Miao *et al.*[16] have carried out soft X-ray (wavelength ≅ 1.7 nm) diffraction using an undulator beam line at NSLS. The specimen examined was a collection of gold dots (≅ 100 nm^3 each) arranged in the form of English alphabets. The diffraction was registered using a 512 × 512 CCD pixels detector. Data analysis is based on reconstruction using oversampling technique proposed by Bates[17] earlier. Figure 3 *a* shows the reconstructed image of the specimen shown in Figure 3 *b*, demonstrating successfully that soft X-ray diffraction is indeed useful in such imaging requirements.

It is believed that one can extend this technique to image large cell structures in biological systems and other materials.

1. Rietveld, H. M., *Acta Crystallogr.*, 1967, **22**, 151.
2. Rietveld, H. M., *J. Appl. Crystallogr.*, 1969, **2**, 65.
3. For a brief review see Manohar, H., *Curr. Sci.*, 1983, **52**, 39.
4. Cascarano, C. *et al.*, *J. Appl. Crystallogr.*, 1992, **25**, 310.
5. David, W. I. F., *Nature*, 1990, **346**, 731.
6. Newdam *et al.*, *NIST Spl Publn 846*, 1992, 80.
7. Giacovazzo, C., *Acta Crystallogr.*, 1996, **A52**, 331.
8. Wessels, T., Baerlocher, C. and McCusker, L. B., *Science*, 1999, **284**, 477.
9. Faigel, G. and Tegze, M., *Rep. Prog. Phys.*, 1999, **62**, 355.
10. Gabor, D., *Proc. R. Soc.*, 1949, **97**, 45.
11. Howells *et al.*, *Synchrotron Radiat. News*, 1990, **3**, 23.
12. Tegze, M., Faigel, G., Marchesini, S., Belakhovsky and Chumakov, A. I., *Phys. Rev. Lett.*, 1999, **82**, 4847.
13. Coppens, P., *X-ray Charge Densities and Chemical Bonding*, Oxford, NY, 1997.
14. Zuo, J. M., Kim, M., O'Keeffe, M. and Spence, J. C. H., *Nature*, 1999, **401**, 49.
15. Zuo, J. M., *Mater. Trans. IIM*, 1998, **39**, 938.
16. Miao, J., Charalambous, P., Kirz, J. and Sayre, D., *Nature*, 1999, **400**, 342.
17. Bates, R. H. T., *Optik*, 1982, **61**, 247.

K. R. Rao lives at 'Gokula', 29/2, 11th Cross Road, III Main Road, Malleswaram, Bangalore 560 003, India.

Three dimensional structures at the heart of the central dogma of molecular biology: An end of millennium gift-pack from crystallographers

Dipankar Chatterji and Preethi Chander

The fundamentals of gene expression were established by the mid-1960s, primarily from the enunciation of the central dogma of molecular biology, i.e. DNA → RNA → protein and the pioneering work of Francois Jacob and Jacques Monod on the regulation of this circuit. Thus, several research groups all over the world spent over 30 years to understand every step associated with the central dogma. A few organisms were the hot favourites like the bacterium *Escherichia coli* and the common yeast *Saccharomyces cerevisiae*. Around the mid-1980s, all the broad outlines of the steps leading to protein synthesis from DNA, were unveiled. However, there remained major gaps in our understanding of the molecular details of the processes of transcription and translation.

Whenever there is a technological hang-up scientists spend money on a different route or a different technology. We knew about the DNA replication cycle and the role of DNA polymerase, the complexity of eukaryotic genome like chromatin was well established, the mechanisms of RNA synthesis and protein synthesis on mRNA templates was clearly understood. However, the structure of chromatin was not known and therefore we were surprised by the rapidity with which transcription processes can continue over a mass of DNA so tightly bound with the histone octamer. Similarly, the structure of RNA polymerase which actively transcribes the DNA chain or the structure of the ribosome eluded us and as a result the beauty of the cell machinery remained obscure. From our knowledge and experience, it was clear that single crystal X-ray diffraction analysis at atomic dimensions, is a necessity for a clear understanding of these basic biochemical events. There were however, formidable obstacles to be overcome in the crystallization of the immensely complex entities involved in protein synthesis, based on direction encoded in the DNA template. Alternate approaches like NMR measurements, molecular modelling, analysis of the structure in parts, photochemical cross linking, fluorescence energy transfer, etc. yielded a large volume of data on the structures of the nucleosome, RNA polymerase and ribosome; but these were almost like listening to cricket commentaries on the radio when you are deprived of a ticket at the venue (in pre-television days)! It is needless to say we have been all waiting for X-ray structures of these macromolecules and news (gossip?) arrived at regular intervals that things were happening!

But, what was the problem? The molecular complexity of the nucleosome, RNA polymerase and ribosomes is the reason for their obscurity! They have too many components within them; the nucleosome has a histone octamer bound to short DNA, RNA polymerase from bacteria have 5–6 subunits whereas the yeast enzyme has 10–12 subunits, and the ribosome from bacteria has 54 proteins with 3 RNA molecules. In addition, there is no apparent symmetry in these molecular assemblies, unlike viruses. Furthermore, the choice of the bacterium *E. coli*

as a source of these cellular constituents in most early studies may have been an unfortunate one.

The first breakthrough may be traced to the mid-1980s when the structures of the nucleosome core particle at 7 Å resolution and histone octamer at 3.1 Å resolution were revealed (Figure 1)[1,2]. First of all, it should be mentioned that a structure at 7 Å resolution is in no way telling us anything at atomic dimensions. Similarly at 3.1 Å, a protein structure is blurred providing at best the conformation of the polypeptide backbone. Or in other words, the complexity of the side chains completely lost. Nonetheless, these low resolution pictures provide the first glimpses of the functional assemblies. About a decade later, 2.8 Å resolution map of the nucleosome core particle was available which showed the relative arrangements of histone octamer and the DNA[3] (Box 1).

The first important observation that came out of the nucleosome structure determination was the visualization of right-handed B-DNA on the outside of

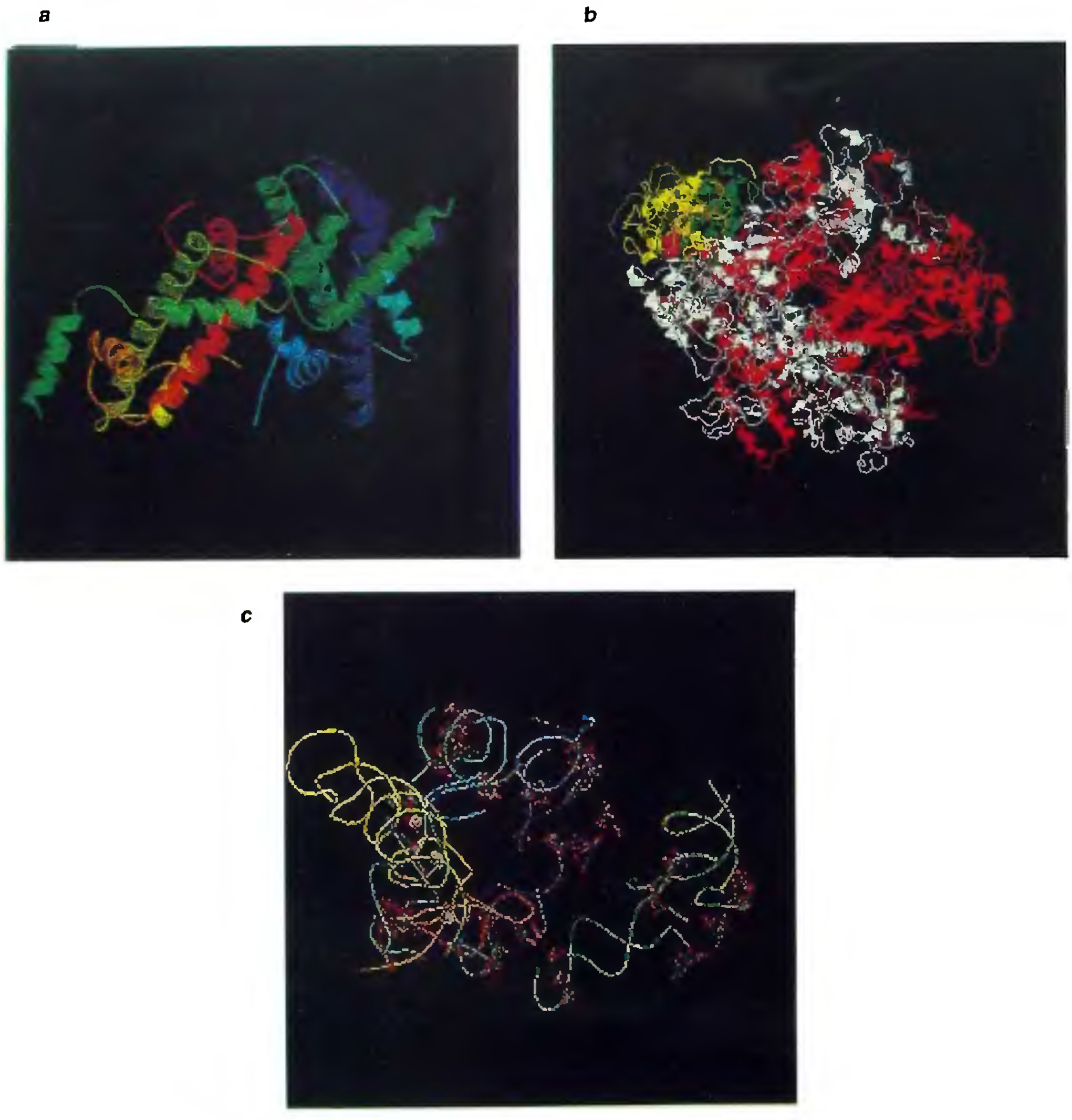

Figure 1. The histone octamer (*a*), RNA polymerase (*b*) and 70S ribosome (*c*) (RNA only) taken from protein data bank. Different colour codes show different protein subunits or RNA molecules. Readers are adviced to take a look at the original articles for electron-density map.

the histone octamer. The DNA backbone contained several sharp bends resulting in numerous interactions with the histone octamer within. The central turn of the superhelix and H3.H4 tetramer have a two-fold axis of symmetry whereas $H_2A.H_2B$ has no symmetry element within. Interestingly, this structure also predicted how H1 may be bound with the nucleosome. The structure of the histone octamer at higher resolution (Å) showed that the folded histone chains are elongated rather than globular and are assembled in 'handshake' motif. The individual polypeptides share a common central structural element of the helix–loop–helix type which was named as the 'histone fold'[2,3].

We had to wait a complete decade to come to terms with a prokaryotic RNA polymerase. Even as the millennium was almost slipping away, in September 1999, the single crystal structure of RNA polymerase was deciphered at 3.3 Å resolution[4–6]. The enzyme was isolated and crystallized from *Thermus aquaticus,* a thermophilic (high temperature stable) organism, not *E. coli.* However, the enzymes from these two sources are homologous both in terms of amino acid sequences and by polyacrylamide gel electrophoresis analysis. We will see later that even the ribosome was crystallized from *T. aquaticus*, and not *E. coli* (Box 2).

The structure as shown in Figure 1 is termed as a 'crab-claw' model. Characteristically, only the back-bone folds were noticeable but it showed how the junction of β and β' subunit holds the active site with the participation of a central Mg(II) ion. This is the site which contains the conserved catalytic centre of RNA polymerase consisting of the -NADFDGD-motif. The N-terminal domain of the α subunit forms the base of the structure over which β and β' stand. Most notably, two identical α subunits are asymmetric in the overall structure of RNA polymerase; one contacting the β subunit whereas the other binds β'. The stoichiometric presence of the ω subunit with the core RNA polymerase was shown for the first time from the X-ray structure. ω appears to help in folding of RNA polymerase. However, the major lacuna of this structure was the absence of all important σ subunit, which is present only with the holoenzyme. The structure of the holoenzyme alone and bound with the promoter DNA is now anxiously awaited. It will be indeed interesting to see the conformation changes that the core enzyme undergoes upon σ-association. Similarly, yeast RNA polymerase II has been resolved[6] at 5 Å resolution. Both prokaryotic and eukaryotic RNA polymerase have a lot of structural elements common amongst them.

During the same month when we were just beginning to appreciate the intricacies of RNA polymerase structure, we were struck again by crystallographers with a sense of finality. This time it was the ribosome from *T. aquaticus*[7–9]. The tiny particle in the cell that translates the message in the form of DNA base sequence into all the proteins needed for life was lying in front of us with its divine beauty! (Box 3).

However, the story of solving the ribosome structure was a long and arduous path with international players involved in a close race for the ultimate glory. An Indian-born scientist Venki Ramakrishnan, working first at the University of Utah and later at the MRC laboratory, Cambridge first showed the structure of the 30S subunit[7] followed by 50S and 70S structures[8,9]. Behind all this glory, there was Ada Yonath (I just cannot stop making a comparison with Rosalind Franklin of DNA fame) of the Weizmam Institute, Israel who doggedly worked for several decades to make progress in ribosome crystallography[10,11]. She started the field in the 1970s and she was the one who first thought it could be done. However, she was not the first one who got the structure into the print!

There was another interesting twist to this whole game and that was Soviet effort to solve the ribosome structure in Alexander Spirin's laboratory in Protein Research Institute at Puschino, Russia in the pre-Gorbachev era. By 1987, names Yusupov had crystals of the 70S ribosome, but they realized that it was impos-

BOX 1. Replication

Replication of duplex DNA is the first step of the central dogma where the origin of replication is recognized by a complex set of proteins. Before replication proceeds, the original duplex is separated into single strands and stabilized, which looks like a fork. The synthesis of the daughter strands of DNA then can take place over these single-stranded regions of the parent DNA molecule. However, in eukaryotes, DNA sequences are inaccessible and remain inactive, due to compaction of the DNA sequences in chromatin assembly. The fundamental unit of chromatin is nucleosome which has the same architecture for all eukaryotes. They contain 200 base pairs of DNA wrapping around octamers of basic proteins called histones in a bead-like fashion.

BOX 2. Transcription

The second step of the central dogma is the synthesis of the intermediate RNA molecules from the duplex DNA. The process is known as transcription and is carried out by the enzyme called DNA-dependent RNA polymerase. This enzyme is a multi-subunit protein capable of correctly reading the DNA base sequence and transcribing them to RNA message. The process of transcription also involves separation of the duplex DNA into single strands, where upon only one strand is read by the enzyme. There are three kinds of RNA molecules: messenger RNA, ribosomal RNA and transfer RNA. Out of the three, only the messenger RNA is directly read for protein synthesis, whereas, ribosomal RNA and transfer RNA help in the assembly of the protein synthesizing machinery.

BOX 3. Translation

The last step of the central dogma is known as translation. In this step, DNA base sequence in the form of messenger RNA is translated into amino acids which are subsequently joined together to form a functional protein. The process of translation proceeds in a quantized fashion, i.e. at a time three bases on the messenger RNA are read (genetic code) and translated into a single amino acid. The synthesis of proteins or translation takes place over an organelle called the ribosome. From prokaryotes like *Escherichia coli*, the ribosome is called 70S, defined according to its sedimentation value during centrifugation (Svedberg unit). It has two subunits, 50S and 30S, which have two ribosomal RNA molecules and 31 proteins and one RNA and 21 proteins, respectively.

sible to collect synchrotron X-ray data on the crystals in Russia. They eventually moved with the crystals in 1996 to the USA, where beam times were available. They finally solved the structure as a part of a large team led by Harry Noller[9]– (Figure 1).

In the case of the ribosome structure determination, a wealth of data obtained earlier from immuno-electron microscopy of the whole 70S particle was of great help. However, X-ray analysis for the first time showed the functional relay that was involved in numerous contacts at the subunit interface.

Just as immuno-electron microscopy helped in interpreting X-ray data in the case of the ribosome, 2-D crystallography of RNA polymerase oriented on a lipid bilayer, gave the initial models that helped in the total structure determination of this macromolecule. However, single X-ray structure analysis has repeatedly proved that there is no real substitute. It is interesting to understand how the results of several major efforts, to obtain structural information by chemical and biochemical methods evaporate when the near atomic resolution structures are determined. Usually, if the previous data support evidence obtained from X-ray crystallography then such studies are acknowledged; on the other hand, if they do not then they are sacrificed and forgotten! It is indeed a part of history today that not too long ago, there was an institute at Germany totally devoted to work on the ribosome and they produced several hundred papers on its structure by various means other than X-ray crystallography.

Where does one go now from this point? With the background of these structures and more refinements to come very soon at atomic dimensions, it is expected that we will know exactly how the transcription and translation processes occur, mechanistically. Single-molecule studies[12] of RNA polymerase and ribosome movements over respective templates will finally tell us about the detailed working of the nature's tiny machines.

1. Richmond, T. J., Finch, J. T., Rushton B., Rhodes, D. and Klug, A., *Nature*, 1984, **311**, 532–537.
2. Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E. and Moudrianakis, E. N., *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 10148–10152.
3. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond T. J., *Nature*, 1997, **389**, 251–260.
4. Mooney, R. A. and Landick, R., *Cell*, 1999, **98**, 687–690.
5. Zhang, G., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K. and Darst, S. A., *Cell*, 1999, **98**, 811–824.
6. Fu, J., Gnatt, A. L. Bushnell, D. A., Jensen, G. J., Thompson, N. E., Burgess, R. R., David, P. R. and Kornberg, R. D., *Cell*, 1999, **98**, 799–810.
7. Clemons Jr., W. M., May, J. L. C., Wimberly, B. T., McCutcheon, J. P., Cekel, M. S. and Ramakrishnan, V., *Nature*, 1999, **400**, 833–840.
8. Ban, N., Nissen, P., Hansen, J., Cakel, M., Moore, P. B. and Steiz, T. A., *Nature*, 1999, **400**, 841–847.
9. Cate, J. H., Yusupov, M. M., Yusupova, G. Zh., Earnest, T. N. and Noller, H. F., *Science*, 1999, **285**, 2095–2104.
10. Pennisi, E., *Science*, 1999, **285**, 2048–2051.
11. Yonath, A. *et al.*, *J. Mol. Biol.*, 1988, **203**, 831–834.
12. Weiss, S., *Science*, 1999, **283**, 1676–1682.

Dipankar Chatterji and Preethi Chander are in Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India.

Random selections

Molecular motors

'Molecular architecture of the rotary motor in ATP synthase'
D. Stock, A. G. W. Leslie and J. E. Walker
Science, 1999, **286**, 1700

The ATP synthase of mitochondria is a key element in the process by which the energy stored in a transmembrane proton (electrochemical) gradient is converted into chemical energy in the form of adenosine-5′-triphosphate. The complex mechanisms involved are slowly coming to light as X-ray diffraction coaxes the monstrously large, multisubunit, membrane protein, to yield its secrets. These molecules are probably the smallest known rotary motors. This report defines the structure of the *Saccharomyces cervisiae* ATP synthase at 5 Å resolution, pro-