

Solving the Structure of a Histone Acetyltransferase: Pushing the Limits of MAD

R.N. Dutnall, S.T. Tafrov, R. Sternglanz, and V. Ramakrishnan
(U. of Utah School of Medicine)

Chromatin is a complex between the genetic material DNA and some highly charged proteins called histones. It is built up from a repeating units called nucleosomes, which under the electron microscope appear like beads on a string of DNA. Within each nucleosome the DNA is wrapped nearly twice around a core structure composed of two copies each of the core histones H2A, H2B, H3 and H4. Beyond this fundamental level, nucleosomes form complex interactions with each other and with other chromatin components to form chromosomes.

Organizing DNA this way has two major advantages. Firstly it means that the large amount of information contained within the genomes of higher organisms can be packaged into a nucleus. For example, the DNA in the chromosomes of a human genome, if laid out end to end, would measure roughly two metres long but has to be contained within a nucleus that is only about 10-20 μm in diameter. Secondly it is a powerful way of regulating how the information within a genome is expressed.

The first step in gene expression involves the formation of mRNA by a complex machinery of proteins that interact with DNA during the process of transcription. Eventually the mRNA serves as a template for the production of specific proteins. By changing the structure of chromatin so that it is more or less accessible to components of the transcription machinery, cells can regulate which genes are expressed and which are not. Changes in the pattern of genes that are expressed drive the differentiation of different cell types during the development of an organism. After development is complete, maintaining these patterns helps to define the behaviour of the various cell types found within a complex organism. Failures in this complex control can result in developmental disorders, cancer, or other diseases.

One way that cells regulate the structure of chromatin is by modifying specific lysine residues of the

histones by attaching an acetyl group. Each core histone is made up of a globular region that is involved in interaction with the other histones and with DNA, and more flexible, highly charged tails at each end of the globular region. The N-terminal tails in particular contain many positively charged arginine and lysine residues. Many of the lysine residues have been found to be acetylated under different conditions, and it has been known for a long time that the level of acetylation of histones in different regions of chromatin correlates well with the activity of the chromatin in terms of gene expression. For example, genes that are active often are associated with chromatin in which the histones are highly acetylated, while inactive genes are associated with chromatin made up from less acetylated histones. Despite this correlation it still is not known for certain how histone acetylation alters the structure of chromatin. One possibility is that acetylation reduces the positive charge of the histone tails, making them less able to bind DNA where they could interfere with binding of various transcription factors. A second is that it alters the way in which the nucleosomes interact with each other so that chromatin becomes generally more open to the large protein complexes involved in transcription. Finally, it may act as a signal or flag, similar to phosphorylation of proteins during signal transduction, so that it alters the interaction of other proteins with nucleosomes.

Histone acetylation is controlled by specific histone acetyltransferase (HAT) enzymes and histone deacetylase enzymes. Many of the genes that encode these enzymes have been identified within the past few years, which has enabled a more detailed examination of the way in which they function. One difference between HAT enzymes that has emerged from biochemical studies relates to their specificity toward histones. Some have very broad specificity, being able to acetylate lysine residues in any of the four core histones, while others discriminate more highly, showing a marked preference for modifying one

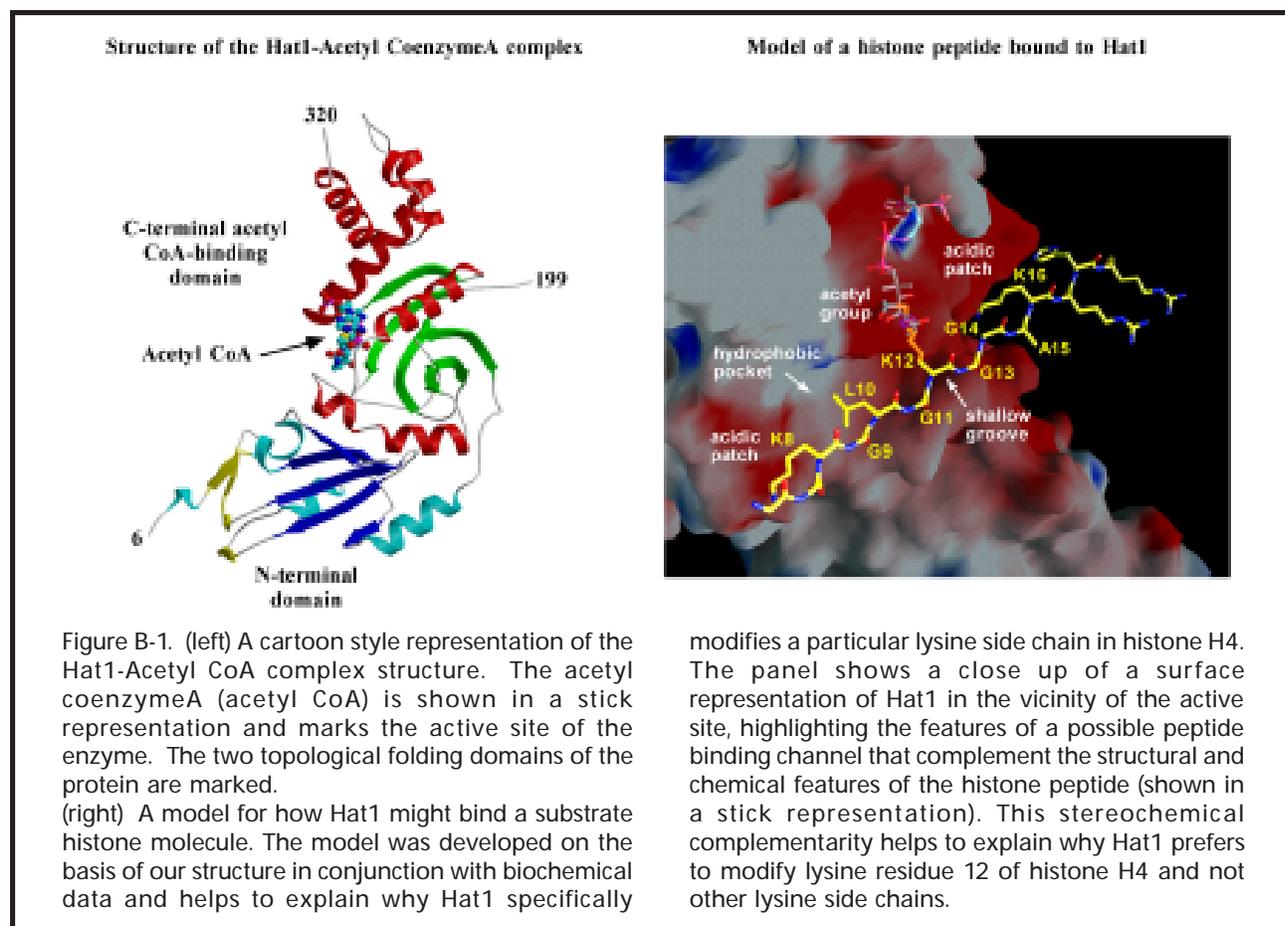
or two lysine residues within one or sometimes two of the histones. Some HATs also can acetylate non-histone substrates such as proteins involved in the transcription machinery itself.

There are examples of chromosome translocations that fuse a HAT gene with another gene. These translocations often are associated with leukemia, indicating that uncontrolled acetylation could be a major factor in promoting cancer formation. Therefore it is important to understand how these enzymes work, in particular how they achieve specificity for a particular target protein. From comparisons of protein sequences it also has been shown that many of the known HATs belong to a large superfamily of enzymes that acetylate a variety of different types of substrate, including antibiotic compounds and factors involved in regulating circadian rhythm. There is therefore more widespread interest in this class of enzymes because of their roles in gene regulation, bacterial antibiotic resistance, and control of day/night cycles.

In collaboration with Rolf Sternglanz from SUNY Stonybrook we set out to determine the structure of the enzyme Hat1, a HAT gene from baker's yeast. Hat1 was the first HAT gene to be identified, both by Sternglanz and separately by Dan Gottschling's group at the

University of Chicago. We initially tried to crystallize the full length 374 residue protein, and eventually ended up with some large crystals, but they took several months to grow and diffracted X-rays only to medium resolution. To fix the problem we used an approach that is quite often adopted by structural biologists. We used proteases to probe the Hat1 protein for flexible regions that might be interfering with the formation of crystals. However we wanted to be careful not to destroy the catalytic activity of the protein, so we combined these proteolytic assays with HAT activity assays to be sure that we retained an active fragment. From this analysis we were able to determine that 54 residues at the C-terminal end of the protein were probably more flexible, and that we could remove them without reducing the activity of the protein. When we made this new protein (residues 1-320) we quickly obtained crystals that diffracted to high resolution ($\sim 2\text{\AA}$ at Beamline X12C of the NSLS) giving us the opportunity to solve the structure.

However there was still one potential obstacle in our way, which was how to solve the phase problem. Two methods are available for determining phases for a new structure. The first is multiple isomorphous replacement (MIR) which involves soaking crystals in compounds of various heavy metals and comparing the diffraction



patterns from these “derivative” crystals to those of “native” untreated crystals to yield phase information. One drawback of this approach is that attaching the heavy metal to the protein must not severely disrupt the crystal lattice, otherwise it is impossible to compare the “derivative” and “native” diffraction patterns. This often means that a large number of heavy metal compounds have to be screened to find a reasonable number of suitable derivatives. An alternative approach is to incorporate an atom into the protein that scatters X-rays anomalously. X-Ray diffraction patterns then are collected using X-rays at different wavelengths, typically two wavelengths around the X-ray absorption edge of the atom and another at a remote wavelength. The information from these data sets collected at different wavelengths can be treated in a fashion analogous to the “native” and “derivative” data sets of MIR to calculate phases. This multiple wavelength anomalous diffraction (MAD) method is becoming an increasingly popular way to solve protein structures. One reason for this is that it requires only a single crystal of the protein and so the problem of non-isomorphism that often makes MIR difficult does not exist. However, it does require that data can be collected at a tuneable X-ray source such as only exists at synchrotron sources like the NSLS.

Our lab has solved several structures using MAD but at first this did not seem to be the best approach to take for Hat1. One factor influencing the success of MAD is being able to incorporate sufficient anomalously scattering atoms into the protein. An ideal method to do this is to take advantage of protein chemistry and prepare protein in which the sulphur atom in each methionine residue is replaced by selenium. This can be done fairly easily by making protein in bacteria and growing the bacteria on a medium that contains selenomethionine instead of methionine. It usually takes a ratio of 1 methionine for every 100 or so protein residues to give enough anomalous signal to successfully determine a MAD structure. However, Hat1 has only 2 methionines for 320 residues, putting us well below the typical 1 in 100 mark. Therefore we made several attempts to prepare derivative crystals to try solving the structure by MIR, but in most cases the crystals either did not diffract as strongly or were clearly non-isomorphous with native crystals. We therefore reconsidered using the MAD approach (and given the type of protein we were working on there were lots of jokes about “MAD HATters”). We went ahead and prepared Selenomethionine-Hat1 crystals but to try to increase the chance of success we attempted a variation on the MAD approach. We prepared crystals of a methyl-mercury derivative of Selenomethionine-Hat1, which although not isomorphous with underivatized protein, did contain more anomalously

scattering atoms. We then planned to collect a four wavelength MAD data set, collecting data sets at the absorption edges of the Selenium and mercury atoms (2 for Se, 1 for Hg) plus a remote wavelength.

We took these crystals to the NSLS in Feb.1998 and collected a four wavelength MAD data set using a crystal of methylmercury-selenomethionine-Hat1 plus a “typical” three wavelength MAD data set using a crystal of plain selenomethionine-Hat1. We collected the data on beamline X12C using a CCD detector which gave excellent sensitivity in measurements of X-ray data. Being able to measure accurately small differences in the X-ray data is important for the MAD technique. As things turned out we were able to solve the structure using the “typical” three wavelength MAD data set (based on Se atoms alone). In part we were fortunate that both methionine residues in Hat1 are buried in the core of the protein and are well ordered in the structure. The second factor in our favour was that we were able to collect high quality data sets using the CCD detector and measure the very small differences produced by the two selenium atoms. This is, to our knowledge, the lowest selenium content (1 per 160 residues) ever used to solve a structure, but we were still able to produce high quality electron density maps, making interpretation much easier. We later used the four wavelength data set to calculate a map but there were no clearly discernable improvements over the map using phases from selenium alone. Another innovation that was available to us was the program SOLVE, which automatically locates heavy atom positions and calculates phases. This enabled us to calculate an electron density map within only three to four hours of completing the data collection. This almost instant feedback gave us a better assessment of the level of success of our experiment allowing us to make better use of our data collection time and return to Utah, sleepy but happy MAD hatters.

The structure of the Hat1-acetyl CoenzymeA complex has given us a wealth of information about how this enzyme, and other related histone acetyltransferases such as Gcn5, function. Using the structure we have been able to propose a model for how Hat1 binds to a particular histone and places a specific lysine residue within the active site for modification. The structure will play a key role in future structural and biochemical studies aimed at understanding this important group of enzymes. ■

REFERENCE

R.N. Dutnall, S.T. Tafrov, R. Sternglanz, and V. Ramakrishnan, “Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily”, *Cell* **94**, 427-38 (1998).

Chemical Differences in Subchondral Osteoarthritic Bone Observed with Synchrotron Infrared Microspectroscopy

L.M. Miller and M.R. Chance

(Department of Physiology and Biophysics, Albert Einstein College of Medicine)

D. Hamerman

(Department of Medicine, Division of Geriatrics, Montefiore Medical Center)

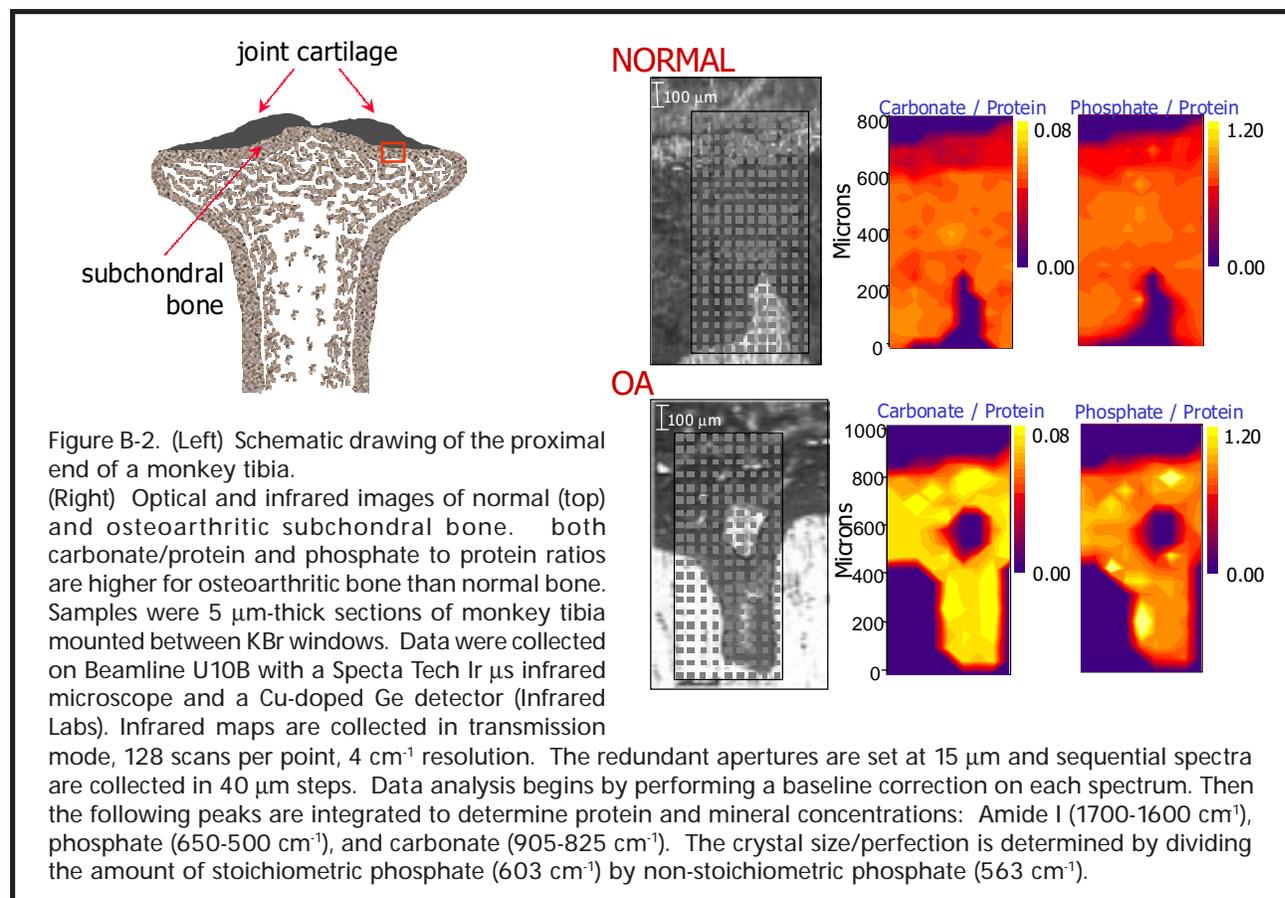
C.S. Carlson

(Department of Veterinary Diagnostic Medicine, College of Veterinary Medicine, University of Minnesota)

Infrared spectroscopy at the NSLS has grown substantially over the last several years as increased use of the existing beamlines has led to construction of new facilities. The use of microscopy to study biological structure is one of the main goals of the new station at U2B, which is operated by the Albert Einstein College of Medicine as an NIH Resource Center. The study of bone structure and function is one of the main projects of the Center, and this article describes some of the progress that has been made so far in understanding bone disease

using synchrotron methods of the NSLS.

Osteoarthritis is the leading cause of disability among people over 65 years old and affects approximately 40 million people in the United States^[1]. It is generally considered a disease of the cartilage. However, it has recently been suggested that the health and integrity of the articular cartilage is dependent upon the mechanical properties of the underlying subchondral bone (**Figure B-2**), where stiffened subchondral bone may lead to the breakdown of articular cartilage^[2].



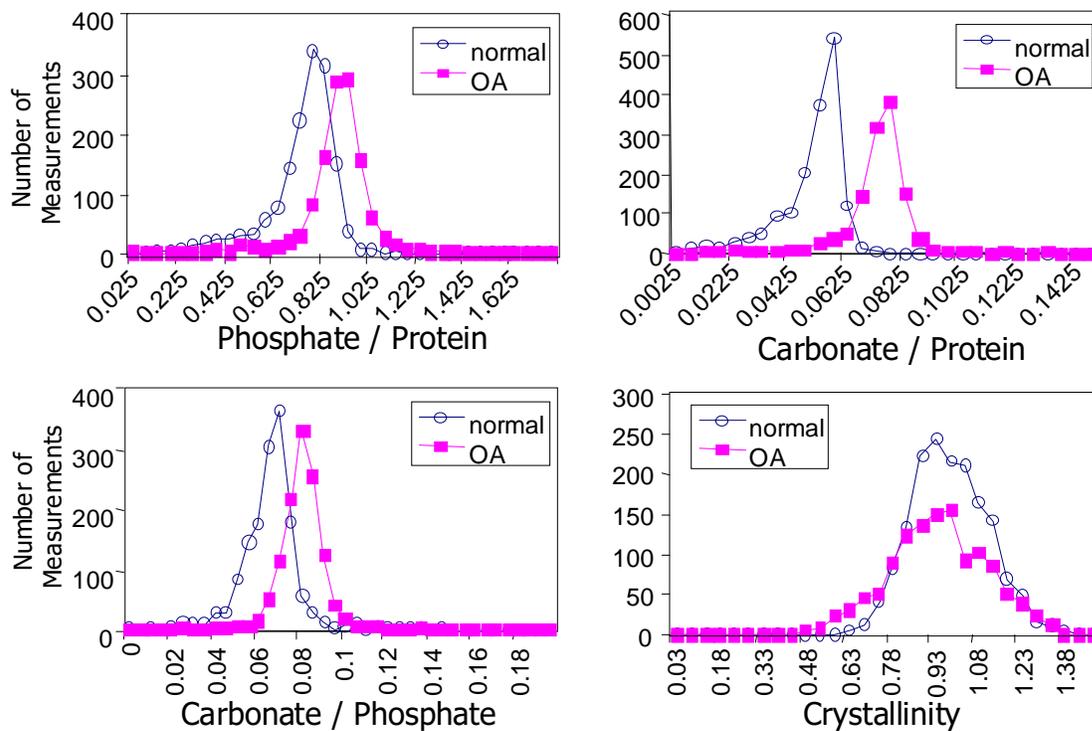


Figure B-3. (Clockwise from top left) Frequency distributions for phosphate/protein, carbonate/protein, crystallinity, and carbonate/phosphate ratios from 5 histologically normal monkey samples (10 infrared maps) as compared to and 2 monkeys with severe osteoarthritis (7 infrared maps). Mineral/protein ratios and carbonate content are significantly higher in osteoarthritic bone than normal bone. Average crystal size/perfection is broadly distributed and not significantly different between normal and osteoarthritic monkeys. Data are expressed as ratios to account for small variations in sample thicknesses.

Osteoarthritic joints display significant thickening of the subchondral bone. To date, the basis for subchondral bone thickening is unknown. But, an extensive study of cynomolgus monkeys has demonstrated that this thickening **precedes** the development of articular lesions^[3]. In addition, there appears to be a threshold subchondral bone thickness, below which articular cartilage lesions are extremely rare. Because of the inherent difficulty in studying the *early* stages of the disease in humans, the use of an appropriate animal model is important. We hope that using the cynomolgus monkey model to examine all ages and stages of the disease will help shed light on the comparable human condition.

The mechanical properties of subchondral bone are dependent upon its chemical structure. Modifications in the chemical composition of bone can affect properties such as its strength, rigidity, density, and flexibility. Thus, we hypothesize that the bone matrix composition and/or bone mineral content and crystallinity are modified *in situ* in a specific manner as a function of subchondral bone thickness.

We are addressing this hypothesis by examining the chemical composition of subchondral bone using synchrotron infrared microspectroscopy. Infrared spectroscopy is an analytical technique that is sensitive to the chemical components in bone; a spectrum of infrared light is absorbed differently depending on the protein (primarily collagen) and mineral (hydroxyapatite) content and composition. More specifically, this technique can be used to determine (1) protein structure and concentration, and (2) mineral concentration, crystallinity, and content (e.g. phosphate, acid phosphate, carbonate). The infrared crystallinity results are well-correlated to hydroxyapatite crystal size and perfection as determined by x-ray powder diffraction. If infrared light is put through a microscope (equipped with special infrared optics), then infrared spectra can be collected on micron-sized regions of bone *in situ* and compared to visible images of the same region. Thus, we can examine the chemical composition of subchondral bone as a function of (a) age, (b) position in the subchondral plate, and (c) thickness of the subchondral plate and compare

these results as a function of severity of the disease. Inherently, the long wavelengths of infrared light limit the spatial resolution achievable with this technique. Our preliminary results show that substantial changes in chemical composition occur within 20 μm of new bone growth^[4, 5]. This type of spatial resolution can only be achieved with a synchrotron infrared source. The National Synchrotron Light Source at Brookhaven National Laboratory provides a source of broadband infrared light that is 1000 times brighter than a conventional infrared source. This permitting rapid data collection at the highest spatial resolution possible (3-5 μm).^[6]

Using synchrotron infrared microspectroscopy, we have compared the chemical composition of subchondral bone *in situ* in histologically normal versus osteoarthritic monkey tibias. Specifically, we have analyzed the protein, phosphate, and carbonate concentrations and bone crystallinity throughout the subchondral bone as a function of disease severity. **Figure B-2** illustrates mineral/protein infrared maps from the subchondral bone of normal (top) and osteoarthritic (bottom) monkeys. As can be seen, both the phosphate/protein and carbonate/protein ratios are higher in the osteoarthritic bone. The results from this single sample were confirmed by examining five radiographically normal monkeys and two monkeys with severe osteoarthritis (as defined by considerable fibrillation and clefting of the joint cartilage.)

This analysis totals approximately 3000 infrared spectra for each disease state. Frequency distributions were calculated for normal bone versus osteoarthritic bone and can be seen in **Figure B-3**. Variance analysis (ANOVA) demonstrates that these both the phosphate/protein and carbonate/protein distributions are significantly different for normal versus osteoarthritic bone, where the osteoarthritic bone is more mineralized than the normal bone.

Analysis of the carbonate/phosphate ratio (**Figure B-3**) also demonstrates that the carbonate content is significantly higher in osteoarthritic bone than normal bone. High carbonate content often correlates with low average crystal size, since carbonate substitution into the hydroxyapatite lattice is usually limited to the surface of hydroxyapatite crystals. However, we find that the average crystal size/perfection are broadly distributed in both normal and osteoarthritic bone, but are not significantly different between the normal and disease states.

These results support our hypothesis that chemical composition changes occur in the subchondral bone in osteoarthritis. Since subchondral bone composition affects both the mechanical and physiological properties of bone, differences between histologically normal versus osteoarthritic monkeys may provide insight into the chemical basis for subchondral bone thickening and articular cartilage breakdown in osteoarthritis. ■

REFERENCES

- [1] Control, C. f. D. *Morb. Mortal. Wkly. Rep.* **43**, 433-438 (1990).
- [2] E.L. Radin, and R.M. Rose, *Clin. Orthop.* **213**, 34-40 (1986).
- [3] C.S. Carlson, R.F. Loeser, C.B. Purser, J.F. Gardin, and C.P. Jerome, *J. Bone Miner. Res.* **11**, 1209-1217 (1996).
- [4] L.M. Miller, C.S. Carlson, G.L. Carr, and M.R. Chance, *Cellular and Molecular Biology* **44**, 117-127 (1998).
- [5] L.M. Miller, C.S. Carlson, G.L. Carr, G.P. Williams, and M.R. Chance, *SPIE* **3153**, 141-148 (1997).
- [6] G.L. Carr, J.A. Reffner, and G.P. Williams, *Rev. Sci. Instr.* **66**, 1490-1492 (1995).

Crystal Structure of a Hepatitis Delta Virus Ribozyme Reveals a Buried Active Site Cleft Reminiscent Protein Enzymes

A.R. Ferre-D'Amare and J.A. Doudna
(Yale University)

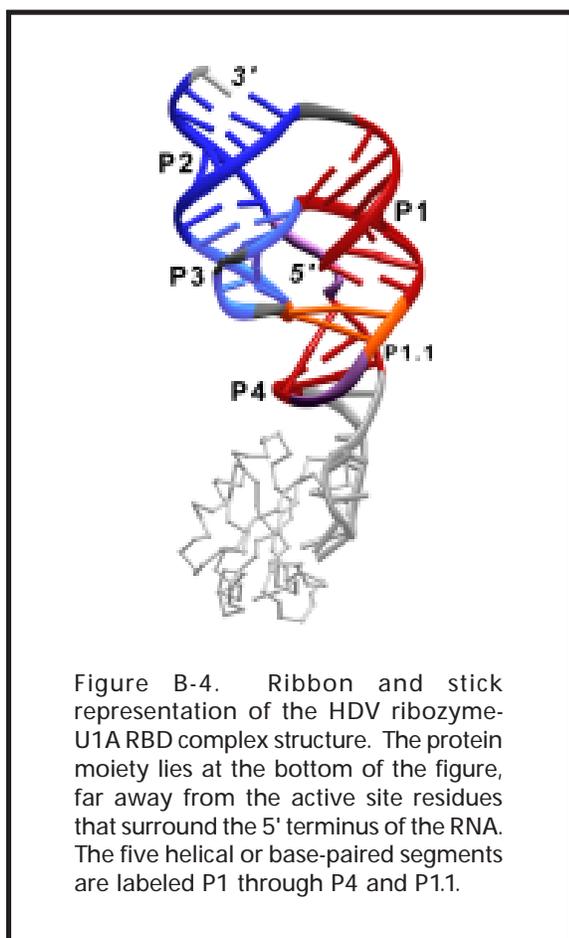
The hepatitis delta virus (HDV) ribozyme is the only catalytic RNA known to be required for the viability of a human pathogen. This self-cleaving RNA is remarkable for its high stability, its limited requirement for divalent cations, and its fast rate. The HDV ribozyme is fully active in conditions, such as 8 molar urea, under which most RNAs and proteins are denatured and strongly inactivated. Because of the negatively charged phosphodiester backbone of RNA, most catalytic RNAs depend on divalent cations to fold into their active conformations. The HDV ribozyme is the fastest natural self-cleaving RNA that has been characterized. Even in very low concentrations of a variety of divalent cations, it

can cleave itself at $\sim 1/\text{second}$, which is two orders of magnitude faster than a typical self-cleaving RNA, and approaches the rate of some protein enzymes^[1].

In order to gain structural insight into the unique catalytic prowess of this RNA, we had to obtain well-diffracting crystals. Good crystals of tightly folded RNAs are notoriously difficult to obtain, presumably because the surface of these molecules is dominated by a periodic array of phosphates. This might allow neighboring molecules to pack subtly out of register in the crystal, resulting in poor order. To overcome this problem, we engineered a solvent-exposed, functionally and structurally dispensable stem-loop of the HDV ribozyme into a high-affinity binding site for the small, basic RNA-binding domain (RBD) of the U1A protein. We expected the complex between the 72 nucleotide RNA and the U1A protein RBD to produce better crystals than the “naked” RNA, because the complex not only has a negatively and positively charged ends, but has the wide variety of surface functional groups of the protein available for making crystal contacts.

Our best crystals of an HDV ribozyme-U1A RBD complex diffracted X-rays to beyond 2.2 Å resolution. We solved the structure by crystallizing RNA-protein complexes in which the methionines of the protein moiety had been biosynthetically substituted with selenomethionine. We then carried out a multiwavelength anomalous diffraction (MAD) experiment around the selenium absorption edge at NSLS Beamline X4A, with help from Drs. C. Ogata and D. Cook. The phases thus obtained produced a good electron density map, and the resulting atomic model has now been refined against diffraction data extending to 2.3 Å resolution^[2]. Use of the RNA-binding protein “crystallization module” might be a powerful general technique, not only for obtaining high quality crystals of large RNAs, but also for determining their structures by MAD.

Our crystal structure revealed that the HDV ribozyme adopts a compact, highly convoluted fold in which five helical segments are linked by five strand crossovers; a connectivity that can be described as a “nested double pseudoknot” (**Figure B-4**). This fold buries the



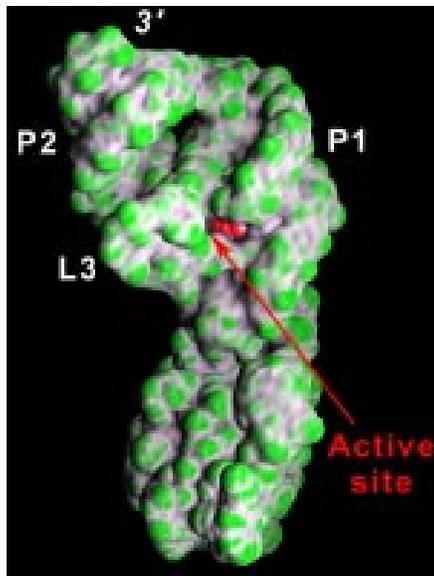


Figure B-5. Solvent accessible surface representation approximately in the same orientation as Figure B-4. The atoms of the 5'-most ribose are shown as red spheres. The active site lies in the cleft behind this sugar.

active site at the bottom of a deep, solvent-inaccessible cleft, in which various base and backbone functional groups appear to be poised to achieve catalysis (**Figure B-5**). Interpretation of previously published biochemical results from other workers in light of the structure suggests that the complex fold of the HDV ribozyme is at least in part responsible for the remarkable stability of this RNA. No tightly bound metal ions were found in the electron density maps. This is consistent with published biochemical data, and indicates that nature can achieve intimate packing of negatively charged RNA without recourse to tight metal ion chelation^[3].

The HDV ribozyme is the first catalytic RNA whose ground state structure demonstrates an active site pre-organized for catalysis. The three-dimensional structure of this remarkable ribozyme constitutes the starting point for further biochemical, biophysical and pharmacological investigations into its mode of action both *in vitro* and in the life cycle of the virus in the human liver. ■

REFERENCES

- [1] M.D. Been and G.S. Wickham. "Self-cleaving ribozymes of hepatitis delta virus RNA", *Eur. J. Biochem* **247**, 741-753 (1997).
- [2] A.R. Ferre-D'Amare, K. Zhou, and J.A. Doudna. "Crystal structure of a hepatitis delta virus ribozyme", *Nature* **395**, 567-574 (1998).
- [3] A.R. Ferre-D'Amare and J.A. Doudna. "RNA folds: insights from recent crystal structures", *Annu. Rev. Biophys. Biomolec. Struct.* In press.

How Sos Activates Ras: The Crystal Structure of the Ras-Sos Complex at 2.8 Å Resolution

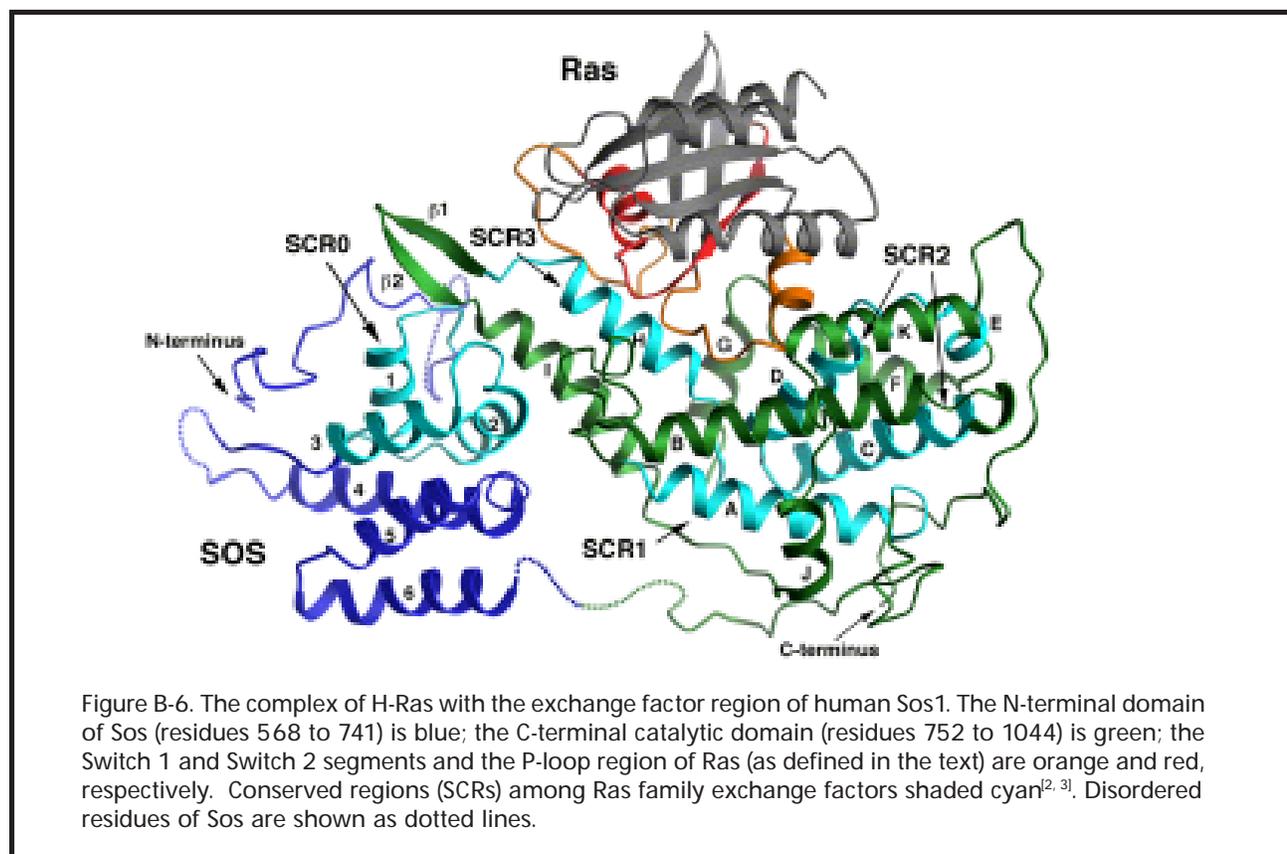
P. A. Boriack-Sjodin, S.M. Margarit, D. Bar Sagi, and J. Kuriyan
(Rockefeller University)

The Ras proteins are structurally conserved guanine nucleotide binding proteins that are involved in intracellular pathways controlling cell proliferation and differentiation^[1,2]. When bound to GTP, Ras is “on” and is able to transmit the cellular signal initiated by ligand binding to receptor tyrosine kinases. When bound to GDP, Ras is “off” and is unable to propagate the cellular signal. The intrinsic rates of GTP hydrolysis and nucleotide dissociation from Ras are low, therefore, two classes of proteins regulate the nucleotide-bound state of Ras: GTPase activating proteins catalyze the hydrolysis of GTP to GDP, turning Ras “off”, while guanine nucleotide exchange factors facilitate the exchange of GDP for GTP, switching Ras “on”. Although many

biochemical studies have been done on Ras exchange factors, the mechanism for exchange was largely unknown due to the lack of structural information.

The X-ray crystal structure of Ras complexed to the exchange factor Son of Sevenless (Sos) was recently determined (**Figure B-6**) using multiple isomorphous replacement methods. Data used to solve the structure were collected at beamline X25 at the National Synchrotron Light Source at Brookhaven National Laboratories using the Brandeis B4 2x2 module CCD detector. The structure provides the first glimpse of how Ras exchange factors function to (1) remove nucleotide from Ras and (2) allow nucleotide to rebind to Ras.

The structure reveals that Sos interacts extensively



with Ras to exclude nucleotide and magnesium ion and stabilizes Ras in the nucleotide-free state. Residues 25-40 of Ras (the Switch 1 region), which normally form part of the nucleotide binding pocket, are completely removed from the nucleotide binding site. Conformational changes in residues 57-75 of Ras (the Switch 2 region) result the occlusion of the magnesium binding site of Ras by Ala 59 (**Figure B-7**). Sos also introduces a hydrophobic residue (Leu 938) into the magnesium binding pocket, further changing the electrostatic environment to one unsuitable for a divalent ion. Residues which normally

coordinate the nucleotide (Lys 16 and Gly 60) are in alternate hydrogen bonding interactions and Sos introduces an acidic residue (Glu 942) into the α -phosphate binding site, further disrupting the binding site of the nucleotide. Importantly, the complex leaves the guanine base and ribose binding sites unobstructed, thus providing a mechanism for the rebinding of nucleotides to Ras to complete the exchange reaction.

This structure was recently published in *Nature*: P. A. Boriack-Sjodin, S.M. Margarit, D. Bar Sagi, and J. Kuriyan, "Structural basis for the activation of Ras by Sos", *Nature* **394**, 337-343 (1998). ■

REFERENCES

- [1] H.R. Bourne, D.A. Sanders, and F. McCormick, "The GTPase superfamily: conserved structure and molecular mechanism", *Nature* **349**, 117-127 (1991).
- [2] M.S. Boguski and F. McCormick, "Proteins regulating Ras and its relatives", *Nature* **366**, 643-654 (1993).
- [3] C.-C. Lai, M. Boguski, D. Broek, and S. Powers, "Influence of guanine nucleotides on complex formation between Ras and Cdc25 proteins", *Molec. Cell. Biol.* **13**, 1345-1352 (1993).
- [4] E.F. Pai, *et al.* "Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis", *EMBO J.* **9**, 2351-2359 (1990).

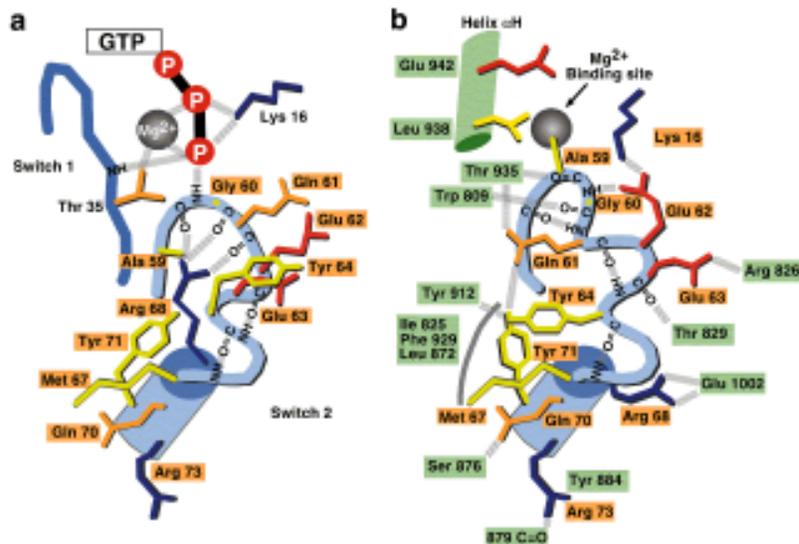


Figure B-7. A schematic of the differences in the Switch 2 regions of a) Ras-GTP analogue (5P21)^[4] and b) Ras-Sos. Selected polar interactions are shown as dashed lines, hydrophobic interactions are shown as solid arcs.

Structure of the HIV gp120 Envelope Glycoprotein in Complex with CD4 and a Neutralizing Human Antibody

P. D. Kwong (Columbia University)

R. Wyatt and J. Sodroski (Dana-Farber Cancer Institute)

W.A. Hendrickson (Columbia U. and the Howard Hughes Medical Institute)

As the 20th century draws to a close, the human immunodeficiency virus (HIV) continues to spread, having recently surpassed both malaria and tuberculosis as the leading infectious disease killer. Since 1983, when HIV was first discovered, the UNAIDS organization estimates that over 40 million people have been infected of which almost 15 million have died. Despite global awareness and recent therapeutic advances, the only hope for containment in many countries lies in the development of a vaccine. In spite of enormous effort, all such vaccine candidates have thus far met with failure.

HIV is an enveloped virus, which establishes a persistent infection in its human hosts. Although neutralizing antibodies are generated, HIV eludes the immune system. This ability of HIV to evade the immune system is the basis of both the well noted vaccine failures as well as of the ability of HIV to establish a persistent infection. Over a period of years, the constant battle with the virus impairs the immune system enough to allow opportunistic infections to kill the human host.

The HIV lipid envelope, which is derived from the membrane of the host cell, hides most of the infectious virus particle from immune surveillance. Only two viral proteins protrude beyond the protective lipid bilayer, the exterior envelope glycoprotein, gp120, and its membrane-spanning partner, gp41. Of these, gp120 appears to occlude much of the gp41 ectodomain and is the target of virtually all neutralizing antibodies.

The exposure of gp120 on the virion surface is a functional consequence of its central role in virus-cell entry. The gp120 glycoprotein binds sequentially to the CD4 glycoprotein and a chemokine receptor on the surface of specific cells of the immune system, and then in concert with gp41 orchestrates a fusion of the viral and cellular membranes. This fusion is critical for the introduction of HIV's genetic material into the target cell.

Because of the pivotal role of gp120 in receptor binding and interactions with neutralizing antibodies, we undertook an investigation of its atomic structure.

The size of gp120 as well as its extensive glycosylation and expression in eukaryotic cells limited

the choice of high resolution structural techniques to X-ray crystallography. X-ray crystallography in turn depends on obtaining well ordered crystals of the molecule of interest. For gp120, this proved to be extraordinarily difficult.

It was apparent from functional and sequence analyses that many of the same mechanisms that HIV uses to avoid specific protein:protein contacts in eluding the immune system, would also stymie attempts at forcing gp120 into the specific protein:protein lattice interactions required to produce a well ordered crystal.

To surmount these difficulties, we used a procedure, which we term variational crystallization, that focuses on protein modification to enhance the overall probability of producing crystals. Variational crystallization takes advantage of the ability of molecular biology to provide almost limitless variation in the protein of interest. To guide our efforts, we also derived the theoretical underpinnings for an approximate probability theory, based on the comparison of crystallization probabilities before and after a modifying procedure^[1].

This combination of probability analysis and variational crystallization produced very small crystals from four different gp120 crystallization variants, in six different crystallization conditions. Because the larger crystals necessary for high resolution structural analysis are often difficult and time-consuming to grow, we used the NSLS beamline X4A, to shortcut the process. The intense X4A x-rays, several orders of magnitude stronger than standard laboratory sources, enabled us to directly assay the suitability of the very small initial crystals for structural analysis. Of the six, only one proved sufficiently well-ordered for further study. This crystallization variant was composed of a ternary complex of core gp120 (gp120 with deletions at the N and C termini as well as at two internal variable loops), the N terminal two domains of CD4, and the antigen binding fragment of the neutralizing human antibody, 17b.

Small needle-shaped crystals, 5-10 μm in cross-section, could readily be grown. These were used to screen for appropriate cryocrystallography conditions and heavy atom derivatives. Despite several months of concentrated

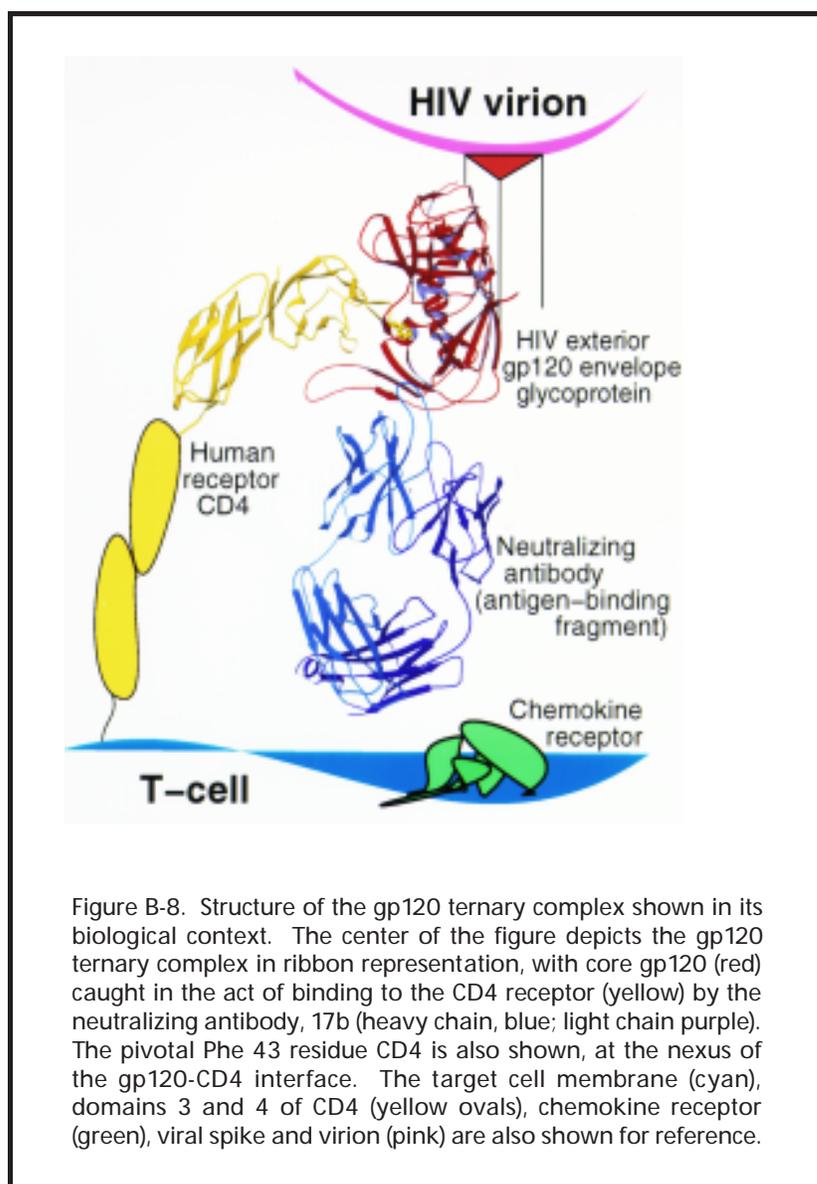
effort, only five larger crystals, with cross-sections of $\sim 50 \mu\text{M}$, could be grown. Nonetheless, use of the NSLS X4A beamline made possible the collection of a native and three derivative data sets. The structure was solved to 2.5 \AA , a resolution sufficient to decipher the gp120 atomic structure (Figure B-8).

Analysis of this structure revealed the detailed features of a cavity-laden CD4-gp120 interface, highlighted by an extraordinary 150 \AA^3 cavity which reached from the pivotal Phe 43 residue of CD4 into the heart of gp120. Also revealed were a conserved binding site for the chemokine receptor, evidence for a conformational change upon CD4 binding, and the nature of a CD4-induced antibody epitope^[2]. In addition, epitope maps in conjunction with the crystal analysis defined the antigenic structure of gp120 and specific mechanisms of immune evasion^[3]. These included multiply redundant mechanisms of steric hindrance and conformational occlusion as well as a completely unanticipated, immunologically “silent” face masked by carbohydrate.

Whereas previous attempts to devise an HIV vaccine relied on indirect hints and clues, our results visualize gp120 at atomic resolution. Understanding the structural basis for the remarkable ability of HIV to elude

the immune response should assist in the design of a vaccine. Our results thus provide not only a framework

for understanding the biology of HIV entry into cells, but also a guide in attempts to intervene. ■



REFERENCES

- [1] P.D. Kwong, R. Wyatt, E. Desjardins, J. Robinson, J.S. Culp, B.D. Hellmig, R.W. Sweet, J. Sodroski, and W.A. Hendrickson, “Probability analysis of variational crystallization and its application to gp120, the exterior envelope glycoprotein of type 1 human immunodeficiency virus (HIV-1)”, *J. Biol. Chem.* in press.
- [2] P.D. Kwong, R. Wyatt, J. Robinson, R.W. Sweet, J. Sodroski, and W.A. Hendrickson, “Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody”, *Nature* **393**, 648-650 (1998).
- [3] R. Wyatt, P.D. Kwong, E. Desjardins, R.W. Sweet, J. Robinson, W.A. Hendrickson, and J. Sodroski, “The antigenic structure of the HIV gp120 envelope glycoprotein”, *Nature* **393**, 705-711 (1998).

Structural Basis for the Interaction of Ras with RalGDS

L. Huang, F. Hofer, G.S. Martin, and S.-H. Kim
(U. of California at Berkeley)

Ras is central within a web of signaling circuits, thus understanding the basis for Ras control is key to unraveling the principles of signal transduction. Indeed, *ras* is the most prevalent of all the oncogenes isolated from human tumors. Mutationally activated Ras plays a role in 90% of pancreatic cancers, 50% of colon cancers, and 25% of lung cancers.

Ras and a large family of related GTPases function as molecular switches that control a variety of signaling pathways that regulate cell growth, motility, intracellular transport, and differentiation. Interaction of the GTPase with its effectors is dependent on binding of GTP. Binding of GTP to human Ras results in major conformational changes in two highly conserved regions: switch I and switch II.

The Ras protein signals to a number of distinct pathways by interacting with diverse downstream effectors. Among the effectors of Ras are the Raf kinase and RalGDS, a guanine nucleotide dissociation stimulator specific for Ral (Ras like GTPase). Ral may regulate Cdc42, which plays a role in the actin cytoskeleton organization. Ral also signals to Phospholipase D pathway. Despite the absence of significant sequence similarities, both effectors bind directly to Ras, but with different specificities. The solution of the Ras/effector complex would lay a foundation for understanding Ras signaling, and thus the role of malfunctioning Ras in oncogenesis.

The X-ray data taken at the BNL beam line X12B and X25 enabled us to solve the structure of the complex between Ras(E31K) and the Ras-interacting domain (RID) of RalGDS in the presence of bound non-hydrolyzable GTP to a resolution of 2.1 Å^[1]. The small crystals of the complex diffracts only weakly at home source, which is not sufficient to determine the precise side chain interactions in the complex. Access to the high beam intensity of beam line X25 was mandatory in order

to determine its structure to such a high resolution.

This structure reveals that the β -sheet of the RID joins the switch I region of Ras to form an extended anti-parallel β -sheet with a topology similar to that found in the Ras homologue Rap-Raf complex. In addition, the overall structure of RalGDS-RID is similar to that of the Ras-binding domain (RBD) of Raf and ubiquitin. However, the side chain interactions at the joining junctions of the two interacting systems, the electrostatic complementarity at the interface, and the relative orientation of the two binding domains are distinctly different. Furthermore, in the case of the Ras-RID complex a second RID molecule also interacts with a different part of the Ras molecule, the switch II region. These findings account for the cross-talk between Ras and Ral pathways. Additionally, comparison with the Rap1A/Raf-RBD structure enabled us, for the first time, to explain the multiplicity and specificity of Ras signaling at an atomic level. It also helps to distill the common structural motif underlying protein-protein interaction in the signal transduction process.

The mechanism by which RalGDS is activated by Ras is implied by the observation that the interaction with Ras perturbs the structure of RalGDS-RID, but not Ras itself. The free RalGDS-RID was solved by the MAD method with data taken at beam line X4A using a selenomethionine mutant protein crystal^[2]. This conformational change might activate the catalytic domain of RalGDS structurally, which in turn would activate Ral.

The structural characterization of the Ras/effector interface will assist in the rational design of small drug molecules capable of disrupting this interface, thereby halting the oncogenic effect of the constitutively activated Ras. Furthermore, the discovery of this interaction motif will aid in recognizing other pairs of signal transduction proteins that may interact via these mechanisms. ■

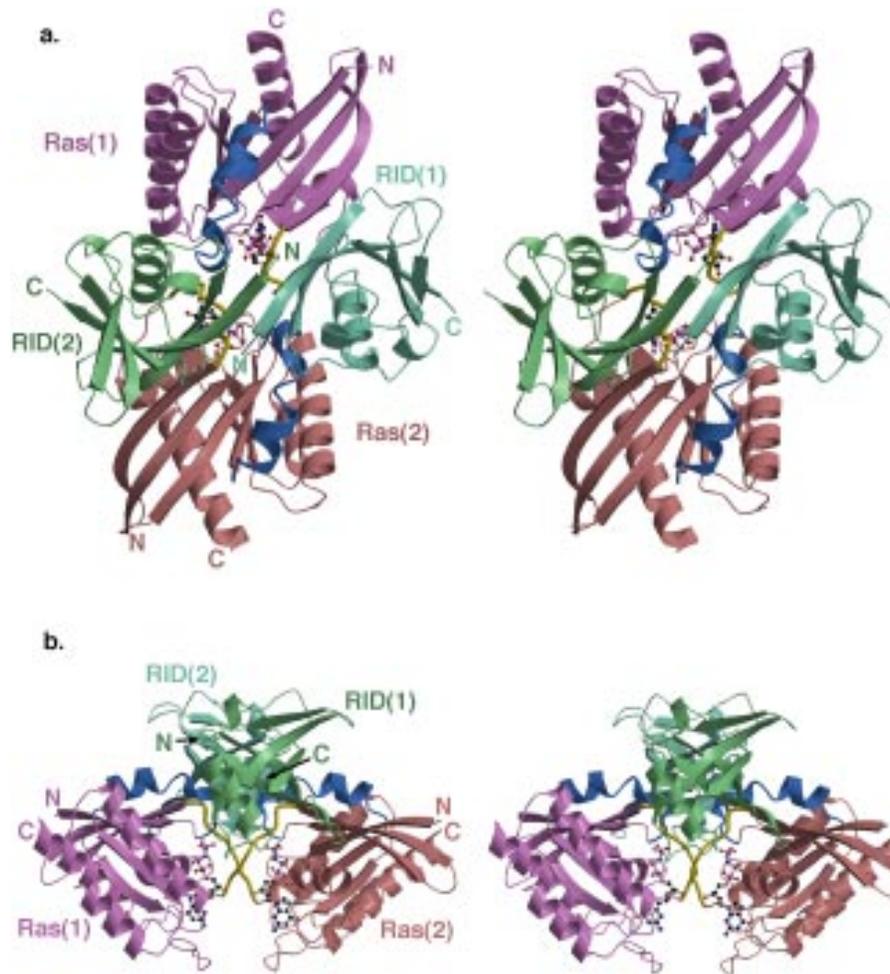


Figure B-9. Ribbon representation of the complex between RaIGDS-RID and Ras(E31K).

(a) Front view of the complex;

(b) a different view of the complex.

Ras molecules are in purple (Ras1) and brown (Ras2), and RID molecules are in cyan (RID1) and green (RID2). The switch I region (residues 30-37) is colored yellow, while the switch II region (residues 60-76) is colored blue. Non-hydrolyzable GTP (GMPPNP) and magnesium are drawn in ball-and-stick representation.

REFERENCES

- [1] L. Huang, F. Hofer, G.S. Martin, and S.-H. Kim, *Nature Struct. Biol.* **5**, 422-426 (1998).
- [2] L. Huang, X. Weng, F. Hofer, G.S. Martin, and S.-H. Kim, *Nature Struct. Biol.* **4**, 609-615 (1997).

Structure of the Large Ribosomal Subunit from *H. marismortui* at 5.5 Å Resolution

B. Nenad, P. Nissen, P.B. Moore and T.A. Steitz
(Yale University)

The ribosome plays a pivotal role in the process of protein synthesis. Ribosomes mediate the interaction between mRNA and tRNA, and catalyze peptide-bond formation. Protein synthesis is one of the most studied cellular processes. Intense research over the past 30 years led to many important discoveries regarding sequences of ribosomal components, folding of the rRNAs, spatial proximity of ribosomal proteins, and overall shape of the ribosome^[1, 2, 3, 4]. Structural information about the nature of protein-nucleic acid interactions in the ribosomes and the structure of the entire organelle is a prerequisite for a mechanistic analysis. Undoubtedly this structure also will reveal fundamental architectural principles that govern the assembly and maintain the stability of all ribonucleoproteins. The progress on the research of ribosomes is reviewed periodically in book form^[5, 6, 7, 8]. The general interest has shifted from investigations of the biochemical properties of the ribosomes to high-resolution studies of their structure and function.

Prokaryotic ribosomes (70S) are composed of two subunits and have a molecular weight of roughly 2.5×10^6 . Each of the subunits is a complex of ribosomal RNA (rRNA) and proteins. The small subunit (30S) contains a single 16S rRNA molecule and 21 different proteins, while the large subunit (50S) contains two rRNA molecules and 32 different proteins^[9]. The 50S subunit has a molecular weight of about 1.5×10^6 Daltons and catalyzes peptide bond formation, possibly functioning as a ribozyme^[10]. Eukaryotic ribosomes are similar structurally, but they are significantly bigger and more complicated^[11]. Although the research on which we are embarked focuses on the large subunit from an archebacterial ribosome, there is no reason to believe that bacterial and eukaryotic ribosomes are fundamentally different^[12], and many of the conclusions are likely to be transferable.

We have generated X-ray crystallographic electron density maps of the large ribosomal subunit from *Haloarcula marismortui* at various resolutions up to 5.5 Å, having measured data from crystals that diffract to 3 Å resolution. Initial steps in the solving of this structure demanded a marriage of two techniques, electron microscopy and X-ray diffraction. We were able to place a 20 Å resolution, three-dimensional model of the 50S

subunit determined by EM methods into the crystal unit cell. To accomplish this required that we had measured the lowest-angle X-ray diffraction data from the crystals, in the 80-20 Å range, which often are ignored by crystallographers. This provided initial phases for low resolution reflections. Using these EM-derived phases, we were able to use difference-Fourier electron-density maps to locate the high-occupancy sites in heavy atom-derivatized crystals. This allowed us to initiate the process of heavy-atom refinement, leading to our first interpretable electron density map of the 50S subunit. The map at 9 Å resolution revealed long, continuous, but branched features whose shape, diameter and right-handed twist are consistent with segments of double helical RNA that crisscross the subunit^[13].

High-resolution structure determination of the large ribosomal subunit will be initiated using crystallographically determined phases calculated at 5.5 Å resolution by multiple isomorphous replacement and anomalous scattering (MIRAS) methods. Electron-density maps calculated using these phases begin to unveil the molecular architecture that underlies this large, ribonucleoprotein complex. First glimpses of the 7.0 Å resolution map show ribbons of electron density corresponding to the RNA backbone, punctuated at regular intervals by bumps arising from the phosphates (**Figure B-10**). Some of the proteins can be distinguished clearly from RNA at this resolution. Thus, we believe that a critical step has been taken on the path leading to an electron-density map of the large ribosomal subunit that can be interpreted in atomic detail.

X-ray crystallographic data collection for this project represents a significant technical challenge which, it appears, can be met by the facilities at NSLS. All the data collection to date has been done at NSLS beamlines X12C, X12B, and X25. For this purpose, numerous adjustments and modifications have been provided by the helpful staff at these beamlines. Initial heavy-atom data were taken at beamline X12C with the Brandeis four-module CCD-based detector. Using a MAR300 imaging-plate detector on beam line X12B we have been able to collect excellent 3.8 Å resolution data. We hope that in the future, when some additional modifications will be made, it will be possible to collect data to the diffraction limit of the crystals. ■

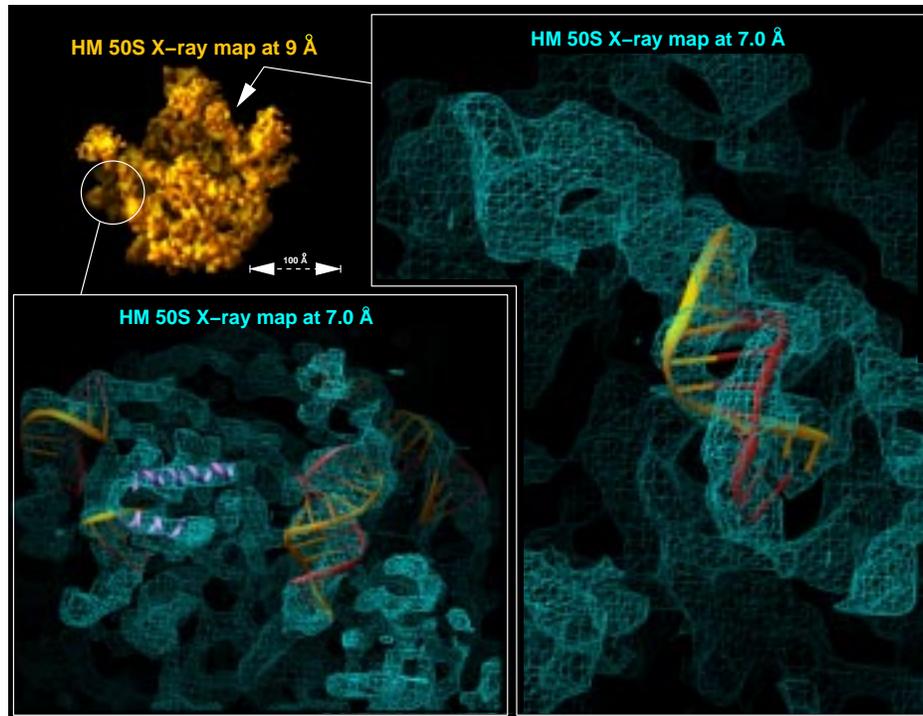


Figure B-10: Results from the study of the 50S ribosomal subunit from *Haloarcula marismortui*. Preliminary electron-density maps at 7Å resolution are shown and compared to the previously published 9Å resolution map^[13].

REFERENCES

- [1] H.F. Noller, *Annu. Rev. Biochem.* **53**, 119 (1984).
- [2] J.A. Lake, *Annu. Rev. Biochem.* **54**, 507 (1985).
- [3] A. Yonath, (1992). *Ann. Rev. Biophys. Biomol. Struct.* **21**, 77 (1992). Z.H. Zhou, S.J. Macnab, J. Jakana, L.R. Scott, W. Chiu, and F.J. Rixon, *Proc. Natl. Acad. Sci. USA* 95:2778 (1998).
- [4] J. Frank, *Curr. Opin. Struct. Biol.* **7(2)**, 266 (1997).
- [5] B. Hardesty, and G. Kramer, *Structure, Function, and Genetics of Ribosomes*, Springer-Verlag, New York (1986).
- [6] W.E. Hill, A. Dahlberg, R.A. Garrett, P.B. Moore, D. Schlessinger, and J.R. Warner, *The Ribosome. Structure, Function, and Evolution*. American Society for Microbiology: Washington, D.C. (1990).
- [7] K.H. Nierhaus, F. Franceschi, A.R. Subramanian, V.A. Erdmann, and B. Wittmann-Liebold, *The Translational Apparatus*. Plenum Press, New York (1993).
- [8] A.T. Matheson, J.E. Davies, P.P. Dennis, and W.E. Hill, *Frontiers in Translation. Biochem. Cell Biol.* **73**, 739 (1995).
- [9] B. Wittmann-Liebold, *Ribosomal proteins: their structure and evolution*, in *Structure, Function, and Genetics of Ribosomes*, B. Hardesty & G. Kramer, Editors. Springer-Verlag: New York, pp. 326 (1986). I.G. Wool, *Annu. Rev. Biochem.* **48**, 719 (1979).
- [10] H.F. Noller, V. Hoffarth, and L. Zimniak, *Science* **256**, 1416 (1992).
- [11] I.G. Wool, Y. Endo, Y.-L. Chan, and A. Gluck, "Structure, function and evolution of mammalian ribosomes", in *The Ribosome. Structure, Function and Evolution*, W.E. Hill, A. Dahlberg, R.A. Garrett, P.B. Moore, D. Schlessinger & J.R. Warner, Editors. American Society for Microbiology: Washington, D.C. pp. 203. (1990).
- [12] I.G. Wool, The bifunctional nature of ribosomal proteins and speculations on their origins, in *The Translational Apparatus*, K.H. Nierhaus, F. Franceschi, A.R. Subramanian, V.A. Erdmann & B. Wittmann-Liebold, Editors. Plenum Press, New York. pp. 727 (1993).
- [13] N. Ban, B. Freeborn, P. Nissen, P. Penczek, R.A. Grassucci, R. Sweet, J. Frank, P.B. Moore, P.B., T.A. Steitz, *Cell* **93(7)**, 1105-1115 (1998).