

Take a moment to ponder the form of your body: the shape of your face, the color of your eyes, the length of your fingers, the perfect articulation of your bones and muscles, the way your hair grows curly or straight. Now let your imagination travel inward, and think of the complex shapes and functions of your different cells, and the teeming molecular world inside each one.

Remarkably, this amazing structure and form and function is specified by information in the genome, which encodes a mere 20,000-25,000 protein-coding genes. One of the great puzzles being pieced together by scientists is the mechanism by which these genes, and the methods used to control their expression, specify all of these different aspects of life.

About the  
*RCSB PDB Molecule of the Month*

Using selected molecules from the PDB archive, each feature includes an introduction to the structure and function of the molecule, a discussion of its relevance to human health and welfare, and suggestions for viewing and accessing further details.

The *RCSB PDB Molecule of the Month* is read by students, teachers, and scientists worldwide at [www.pdb.org](http://www.pdb.org).

This February 2010 edition was written and illustrated by David S. Goodsell (RCSB PDB and The Scripps Research Institute).

## Combinatorial Control

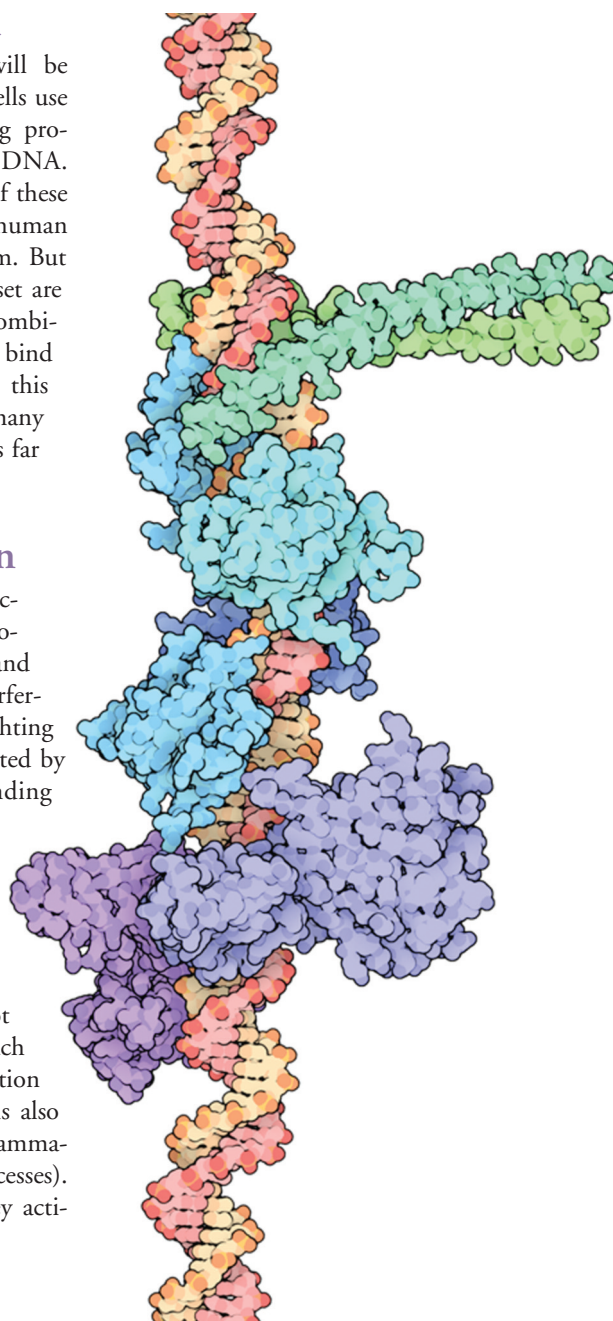
In order to specify which gene will be expressed in a given situation, your cells use a diverse collection of DNA-binding proteins to control access to the DNA. Surprisingly, there are relatively few of these proteins: by some estimates, the human genome encodes about 2,600 of them. But then, the capabilities of this limited set are greatly expanded by using them in combination, by requiring two or more to bind simultaneously to activate a gene. In this way, each protein may be used in many ways and the spectrum of responses is far more varied.

## Enhancing Transcription

The assembly of DNA and proteins pictured here is a transcriptional enhanceosome (PDB entries 1t2k, 2pi0, 2o6g and 2o61) that controls expression of interferon-beta, an important protein for fighting viral infection. When the cell is infected by viruses, several different DNA-binding proteins are produced, including ATF-2/c-Jun (in green at the top), interferon response factors (IRF, shown in turquoise at the center), and nuclear factor kB (NF-kB, shown in blue and magenta at the bottom). Individually, each one is not sufficient to activate the gene, and each one also plays other roles in the activation of other genes (for instance, Nf-kB is also important in immune responses, inflammation, apoptosis, and many other processes). But when they all bind together, they activate the gene and interferon is made.

## Integrating the Signal

Once the transcription factors bind to the different sites in the enhancer DNA sequence, the signal must somehow be sensed and used to activate the gene. In many cases, this is performed using CREB-binding protein or the similar protein p300. This protein



is composed of many connected domains, (PDB entries 1l8c, 1kdx, 1jsp, 3biy, 2ka6 and 1kbh), which bind to different proteins in the assembled enhanceosome. Then, a large domain in the center acts as a histone acetyl-

## RCSB Protein Data Bank

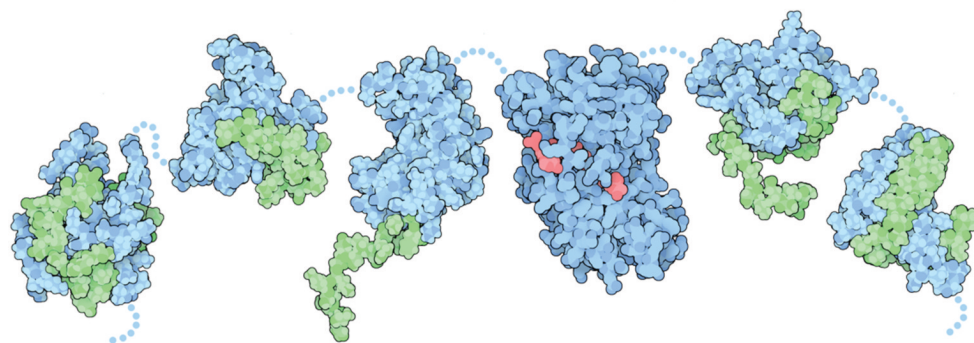
The Protein Data Bank (PDB) is the single worldwide repository for the processing and distribution of 3D structure data of large molecules of proteins and nucleic acids. The RCSB PDB is operated by Rutgers, The State University of New Jersey and the San Diego Supercomputer Center and the Skaggs School of Pharmacy and Pharmaceutical Sciences at the University of California, San Diego—two members of the Research Collaboratory for Structural Bioinformatics (RCSB).

It is supported by funds from the National Science Foundation, the National Institute of General Medical Sciences, the Office of Science, Department of Energy, the National Library of Medicine, the National Cancer Institute, the National Institute of Neurological Disorders and Stroke and the National Institute of Diabetes & Digestive & Kidney Diseases.

The RCSB PDB is a member of the worldwide PDB ([www.pdb.org](http://www.pdb.org)).

## References:

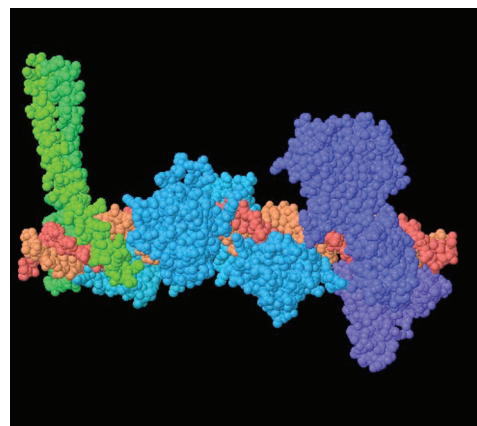
- 1t2k**: D. Panne, T. Maniatis, S.C. Harrison (2004) Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer. *Embo J.* **23**, 4384-4393
- 2pi0**: C.R. Escalante, E. Nistal-Villan, L. Shen, A. Garcia-Sastre, A.K. Aggarwal, (2007) Structure of IRF-3 bound to the PRDIII-I regulatory element of the human interferon-beta enhancer. *Mol.Cell* **26**, 703-716
- 2o6g**, **2o61**: D. Panne, T. Maniatis, S.C. Harrison (2007) An Atomic Model of the Interferon-beta Enhanceosome. *Cell(Cambridge,Mass.)* **129**, 1111-1123
- 1l8c**: S.A. Dames, M. Martinez-Yamout, R.N. De Guzman, H.J. Dyson, P.E. Wright (2002) Structural basis for Hif-1 alpha /CBP recognition in the cellular hypoxic response. *Proc.Natl.Acad.Sci.USA* **99**, 5271-5276
- 1kdx**: I. Radhakrishnan, G.C. Perez-Alvarado, D. Parker, H.J. Dyson, M.R. Montminy, P.E. Wright (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. *Cell(Cambridge,Mass.)* **91**, 741-752
- 1jsp**: S. Mujtaba, Y. He, L. Zeng, S. Yan, O. Plotnikova, Sachchidanand, R. Sanchez, N.J. Zeleznik-Le, Z. Ronai, M.M. Zhou (2004) Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol.Cell* **13**, 251-263
- 3byj**: X. Liu, L. Wang, K. Zhao, P.R. Thompson, Y. Hwang, R. Marmorstein, P.A. Cole (2008) The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature* **451**, 846-850
- 2ka6**: J.M. Wojciak, M.A. Martinez-Yamout, H.J. Dyson, P.E. Wright (2009) Structural basis for recruitment of CBP/p300 coactivators by STAT1 and STAT2 transactivation domains. *Embo J.* **28**, 948-958
- 1kbb**: S.J. Demarest, M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H.J. Dyson, R.M. Evans, P.E. Wright (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-553
- 1t2k**: D. Panne, T. Maniatis, S.C. Harrison (2004) Crystal structure of ATF-2/c-Jun and IRF-3 bound to the interferon-beta enhancer. *Embo J.* **23**, 4384-4393



transferase, modifying histones in nucleosomes and causing them to disassemble and reveal the gene. In the interferon- $\beta$  gene, a nucleosome normally covers the start site of transcription, blocking transcription. Assembly of the enhanceosome leads to removal of this nucleosome, allowing the gene to be expressed.

## Exploring the Structure

The 3D model of the entire interferon- $\beta$  enhancer was created by Daniel Panne based



on several crystallographic structures (PDB entries **1t2k**, **2pi0**, **2o6g** and **2o61**). It reveals several surprising aspects of the assembly. First, it shows that the DNA is nearly straight—some researchers had predicted that the DNA would be significantly bent. Second, the interaction between the proteins is very limited, even though the binding of the proteins is known to be cooperative, such that binding of one protein promotes the binding of others. This may be a consequence of several things. Several of the proteins bind on opposite sides of the same part of the DNA, so the proteins may promote binding by distorting the DNA slightly into a preferred conformation. Also, the crystal structures of ATF-2/c-Jun and the IRF proteins were solved using only the DNA-binding portions of the proteins, so the missing portions of the proteins may interact cooperatively with their neighbors. Finally, inside cells, CREB-binding protein or p300 (shown on the previous page) might generate a strong cooperative effect through its binding to the different proteins. Visit [www.pdb.org](http://www.pdb.org) to view an interactive Jmol version.

## Topics for Further Exploration

1. Additional structures of several of these proteins are available in the PDB. Are they similar in structure to the proteins in the enhanceosome complex?
2. Many different protein folding motifs are used to recognize DNA. Can you find other examples in the PDB?

## Additional Information on Enhanceosomes

- D. Panne, T. Maniatis and S. C. Harrison (2007) An atomic model of the interferon- $\beta$  enhanceosome. *Cell* **129**, 1111-1123
- D. Panne (2008) The enhanceosome. *Current Opinion in Structural Biology* **18**, 236-242
- L. Wang, Y. Tang, P. A. Cole and R. Marmorstein (2008) Structure and chemistry of the p300/CBP and Rtt109 histone acetyltransferases: implications for histone acetyltransferase evolution and function. *Current Opinion in Structural Biology* **18**, 741-747
- Number of genes in the human genome: International Humane Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931-945
- Number of transcription factors in the human genome: M. M. Babu, N. M. Luscombe, L. Aravind, M. Gerstein and S. A. Teichmann (2004) Structure and evolution of transcriptional regulatory networks. *Current Opinion in Structural Biology* **14**, 283-291