

HOW UBIQUITIN TAKES ITS CUE

Proteins often just cannot get away from ubiquitin, a versatile protein tag that might mark them for destruction, sorting, or enlistment into gene repair or cell division. Now researchers have determined how a ubiquitin-binding component of several proteins, called CUE, recognizes that tag. They solved the crystal structure of CUE and found that it groups in pairs capable of binding single ubiquitins, and that multiple pairs can team up to form a series of pockets for binding multiple ubiquitin molecules.

Researchers originally discovered ubiquitin, named for its nearly identical form in cells from yeast to human, as a tag marking other proteins for destruction. Some-times the tag is a single ubiquitin; other times it is a string of them. They have since learned that these tags help regulate a range of basic cellular processes by attracting different proteins that recognize and bind to ubiquitin. The discovery of new ubiquitin-binding modules brought to light additional functions for ubiquitin in the recent past, but researchers still know little about how such modules work. One mystery is how they can tell a single ubiquitin molecule from a string of them. Hoping to shed light on that ability, a group from the National Institutes of Health and the Mayo Clinic crystallized and solved the structure of a yeast protein's CUE domain, which recognizes single ubiquitin tags.

Based on x-ray diffraction data obtained at the SBC-CAT 19-ID beamline and the SER-CAT 22-ID beamline at the APS, the group found that CUE has a short, rod-like structure consisting of three helices, very similar to another ubiquitin-binding domain called UBA. They were surprised to find that CUE domains pair up to bind ubiquitin. Individual CUE molecules, or monomers, join together by swapping one helix for the identical one from their partner. The researchers found that such a pair, or dimer, undergoes a dramatic structural shift upon binding, flexing into a basket shape that cups the protein tag (Fig. 1). CUE can also dimerize in real cellular conditions as part of a full-length protein, they found.

To understand why the dimer is responsible for binding ubiquitin, the group constructed mutant proteins whose key binding amino acids were swapped for amino acids that interfere with binding. The first and last helices of the CUE monomer nestle against ubiquitin, whereas the middle helix faces away. In the dimer, however, the middle helix of each monomer snuggles closer to ubiquitin. The group mutated either the middle helix or the flanking helices of the dimer and found that both kinds of mutation made CUE much less sticky to ubiquitin, indicating that the middle and flanking helices both take part in binding. Because the monomer can bind only with the flanking helices, the group conclude that this interaction is not enough to cause tight binding.

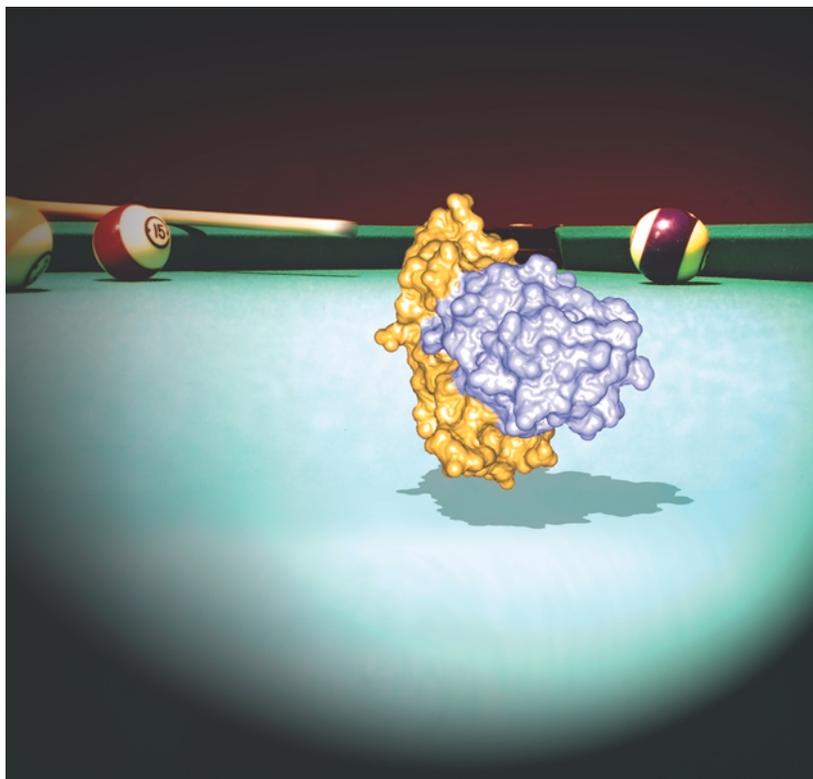


Fig. 1. The CUE domain (gold) joins in pairs to form a basket-like pocket that cups the ubiquitin protein (blue). (Image by G. Prag.)

Surprisingly, each dimer was joined to two ubiquitin molecules in the crystal structure, despite forming a basket designed to hold just one. Because of the way the dimer forms, the same amino acids in one monomer that bind to ubiquitin on the inner, concave surface of the basket are also present in the other monomer, but on the outer, convex surface. If two baskets are side by side, their outer surfaces complement each other and form essentially the same binding pocket as the inner basket does. Two CUE dimers would therefore have three basket-like pockets and be able to bind three ubiquitin molecules. The researchers say that the existence of such a binding pattern could allow domains like CUE to bind individual ubiquitin molecules or strings of them, depending on its state—monomer or dimer. ○

See: G. Prag¹, S. Misra¹, E.A. Jones¹, R. Ghirlando¹, B.A. Davies², B.F. Horazdovsky², and J.H. Hurley¹, "Mechanism of Ubiquitin Recognition by the CUE Domain of Vps9p," *Cell* **113**, 609–620 (30 May 2003).

Author affiliations: ¹National Institutes of Health, ²The Mayo Clinic Cancer Center

This study was supported in part by NIH grant number GM55301 to B.F.H. and a predoctoral fellowship from the HHMI to B.A.D. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

HOW A PROMISCUOUS ENZYME RECOGNIZES COCAINE AND HEROIN

The first three-dimensional structure of a human protein bound to chemicals that mimic cocaine and heroin has been determined by researchers from the University of North Carolina in Chapel Hill and St. Jude Children's Hospital. The enzyme, called hCE1, cleaves apart both narcotics and detoxifies the nerve gas sarin and similar chemical weapons. The structure reveals that hCE1 is capable of acting on structurally diverse chemicals because it has both specific and indiscriminate binding pockets. The enzyme might be used to treat drug overdose and defend against chemical weapon exposure.

Human carboxylesterase (hCE1, also called egasyn), which is found in the liver and several other tissues, metabolizes foreign chemicals—from narcotics to clinical drugs to poisons—by surgically removing small carbon-oxygen chemical groups. The United States military is working to develop the protein into a defense against chemical weapons such as sarin, soman, tabun and VX gas. Not only does the enzyme recognize a wide range of chemicals, it also rejects molecules with subtly different structures, called isomers. To understand how the enzyme could be so general yet so specific, and how it could be improved, the group crystallized mixtures of hCE1 with analogs of cocaine and heroin, called homatropine and naloxone (Fig. 1). They solved the resulting structures from x-ray diffraction patterns obtained at the SER-CAT beamline 22-ID at the APS and beamline 9-1 at the Stanford Synchrotron Radiation Laboratory.

The protein works its magic by juxtaposing a large, flexible binding pocket to a small, rigid one. These are buried deep in the middle of the molecule, in a gorge called the "active site." The small pocket prefers slender chemical linkages, making the protein selective, whereas the larger one can accommodate a variety of bulky chemical groups, making it promiscuous. The enzyme's binding pocket is unique compared to other known proteins that bind cocaine, so it may serve as a model for understanding how the drug binds to targets in the brain and other tissues.

The researchers modeled the orientations of cocaine and heroin in the enzyme's active site based on the positions of their analogs. Cocaine would be positioned to have one of its chemical bonds dissolved by amino acids in the rigid pocket, whereas the corresponding chemical group in one of its isomers (similar to a mirror image of the drug) would clash with the surrounding amino acids. The heroin analog naloxone bound to the active site in two different orientations but preferred one orientation over the other, which explains why hCE1 can metabolize heroin into either of two compounds but carries out one reaction more efficiently than the other. hCE1 is also responsible for creating an especially toxic derivative of cocaine, called cocaethylene, in the presence of alcohol. The researchers propose that ethanol in the blood accesses the protein's active site through a side-door on the surface of the enzyme that opens up into a secondary pore.

Two other human enzymes metabolize cocaine, but one of these prefers the drug's mirror image isomer, and only hCE1

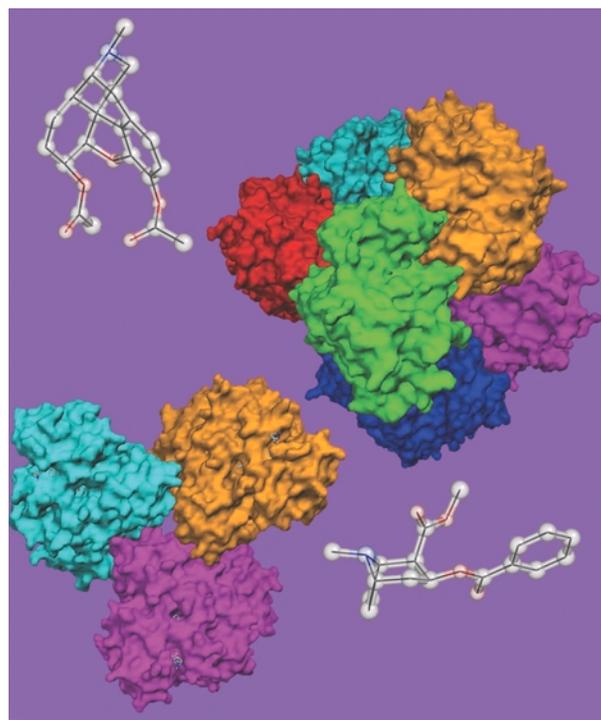


Fig. 1. The "promiscuous" human drug-processing enzyme carboxylesterase 1 (hCE1), which exists as both a three-membered form (lower left) and a six-membered form (upper right), catalyzes the metabolism of the dangerous narcotics heroin (upper left) and cocaine (lower right).

converts it into a compound that is easily removed from the body in urine. A bacterial enzyme proposed to treat cocaine overdose deteriorates rapidly in the blood and may cause an allergic reaction, which hCE1 would be unlikely to do because it is native to humans. The group observed that hCE1 partially clusters into symmetric groups of six (Fig. 1) in the presence of naloxone. Encouraging such clustering may make the purified protein stable enough to provide days worth of protection against chemical weapons. The researchers are now trying to develop a more selective and efficient form of the protein. ○

See: S. Bencharit¹, C.L. Morton², Y. Xue¹, P.M. Potter², and M.R. Redinbo¹, "Structural basis of heroin and cocaine metabolism by a promiscuous human drug-processing enzyme," *Nat. Struct. Biol.* **10**(5), 349-366 (May 2003).

Author affiliations: ¹University of North Carolina at Chapel Hill, ²St. Jude Children's Hospital

This research was supported by the NIH (including a Core Grant), a Burroughs Wellcome Career Award in the Biomedical Sciences, and by the American Lebanese Syrian Associated Charities. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

A CLEAR PICTURE OF THE SARS MAIN PROTEASE STRUCTURE

Severe Acute Respiratory Syndrome (SARS) achieved global notoriety as the first new epidemic of the 21st century. Three months after this new coronavirus was identified and characterized, the disease was declared contained, but only after it had infected over 8,400 people from 30 countries and killed more than 800. SGX (Structural GenomiX, Inc.), in collaboration with the Genome Institute of Singapore, determined the crystal structure of the SARS-CoV main protease to 1.86 Å resolution as a first step toward developing an antiviral therapeutic using structure-based drug discovery.

Residues 1-304 were expressed in *E. coli* as a fusion protein. After purification and cleavage of the fusion partner, the protein was crystallized by hanging drop vapor diffusion. Diffraction data were collected at the SGX-CAT beamline 31-ID at the APS, and the structure was determined via molecular replacement with the program EPMR at SGX, San Diego. The structure was deposited in the Protein Data Bank with accession ID 1Q2W and was released to the public immediately.

In the SARS-CoV M^{pro} dimer, the chain terminus of one molecule falls near the active site of the dimeric partner. Geometric constraints do not allow autoproteolysis. Dimer formation is, therefore, absolutely necessary for trans intradimer cleavage of the opposing molecule from the viral polyprotein.

The natural SARS-CoV M^{pro} substrate is a peptide corresponding to the consensus sequence immediately N-terminal to each one of the mature viral proteins (including the protease itself): TSAVLQ. Specificity for substrate glutamine in the P1 peptide recognition pocket and the presence of a hydrophobic residue in the P2 pocket are features conserved in all coronaviruses. The glutamine sidechain in the P1 position could make hydrogen-bonding interactions with the Phe140 carbonyl oxygen and the His163 imidazole sidechain, both of which occur in the P1 pocket of the enzyme. ○

J.M. Sauder¹, J.B. Bonanno¹, D. Lorimer¹, R. Fowler¹, R. Romero¹, J. Hendle¹, S. Gupta², C.-L. Wei², E.T. Liu², S.K. Burley¹, T. Harris¹

¹Structural GenomiX, Inc., ²Genome Institute of Singapore

Relevant Literature: K. Anand, G.J. Palm, J.R. Mesters, S.G. Siddell, J. Ziebuhr, and R. Hilgenfeld, "Structure of coronavirus main proteinase reveals combination of a chymotrypsin fold with an extra alpha-helical domain," *EMBO J.* **21**, 3213-3224 (2002).

K. Anand, J. Ziebuhr, P. Wadhvani, J.R. Mester, and R. Hilgenfeld, "Coronavirus main proteinase (3CLpro) structure: Basis for design of anti-SARS drugs," *Science* **300**, 1763-1767 (2003).

M.A. Marra et al., "The genome sequence of the SARS-associated coronavirus," *Science* **300**, 1399-1404 (2003).

P.A. Rota et al., "Characterization of a novel coronavirus asso-

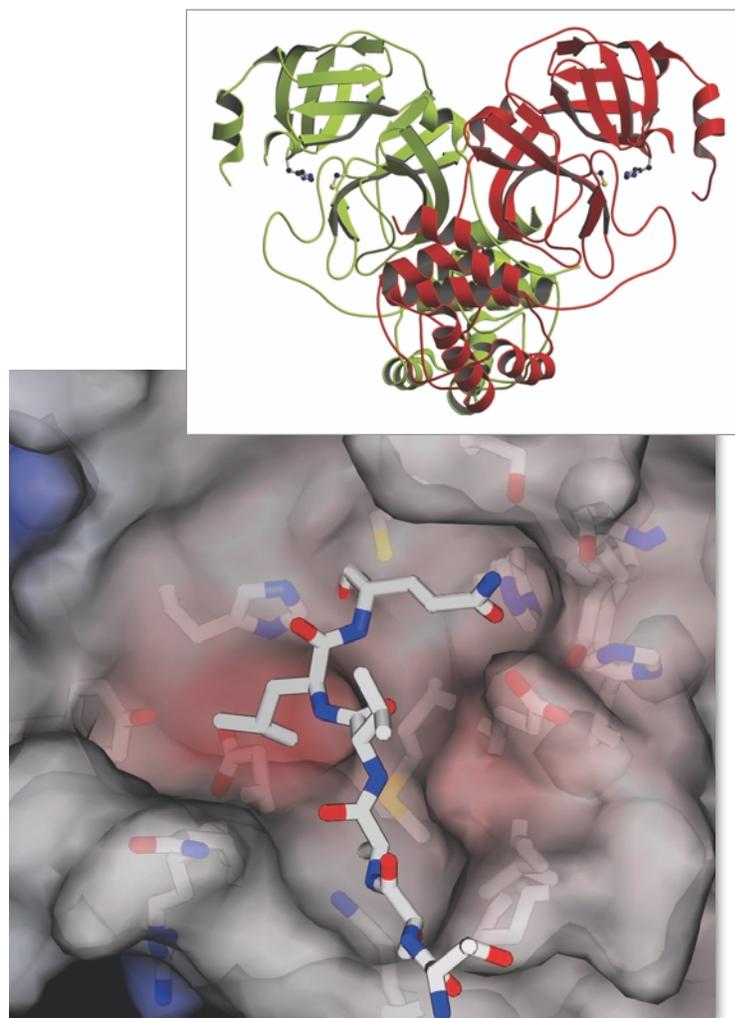


Fig. 1. Top: Ribbon representation of the SARS main protease dimer. Catalytic residues are displayed in ball and stick. Bottom: Surface representation of active site with overlaid model of the consensus peptide.

ciated with severe acute respiratory syndrome," *Science* **300**, 1394-1399 (2003).

Y.J. Ruan et al., "Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection," *Lancet* **361**, 1779-1785 (2003).

E.J. Snijder et al., "Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage," *J. Mol. Biol.* **331**, 991-1004 (2003).

Use of the Advanced Photon source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

THE MOLECULES BEHIND TAY-SACHS DISEASE

Using the BioCARS beamline 14-BM at the APS, a group of researchers from the University of Alberta, the Hospital for Sick Children, the University of Toronto, and Rutgers University have determined how changes in a pair of proteins lead to the family of neurological disorders that includes Tay-Sachs disease. They solved the three-dimensional structure of one protein essential to Tay-Sachs and modeled the structure of a second key protein based on the first. The results help explain why different mutations lead to different forms of the same class of neurological disorder, and may help in developing treatments against them.

Tay-Sachs is caused by an inherited deficiency of the enzyme Hex A, which removes a sugar residue from fatty molecules found in brain cells. Without it, fat builds up in the cells and kills them in a still mysterious way. Three genes must interact to remove the sugar properly: *HEXA* and *HEXB*, which produce the α - and β -subunits of Hex A, respectively; and *GM2A*, whose protein product, GM2A, activates Hex A. Mutations that significantly reduce the activity of these genes can cause one of several related disorders, including Tay-Sachs and Sandhoff disease. One in 35 Ashkenazi Jews carries the mutation that causes Tay-Sachs, and 1 in 300 individuals in general carries it. Understanding the structures of the proteins involved might enable researchers to design drugs to activate or replace them. However, researchers had, to date, only determined the structure of the GM2A protein.

The researchers solved two structural problems at once—Hex A and the closely related Hex B. They first crystallized Hex B, made of two β -subunits, and determined its structure by analyzing its x-ray diffraction pattern. In Hex B, which is easier to crystallize than Hex A, two kidney-shaped β -subunits come together so their active sites, or catalytic regions, face each other but are offset (Fig. 1). Previous studies had indicated that each type of Hex subunit has an active site capable of performing a slightly different type of sugar removal, but that two subunits must come together for either one to be active. To determine how the β -subunit active site binds its substrate, the researchers soaked a Hex B inhibitor that mimics the substrate into the crystallized enzyme. They found that the active sites work cooperatively because several of the amino acids that make up each subunit's active site come from the other subunit. People who have mutations in these amino acids experience a mild, chronic form of Sandhoff disease, or a variant or acute form of Tay-Sachs, depending on which amino acid is mutated. Soaking the crystal with a different inhibitor also confirmed that Hex B uses a similar catalytic mechanism to related enzymes.

Next they modeled the α -subunit based on the β -subunit, which has 60% of the same amino acids. In support of the modeling approach, much of the amino acid similarity occurs in the regions implicated in subunit joining or catalysis, suggesting

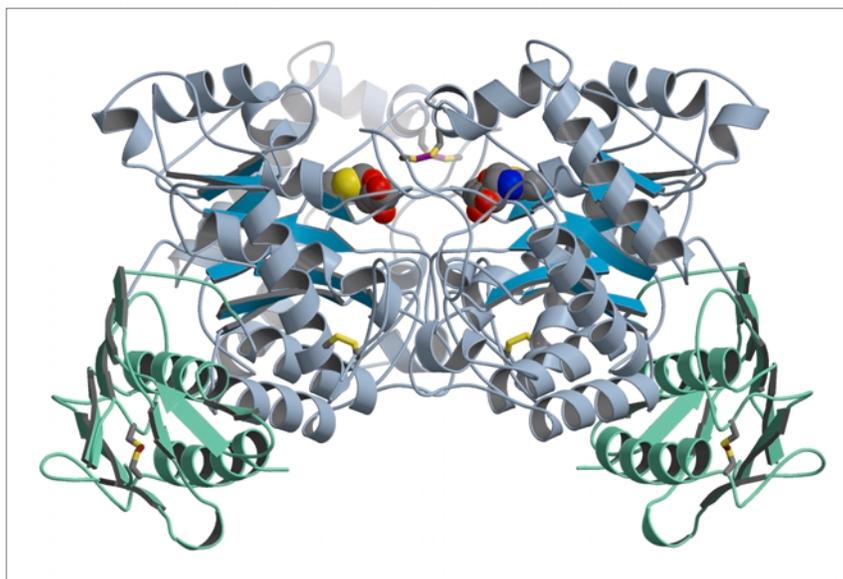


Fig. 1. Schematic image of Hex B showing its two identical, rotationally symmetric subunits. Small molecule inhibitors are shown as black, red, blue, and yellow space-filling models. The enzyme active sites, which contain the inhibitors, are offset perpendicularly to the plane of the image.

that the β -subunit binds the α in the same way it binds itself; Hex A would probably need a similar overall shape to that of Hex B to support the same type of subunit binding. The model indicates that three amino acids in the α -subunit are responsible for the ability of Hex A to bind substrates with a negative charge at one end. Biochemical studies of the enzyme's reaction rate when these amino acids are mutated support the conclusion. The predicted subunit interface also forms a large groove suitable for binding the cup-shaped activator protein Gm2A. Shape and electrostatic complementarities suggest that the activator positions itself over the α -subunit active site. ○

See: B.L. Mark¹, D.J. Mahuran^{2,3}, M.M. Cherney¹, D. Zhao⁴, S. Knapp⁴, and M.N.G. James¹, "Crystal Structure of Human β -Hexosaminidase B: Understanding the Molecular Basis of Sandhoff and Tay-Sachs Disease," *J. Mol. Biol.* **327**, 1093–1109 (2003).

Author affiliations: ¹University of Alberta, ²The Hospital for Sick Children, ³University of Toronto, ⁴Rutgers University

This work supported by Canadian Institutes of Health Research grants (to D.M. and M.N.G.J.). B.L.M. is supported by scholarships from the C.I.H.R., the Alberta Heritage Foundation for Medical Research, and the University of Alberta. M.N.G.J. holds a Canada Research Chair in Protein Structure and Function at the University of Alberta. Use of the BioCARS beamline was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

BINDING OF TWO YEAST TRANSCRIPTION FACTORS

The RNA polymerase enzyme transcribes DNA sequences (genes) into matching RNA molecules. RNA polymerase III produces small RNAs that aid in protein synthesis. Researchers have now determined the three-dimensional (3-D) structure of fragments of two proteins that recruit yeast polymerase III to DNA. The research reveals how these proteins assemble and bind together, despite one of them having a seemingly loose structure. Similar structural features may exist in other recruiting factors for RNA polymerases across the eukaryotes, which include everything from amoeba to humans.

Transcription factor IIIB recruits polymerase III to a region of DNA at the beginning of every gene called the promoter and helps separate the strands of DNA there. It consists of three proteins: a DNA-binding protein called TBP and two connector proteins called Brf1 and Bdp1. Half of Brf1 is very similar to an analogous protein that helps recruit RNA polymerase II, but the other half, especially a region called homology domain II, is more responsible for binding TBP. To understand the structural basis of this binding, a group of experimenters from Yale University and the University of California, San Diego, has crystallized Brf1's homology domain II and the key portion of TBP bound to a short stretch of promoter DNA. They determined the complex's 3-D structure from its x-ray diffraction pattern, which was measured at the SBC-CAT beamline 19-ID at the APS and at the National Synchrotron Light Source beamlines X4A and X25.

The crystal structure shows the symmetric, saddle shaped TBP wedged into the groove between strands of the DNA double helix, straddling the molecule at an angle. Part of Brf1 forms an elongated structure running along the convex outer surface of TBP; the rest forms a bent helix that curls around one of TBP's "stirrups" (Fig. 1). The researchers compare Brf1's fold to that of a vine on a tree, which is markedly different from known structures of polymerase II transcription complexes consisting of TBP and DNA. Brf1 and TBP interact at about 30 locations, which is three to five times more than in the analogous polymerase II complex.

Brf1 is unusually extended along the surface of TBP, nothing like the regular structural domains proteins normally fold into. The researchers found that free Brf1 resists digestion by protein-chewing enzymes (which is only possible if it folds into a more compact, stable structure), and suggest that Brf1 unfolds to grasp TBP. The bent helix may form part of a DNA-binding element, according to earlier research. Researchers suggest its orientation probably positions it for optimal interaction with DNA, Bdp1, and other polymerase III subunits.

Brf1 makes extensive contact with two regions of TBP: a helix on the convex outer surface and a helix and stirrup region on the side. At the first region, Brf1 extensively complements

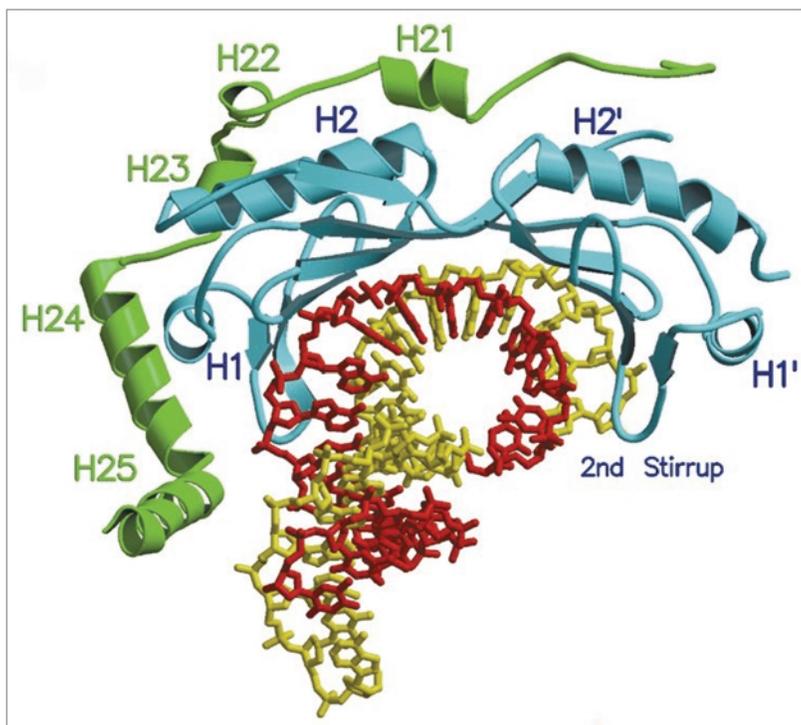


Fig. 1. Image of yeast transcription factor Brf1 (green) bound to the DNA-binding factor TBP (blue). DNA is shown in yellow and red. Brf1 helices H21-23 and their connecting loops surround helix H2 of TBP. Helices H24 and H25 of Brf1 make contact with helix H1 and the first "stirrup" (blue loop, obscured by DNA in red) of TBP.

the charge and shape of TBP, forming 17 hydrogen bonds. The second region contains medium-strength hydrogen bonds formed mostly between the backbones of the two proteins. Proteins normally recognize each other through interactions between amino acid side chains, so the second location probably helps preserve the overall shape of the complex.

Human Brf1 may bind TBP like its yeast counterpart does. On the basis of an amino acid sequence comparison, 28 of 31 hydrogen bonds in the yeast Brf1-TBP complex would exist in the human complex. Yeast Brf1 also has a similar sequence to yeast TAF1, a polymerase II transcription factor (28/31 hydrogen bonds), and human TAF1 (24/31 hydrogen bonds). ○

See: Z.S. Juo¹, G.A. Kassavetis², J. Wang¹, E.P. Geiduschek², and P.B. Sigler^{1(†)}, "Crystal structure of a transcription factor IIIB core interface ternary complex," *Nature* **422**, 534-539 (3 April 2003).

Author affiliations: ¹Yale University, ²University of California, San Diego (†Deceased)

This work was supported in part by the Howard Hughes Medical Institute and by the National Institutes of Health (E.P.G., P.B.S., and T. Steitz). Use of the Advanced Photon Source was funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

STRUCTURE OF A BACTERIAL PROTEIN TRANSPORTER

The three-dimensional (3-D) structure of a protein essential for the tuberculosis bacterium's ability to secrete proteins into its environment has been determined by a team of researchers using the BioCARS 14-BM and 14-ID and SBC-CAT 19-ID beamlines at the APS. The protein, called SecA, may serve as a molecular ratchet, using energy from the breaking of high-energy chemical bonds to physically push other proteins out of a pore in the bacterium's external membrane. The research provides a framework for understanding the protein secretion machinery important for a variety of processes, including transport of membrane proteins and proteins that make tuberculosis infectious. It may also help lead to drugs against the disease, which kills three million people every year.

The system for protein secretion, or translocation, is similar in all types of cells, whether from bacteria or humans. Proteins destined to cross the cell membrane are synthesized with an additional section of protein at one end that alerts the cell to their intended fate. In bacteria, the minimum requirements for translocation are a protein pore, which extends through the membrane, and a membrane enzyme that breaks down ATP, the high-energy molecule that fuels most cellular processes. It is known that SecA binds to other proteins destined for translocation by recognizing the additional piece of protein. Some have proposed that it moves a protein through the pore system by repeatedly inserting and releasing it. A shape change caused by ATP binding would drive the process.

To see if structural data corroborate the hypothesis, the researchers from the Institute of Biosciences and Technology, Texas A&M University, the University of North Carolina, and the Albert Einstein College of Medicine crystallized *Mycobacterium tuberculosis* SecA alone and bound to ADP, the spent form of ATP. They then analyzed its x-ray diffraction pattern, from which they determined its structure. The protein is made of two identical subunits, each of which contains more than 900 amino acids folded into a "motor" domain and a translocation domain (Fig. 1). It is convex on one side and concave on the other, with an elliptical pore running between the two sides. The researchers believe concave surface of SecA, which contains the translocation domain, docks with the membrane pore to form a continuous channel through both proteins.

The group compared the structure of the ADP-bound SecA's motor domain to the very similar domain of a protein that unwinds DNA helices. The motor domain's structure closely matched the "off" state of the unwinding protein (whose mode of action researchers have modeled), implying that SecA undergoes a similar small rotation when switching on or off. The motor domain is composed of two parallel wheel-shaped

pieces, which would rotate about 10° relative to one another. The group believes this small motion would transform into a much larger motion of the translocation domain relative to the motor through a hand-pump-like mechanism. The small rotation would move a long, bent stretch of protein (Fig. 1) up and down like a pump handle, which would then twist the translocation domain back and forth along the direction set of the pore axis. The group speculates that a ratchet effect, resulting from pore contraction or interaction with the membrane channel, would keep translocated proteins moving forward.

The group plans to crystallize the bacterium's other

SecA protein, SecA2, which transports proteins that make tuberculosis virulent. Comparing its structure with that of SecA may help to better understand the disease and aid in the design of new pharmaceuticals. ○

See: V. Sharma^{1,2}, A. Arockiasamy², D.R. Ronning², C.G. Savva², A. Holzenburg², M. Braunstein³, W.R. Jacobs, Jr.⁴, and J.C. Sacchettini^{1,2}, "Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase," Proc. Nat. Acad. Sci. U.S.A. **100** (5), 2243–2248 (4 March 2003).

Author affiliations: ¹Institute of Biosciences and Technology, ²Texas A&M University, ³University of North Carolina, ⁴Albert Einstein College of Medicine

Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract W-31-109-Eng-38. Use of the BioCARS beamline was supported by the National Institutes of Health, National Center for Research Resources, under grant number RR07707. This work was supported by the Robert A. Welch Foundation and National Institutes of Health Grant GM62410. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

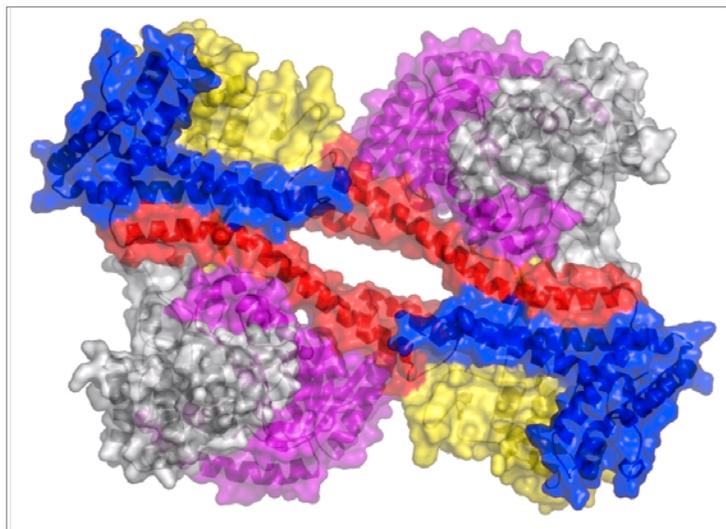


Fig. 1. SecA, made of two identical subunits, is shown here in a schematic representation of its protein fold (ribbons) underneath a translucent image of its molecular surface. The motor domain (purple, gray), translocation domain (blue, yellow), and long "pump handle" (red) surround an elliptical pore.