified a mutant Phe-185-Lys that was quite soluble and was crystallized by Alison Hickman, and the structure was determined by Fred Dyda (8). The structure was homologous to RNase H and to a number of retrotransposases. However, the conditions of crystallization had resulted in a modification of the protein such that it could not bind to magnesium ion, which was a factor in the catalytic action of the integrase. This presented a problem until Yehuda Goldgur found another mutant that crystallized under different conditions that did not involve cacodylic acid and was able to bind magnesium. The old structure was almost identical to the new one except for a distortion of part of the active-site residues owing to the interaction of the cacodylate with a cysteine side chain. The enzyme in this crystal form

bound to a single magnesium (12) and Yehuda was then able to obtain a complex with an inhibitor that was, despite all our efforts, the first and last structure of an inhibitor bound to the HIV integrase active site to date (11). This has been especially frustrating because of the attractiveness of the integrase as a target for therapeutic intervention.

TRYPTOPHAN SYNTHASE: TUNNELS AND CHANNELING

Most of the structures that we determined were the result of collaborations with other NIH scientists. One of these involved Edith Miles, who had been working on the enzyme complex tryptophan synthase, which carries out the two final steps of tryptophan synthesis in bacteria. This enzyme had been studied extensively by Charles Yanofsky, who had observed the occurrence of second-site revertants, which intrigued me. However, as is not uncommon in science, the unexpected intervened and we ended up focusing on quite a different property of the complex.

The enzyme is made up of four subunits arranged $\alpha\beta\beta\alpha$. The α-subunits convert indole glycerol phosphate into indole, which is combined with serine in the β-subunit to make tryptophan. The indole intermediate is uncharged and lipophylic, and if released into the cytoplasm, it could leak out through the cell membrane. It had been demonstrated by Yanofsky that there was "channeling" of the indole from the α- to the β-subunit (reviewed in Reference 7a). Edith obtained crystals of the enzyme from *Salmonella typhimurium* and Craig Hyde determined the crystal structure. We had anticipated that the two active sites would be in close proximity in order to explain the channeling, but to our surprise they were 25 Å apart. The true mechanism emerged when Craig noticed that there was a tunnel that joined the two active sites and was large enough to permit the passage of indole (13) (Figure 4, see color insert). Subsequent high-resolution analysis by Sangkee Rhee showed that the tunnel was partially blocked by the presence of a phenylalanine side chain, but this side chain could be moved out of the tunnel in the presence of K^+ or Cs^+, which suggests that it acts as a gate (25). More recently, other tunnels have been observed in enzyme systems such as carbamoyl synthase and glutamine amidotransferases, which suggests that tunnels are a frequently used mechanism for channeling unstable intermediates in multienzyme complexes (20).

SOME RECENT STRUCTURES

During the past decade or so, in keeping with the rest of protein crystallography, we have picked up our pace and determined several structures in addition to the work on the HIV integrase and the tryptophan synthase. These included TGFbeta II, which Peter Sun showed had an extended structure that has an unusual cysteine knot arrangement. It was a member of an extensive superfamily that included nerve

growth factor, platelet-derived growth factor, and human chorionic gonadotrophin. A comparison of the structures showed that, despite the fact that there is some similarity in the monomer fold, the active form is a dimer and the monomers associate in different ways to form the dimers (4).

In a collaboration with Bob Martin and Lee Rosner, my longtime colleagues in the Laboratory, Sangkee Rhee determined the structure of MarA in complex with its cognate DNA-binding site, the first structure for an AraC family transcriptional activator (24). The structure is rather unusual, having two domains, each with a helix-turn-helix motif that binds to the appropriate base pairs. The two domains are linked by a helix that imposes a certain rigidity on the positions and orientations of the two helix-turn-helix motifs, so that the binding produces a bend in the DNA.

We determined the structure of the enzyme deoxyhypusine synthase. The original structure was solved by Liao Der-Ing (18) and the protein was provided by Edie Wolff and Myung-Hee Park of the Dental Institute. We looked at this protein because of its interesting enzymology, using spermidine as a substrate, and for its remarkable specificity. It modifies a single lysine side chain, Lys-50, on eukaryotic initiation factor 5A (eIF5A). The reaction proceeds in two steps: the first is the attachment of an aminobutane moiety from spermidine to a lysine of the enzyme. In the second step this group is transferred to Lys-50 of the eIF5A. The molecule forms a tetramer, a dimer of dimers with active sites located toward the bottom of tunnels between each dimer pair. Curiously, access to the active sites is blocked by the presence of an N-terminal helix from a third subunit forming a sort of ball-and-chain motif. Crystallization under different conditions by Tim Umland yielded a form in which the active site was accessible with a disordered ball and chain. In this form it was possible to soak in a spermidine analog inhibitor that bound in the pocket and confirmed some of our early predictions about the mechanism of action (36). We are continuing to work on producing a complex with eIF5a.

Frequently, projects do not turn out quite the way one would wish. This was true of a collaboration with Reed Wickner on the yeast prion protein Ure2p, which with Sup35 had been shown by Reed to be prion proteins. Ure2p has an N-terminal prion domain rich in asparagines and glutamines, and we had hoped to get some idea of the native structure and its interactions with the C-terminal domain, which was homologous to the GST superfamily. However, during the expression of the protein we lost the prion domain and ended up with a GST-like structure that is a dimer and that still provided a platform to visualize the sites of interaction between the N- and C-terminal domains (35).

There have also been a number of other structures including the catalytic domain of the *Pseudomonas* toxin (15, 16) and enzyme 1 (17). If these seem to be a motley collection of proteins, I have to confess that I have yet to meet a protein that I did not like and that unanticipated and interesting results have emerged from most of these studies. Currently, our main interest has returned to immunology in an effort to understand the principles underlying the tremendous diversity of the proteins of the innate immune system (2).

CONCLUSION

I sometimes wonder how a country boy from Wales managed to stumble into such a rich research environment. I can only say that luck has played a large role in my life, particularly in the early stages. I was fortunate to attend Oxford and to go to Caltech, and I was extraordinarily lucky to eventually find myself at NIH. Most of all I was lucky to be included in and to contribute to the early days of what we now call structural biology. In those days it took years to determine protein structures, but each structure provided a lot of new information that could be assimilated and discussed and there was plenty of time to do this. It is now quite overwhelming to deal with the 70 or so new structures that are being deposited in the Protein Data Bank every week.

When I came to the NIH in 1955 I never thought that I would stay so long. It was only gradually that I realized what a superb place it is to do research and how many outstanding scientists there are in such a variety of disciplines. For the first six years or so I was a typical postdoc engaged in direct lab work. Then when I moved to our Laboratory of Molecular Biology, I was allowed to have three or four postdocs and began to have less time to do my own research. Many of my postdocs from that period went on to have distinguished careers in academic research. I can truthfully say that any success I have had in achieving our research goals has been entirely due to the many outstanding postdocs who have passed through my lab.

I also owe much to the collaborations I have been fortunate enough to have and the resulting friendships made with people such as Michael Potter, Edith Miles, Carl Piez, and Buzz Baldwin, Ira Pastan, Edith Wolff from outside our Laboratory, as well as Gary Felsenfeld, Martin Gellert, Terrell Hill, Bill Eaton, Bob Martin, Lee Rosner, Todd Miles, and Robert Craigie from within. I also made many friends through sailing and tennis. In particular my sailing experiences with Chris Anfinsen and Brad Thompson stand out, especially a trip to Bermuda with Brad in his thirty-foot boat that we shall never forget.

The conditions for doing science at NIH have also been outstanding. In Volume 26 of the *Annual Review of Biophysics and Biomolecular Structure* Fred Richards describes a day in the life of a typical 35-year-old faculty member at one of today's research universities. The daily grind of this hapless fellow can only be described as horrendous. Such a hypothetical individual at this stage of his career at the NIH would have none of these demands on his time but, once he had penetrated the security barrier, he would be free to devote himself entirely to his research project.

In closing, I am reminded of the eloquent words of Phil Handler, who was President of the National Academy of Sciences from 1969 to 1981: "To be so positioned as to spend a large fraction of your waking hours with those people in the United States whom you most admire and most respect, and some of whom you love, is a most astonishing privilege. It has filled and enriched my life the way nothing else conceivably might have done."

The *Annual Review of Biophysics and Biomolecular Structure* is online at http://biophys.annualreviews.org

LITERATURE CITED

1. Amit AG, Mariuzza RA, Phillips SE, Poljak RJ. 1986. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science* 233:747–53
2. Bell JK, Mullen GED, Leifer CA, Mazzoni A, Davies DR, Segal DM. 2003. Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol.* 24:528–33

2a. Cochran W. 1952. A relation between the signs of structure factors. *Acta Crystallogr*. 5:65–67

2b. Cochran W, Crick FHC, Vand V. 1952. The structure of synthetic polypeptides. 1. The transform of atoms on a helix. *Acta Crystallogr*. 5:581–86

3. Crick FHC, Davies DR, Rich A, Watson JD. 1961. The molecular structure of polyadenylic acid. *J. Mol. Biol*. 3:71–86
4. Daopin S, Piez KA, Ogawa Y, Davies DR. 1992. Crystal structure of transforming growth factor-β2: an unusual fold for the superfamily. *Science* 257:369–73
5. Davies DR. 1960. Polyinosinic plus polycytidylic acid: a crystalline polynucleotide complex. *Nature* 196:1030–31
6. Davies DR, Blum JJ. 1955. The crystal structure of parabanic acid. *Acta Crystallogr.* 8:129–36
7. Davies DR, Pasternak R. 1956. A refinement of the crystal structure of succinamide. *Acta Crystallogr.* 9:334–40

7a. Davis BD. 1958. On the importance of being ionized. *Arch. Biochem. Biophys*. 78: 497–509

8. Dyda F, Hickman A, Jenkins T, Engelman A, Craigie R, Davies DR. 1994. Crystal structure of the catalytic domain of HIV-1 integrase reveals similarity with other polynucleotidyl transferases. *Science* 266:1981–86
9. Felsenfeld G, Davies DR, Rich A. 1957. Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* 79: 2023–24
10. Gellert M, Lipsett MN, Davies DR. 1962. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA* 48:2013–18
11. Goldgur Y, Craigie R, Cohen GH, Fujiwara T, Yoshinaga T, et al. 1999. Structure of the HIV-1 integrase catalytic domain complexed with an inhibitor: a platform for antiviral drug design. *Proc. Natl. Acad. Sci. USA* 96:13040–43
12. Goldgur Y, Dyda F, Hickman AB, Jenkins TM, Craigie R, Davies DR. 1998. Three additional structures of the core domain of HIV-1 integrase: an active site that binds magnesium. *Proc. Natl. Acad. Sci. USA* 95:9150–54

12a. Hauptman H, Karle J. 1952. Crystal-structure determination by means of a statistical distribution of interatomic vectors. *Acta Crystallogr.* 5:48–59

13. Hyde CD, Padlan EA, Ahmed SA, Miles EW, Davies DR. 1988. Three-dimensional structure of the tryptophan synthase multienzyme complex from *Salmonella typhimurium*. *J. Biol. Chem.* 263:17857–71
14. Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, et al. 1960. Structure of myoglobin. *Nature* 185:422–27
15. Li M, Dyda F, Benhar I, Pastan I, Davies DR. 1995. The crystal structure of *Pseudomonas aeruginosa* exotoxin domain III with nicotinamide and AMP: conformational differences with the intact exotoxin. *Proc. Natl. Acad. Sci. USA* 92:9308–12
16. Li M, Dyda F, Benhar I, Pastan I, Davies DR. 1996. Crystal structure of the catalytic domain of Pseudomonas exotoxin A complexed with a nicotinamide adenine dinucleotide analog: implications for the activation process and for ADP

ribosylation. *Proc. Natl. Acad. Sci. USA* 93:6902–06
17. Liao DI, Silverton E, Seok YJ, Lee BR, Peterkofsky A, Davies DR. 1996. The first step in sugar transport: crystal structure of the amino terminal domain of enzyme I of the *E. coli* PEP: sugar phosphotransferase system and a model of the phosphotransferase complex with Hpr. *Structure* 4:861–72
18. Liao DI, Wolff E, Park MH, Davies DR. 1998. Crystal structure of the NAD complex of human deoxyhypusine synthase: an enzyme with a ball and chain mechanism for blocking the active site. *Structure* 6:23–32
19. Matthews BW. 1968. Solvent content of protein crystals. *J. Mol. Biol.* 33:491–97
20. Miles E, Rhee S, Davies DR. 1999. The molecular basis of substrate channeling. *J. Biol. Chem.* 274:12193–96
21. Padlan EA, Silverton EW, Sheriff S, Cohen GH, Smith-Gill SJ, Davies DR. 1989. Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab lysozyme complex. *Proc. Natl. Acad. Sci. USA* 86:5938–42
22. Rich A, Davies DR. 1956. A new two-stranded helical structure: polyadenylic acid and polyuridylic acid. *J. Am. Chem. Soc.* 78:3548–49
23. Rhee S, Han Z-J, Liu K, Miles HT, Davies DR. 1999. Structure of a triple helical DNA with a triplex-duplex junction. *Biochemistry* 38:16810–15
24. Rhee S, Martin RG, Rosner JL, Davies DR. 1998. A novel DNA-binding motif of MarA: the first structure for an AraC family transcriptional activator. *Proc. Natl. Acad. Sci. USA* 95:10413–18
25. Rhee S, Parris KD, Ahmed SA, Miles EW, Davies DR. 1996. Exchange of K^+ or Cs^+ for Na^+ induces local and long-range changes in the three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ complex. *Biochemistry* 35:4211–21
26. Sarma VR, Davies DR, Labaw LW, Silverton EW, Terry WD. 1971. Crystal structure of an immunoglobulin molecule by X-ray diffraction and electron microscopy. *Cold Spring Harbor Symp. Quant. Biol.* XXXVI:413–19
26a. Sayre D. 1952. The squaring method: a new method for phase determination. *Acta Crystallogr.* 5:60–65
27. Segal DM, Padlan EA, Cohen GH, Rudikoff S, Potter M, Davies DR. 1974. The three-dimensional structure of a phosphorylcholine binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Natl. Acad. Sci. USA* 71:4298–302
28. Segal DM, Power JC, Cohen GH, Davies DR, Wilcox PE. 1971. The substrate binding site in bovine chymotrypsin AR: a crystallographic study using peptide chloromethyl ketones as site specific inhibitors. *Biochemistry* 10:3728–37
29. Sheriff S, Silverton EW, Padlan EA, Cohen GH, Smith-Gill SJ, et al. 1987. Three-dimensional structure of an antibody-antigen complex. *Proc. Natl. Acad. Sci. USA* 84:8075–79
30. Sigler PB, Skinner HCW, Coulter CL, Kallos J, Braxton H, Davies DR. 1964. An isomorphous heavy atom substitution at the active site of gamma-chymotrypsin. *Proc. Natl. Acad. Sci. USA* 51:1146–51
31. Subramanian E, Swan IDA, Liu M, Davies DR, Jenkins JA, et al. 1977. Homology among acid proteases: comparison of crystal structure at 3 Å resolution of acid proteases from *Rhizopus chinensis* and *Endothia parasitica*. *Proc. Natl. Acad. Sci. USA* 74:556–59
32. Suguna K, Bott RR, Padlan EA, Subramanian E, Sheriff S, et al. 1987. Structure and refinement at 1.8 Å resolution of the aspartic proteinase from *Rhizopus chinensis*. *J. Mol. Biol.* 196:877–900
33. Suguna K, Padlan EA, Smith CW, Carlson WD, Davies DR. 1987. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: implications for a mechanism of action. *Proc. Natl. Acad. Sci. USA* 84:7009–13

34. Terry WD, Matthews BW, Davies DR. 1968. Crystallographic studies of a human immunoglobulin. *Nature* 220: 239–41
35. Umland TC, Taylor KL, Rhee S, Wickner RB, Davies DR. 2001. The crystal structure of the nitrogen regulation fragment of the yeast prion protein Ure2p. *Proc. Natl. Acad. Sci. USA* 98:1459–64
36. Umland TC, Wolff EC, Park MH, Davies DR. 2004. A new crystal structure of deoxyhypusine synthase reveals the configuration of the active enzyme and of an enzyme-NAD-inhibitor ternary complex. *J. Biol. Chem*. 279:28697–705
37. Zachariasen WH. 1952. A new analytical method for solving complex crystal structures. *Acta Crystallogr*. 5:68–73

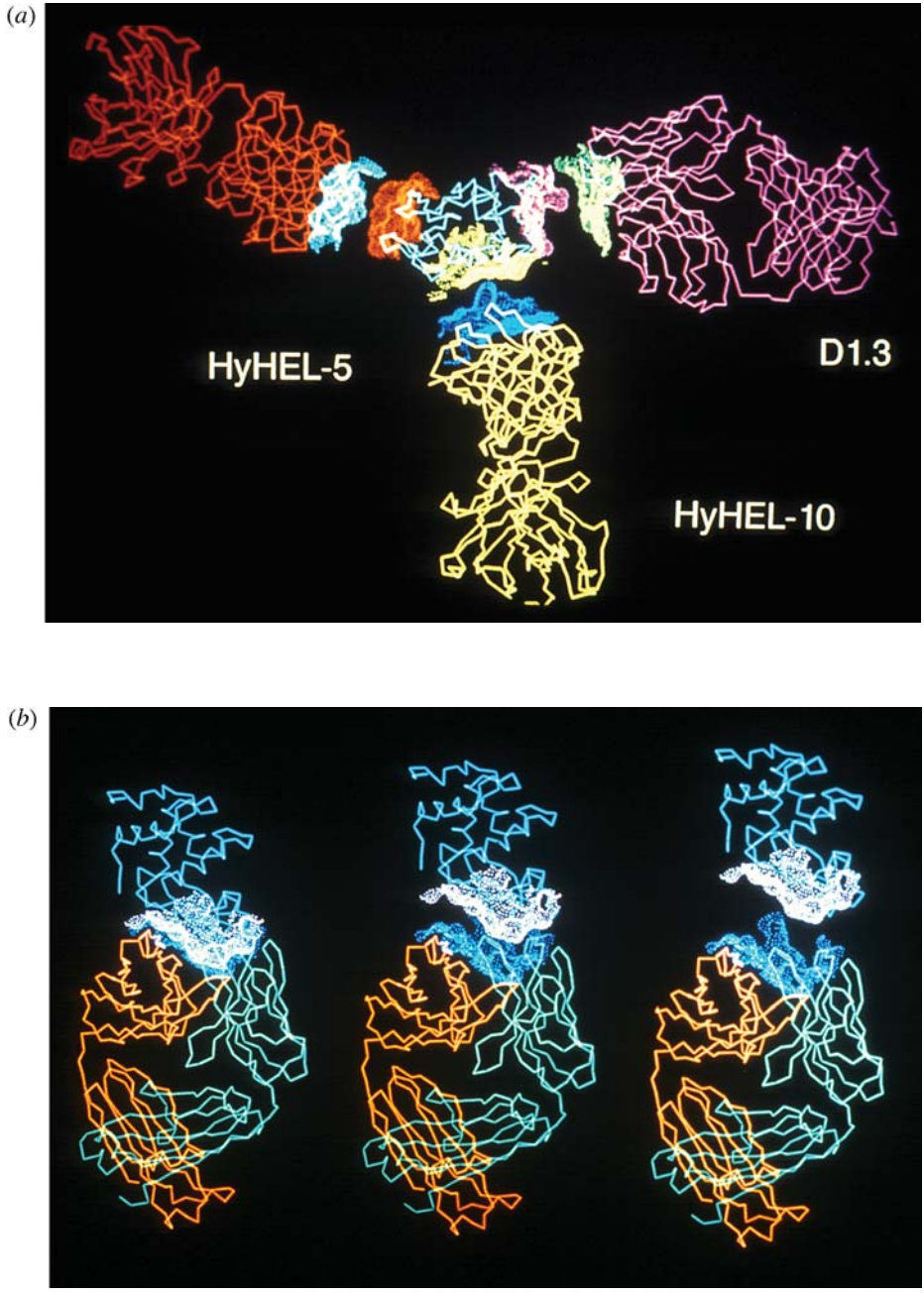

Figure 3 (*a*) Three antibody Fabs that bind to lysozyme. The Fabs are pulled back from contact with the lysozyme to reveal the nonoverlapping epitopes (1, 21, 29). (*b*) HyHEL-5 and lysozyme showing the shape complementarity of the interaction (29).

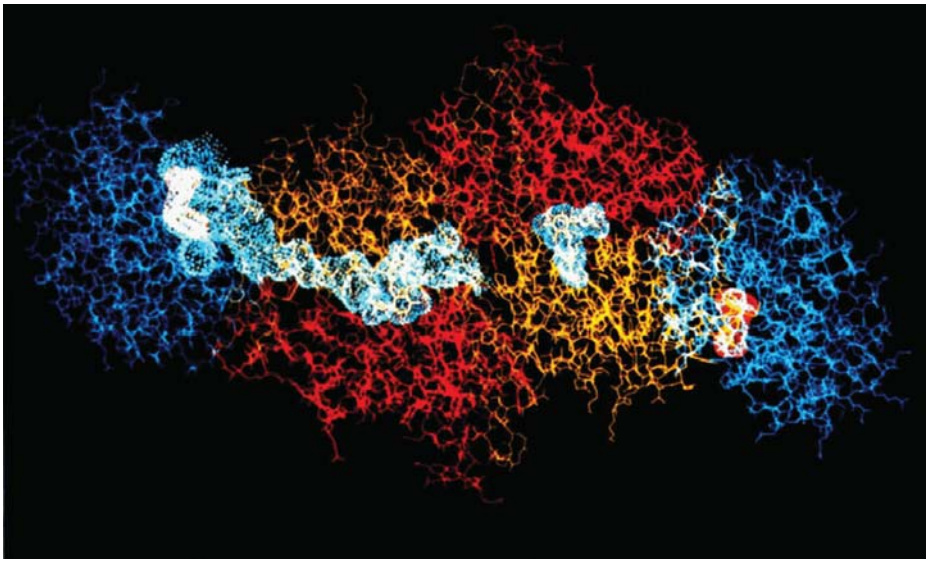

Figure 4 The tryptophan synthase structure. The α-subunits are blue and the two domains of the β-subunits are red and yellow. The active sites of the two subunits are highlighted by the binding of indole-3-propanol phosphate and pyridoxal phosphate, respectively. The indole tunnel is shown in blue and passes from the α active site, between the two β-subunit domains, to the β active site (13).