DNA

Modularity of eukaryotic transcription activators

Mark Ptashne, Grace Gill and Roger Brent

Mark Ptashne, Grace Gill, and Roger Brent

The key experiment described here showed that a eukaryotic transcriptional activator is a modular protein composed of a DNA-binding domain and an "activating region" (2466). The definition of these as distinct, separable elements led to a coherent picture for how such activators work; provided a powerful tool for the analysis of other gene-activating and - repressing functions; laid the basis for the development of the two-hybrid system for detecting protein-protein interactions; and now provides a framework for understanding how specificity is achieved in many cellular information processing networks.

Background

The key experiment we shall describe was performed in yeast, a simple eukaryote. Our understanding of gene regulation in bacteria around 1985 strongly influenced the design of the experiment. By that time, we had in hand three concepts describing how a bacterial regulatory protein could recognize a specific sequence in DNA (called an operator) and activate or repress transcription of a gene. We begin here with a brief statement of these three concepts, as exemplified by the action of lambda repressor, the protein we were studying most intensively (see *EXP, Isolation of repressor*†):

- A regulatory protein (repressor or activator) binds a specific sequence in DNA without significantly changing the structure of either the protein or the DNA.
- A repressor turns off (represses) a gene by excluding RNA polymerase from that gene. This effect requires that an operator site overlap the binding site for RNA polymerase (the promoter).
- An activator turns on (activates) transcription of a gene that bears an operator site immediately adjacent to the promoter. Activation is effected by a protein-protein interaction between the regulator and RNA polymerase, an interaction that we now know recruits RNA polymerase to the gene. Specific amino acids on a surface of the activator, distinct from the DNA interaction surface, are required for the interaction with RNA polymerase.

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

Figure 1 illustrates the third concept, the activation of transcription by lambda repressor.

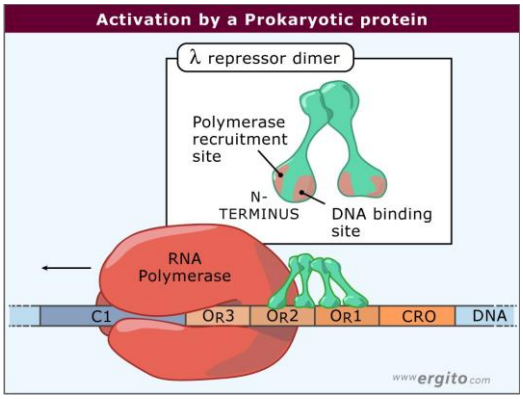

Figure 1 Activation by a prokaryotic protein. The amino terminus of a lambda repressor dimer bound to the operator site OR2 recruits (binds cooperatively to DNA with) RNA polymerase. The polymerase transcribes the repressor gene (*cI*), and thus repressor stimulates transcription of its own gene. The polymerase site contacted by repressor, we now know, is in its sigma70 subunit.

In the mid-1980s, we began to wonder whether this picture from bacteria applied in eukaryotes as well. There were good grounds for thinking it might not. For example, unlike in bacteria, DNA in eukaryotes is wrapped around histone proteins to form nucleosomes, and the DNA is sequestered in a special compartment, the nucleus. Moreover, unlike in the lambda case, eukaryotic regulators often exert their effects over many hundreds, or even thousands, of base pairs. As we shall see, however, the lessons from bacteria stood us in good stead, and taken together with the studies from yeast we now describe, provided a unified view of many aspects of gene regulation, applying to both bacteria and eukaryotes. (2472, 2481, 2471, 3048, 100)

The Yeast Transcriptional Activator Gal4 In the early 1980s, we and others had begun analysis of gene activation in a simple eukaryote, the yeast *S. cerevisiae*. Genetic studies had shown that the Gal4 protein is required for transcription of various Gal genes, the products of which metabolize the sugar galactose. (See, for example, 2480.) Addition of galactose to the medium induced transcription of the various *GAL* genes well over a thousand fold. This feature, taken with the ease with which experiments could be done with yeast, encouraged us to study gene activation by Gal4.

Our first important finding was that Gal4 is indeed a specific DNA-binding protein. Guarente and coworkers had shown that a fragment of about 100 base pairs, if positioned some 250 base pairs upstream of the *GAL1* gene, was sufficient to confer galactose regulation on another promoter, that of the gene *CYC1* (2471). They termed the site of Gal4 action the Galactose Upstream Activating Sequence, or UAS_G.

Using *in vivo* footprinting techniques and deletion analysis, we found that Gal4 bound specifically to repeated 17-base-pair sequences in UAS_G (2497, 2498, 2468). At the time, we did not know how Gal4 recognized DNA. The symmetry of each of the sites suggested that the protein bound as a dimer, but the Gal4 protein sequence did not bear the helix-turn-helix motif characteristic of the bacterial DNA-binding regulators known at that time. This contributed to our concern that protein binding to DNA in a eukaryotic nucleus, where it is wrapped in nucleosomes, might require some still undescribed mechanism. The following preliminary yeast experiment shows that this concern was unfounded.

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

Setting the Stage: Repression by a Bacterial Protein in Yeast

This preliminary experiment was performed with a bacterial repressor called LexA. We and others had shown that LexA binds to well-defined DNA sites, a dimer of the protein rec- ognizing each site (2464, 2479, 2475). Our knowledge was consistent with the idea that DNA-bound LexA works as do other repressors: it turns off transcrip- tion of a gene by competing for binding of polymerase to an overlapping pro- moter. We knew that LexA comprised two domains connected by a flexible linker: an amino terminal domain that contained the DNA-binding function, and a carboxyl terminal domain that contained a dimerization function. Sequence com- parison indicated that LexA recognized DNA using a helix-turn-helix motif simi- lar to that found in lambda repressor. Unlike lambda repressor, there was no evi- dence that LexA activated transcription in prokaryotes; as we now understand, it lacks an activating region.

We placed LexA binding sites between the Gal4 binding sites (UAS_G) and the start site of transcription of a yeast reporter. We introduced these constructs into yeast along with a second construct (now called an effector construct) expressing the LexA gene under the control of a yeast promoter. LexA repressed transcription from some of these constructs. That is, we found certain positions for the LexA sites at which repression occurred, but others also between the UAS_G and the gene, at which LexA had no effect (2465).

The results showed that, in yeast, LexA bound its operator sites. Thus, this bacte- rial protein entered a eukaryotic nucleus and bound its operator sites despite DNA's presumably being wrapped in nucleosomes. The results also argued against so-called "transmission" models for activation by Gal4. According to such models, the activator transmitted some signal (for example, to unwind the DNA) through or along the DNA. Those models predicted that the repressive effects of LexA would not depend on the exact positioning of the LexA bound between the UAS_G and the gene.

The Key Experiment: The Domain Swap

We thus began with two proteins that bound different DNA sites in yeast. One, the yeast activator Gal4, activated transcription; the other, the bacterial repressor LexA, did not. In bacteria, lambda repressor binds immediately adjacent to the gene it activates, whereas in yeast, Gal4 typically binds some 250 base pairs away. We nevertheless imagined that the role of DNA binding might be the same in both cases, namely, to position the activator near the gene. In that case, we might be able to replace the Gal4 DNA- binding domain with that of LexA and activate transcription of a gene with LexA binding sites nearby.

We constructed and expressed in yeast a chimeric protein, LexA-Gal4. The fusion carried the DNA-binding region of LexA (residues 1-87) fused to a C- terminal portion of Gal4 (residues 74-881) that contained more than 90 percent of the Gal4 coding sequence.

The key result is shown in Figure 2. When LexA-Gal4 was expressed in yeast, it activated transcription of a reporter gene near which we had placed LexA DNA-binding sites. LexA-Gal4 did not activate if the gene lacked LexA binding sites, even if it carried a Gal4 binding site (a "17-mer").

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

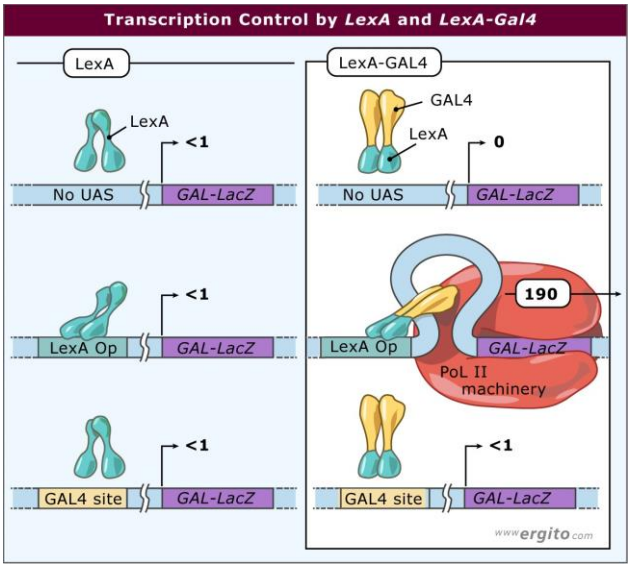

Figure 2 Activation by LexA and LexA-Gal4. The figure shows transcription measured by expression of a *gal1-lacZ* reporter gene. LexA-Gal4 stimulates transcription of a gene that carries a LexA operator, but not of a gene that carries a Gal4-binding site. Numbers denote units of -βgalactosidase activity. (This figure is adapted from Table 3 of Brent and Ptashne, 1985.)

This experiment, whose findings are illustrated in Figure 3, showed the following:

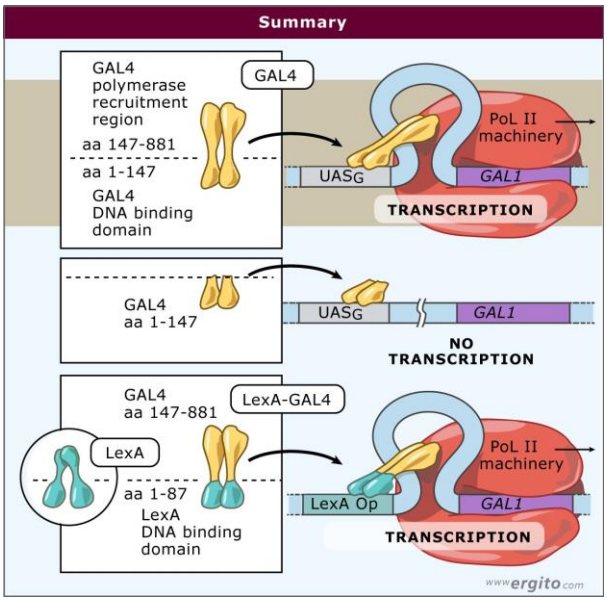

Figure 3 The domain swap experiment summarized: separation of DNA-binding and -activating regions. LexA-Gal4 binds LexA operators and activates transcription, LexA binds LexA operators, and Gal4 (1–147) binds Gal4 sites. None of them, however, activate trancription because none bear an activating region such as that contained in the carboxyl part of Gal4.

- Gal4 contains an "activating region" that can be transferred, without loss of function, to a heterologous DNA-binding domain.
- No special mode of DNA binding is required for activation in eukaryotes beyond that found in a bacterial repressor.
- Specificity of an activator is determined by its DNA-binding address: Gal4 works if a gene bears a Gal4 binding site, and LexA-Gal4 works if a gene bears a LexA site. The role of DNA binding is to tether the activating region near the gene to be activated.
- The activating region itself does not activate unless tethered to DNA.
- Eukaryotic transcriptional activators are modular proteins, surprisingly easily separated into different functional bits.

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

- The amino portion of Gal4 (missing from the LexA-Gal4 fusion) bears the Gal4 DNA binding domain. A Gal4 fragment bearing that domain (e.g., Gal4 [1-147] binds Gal4 sites in yeast but does not activate transcription because it lacks an activating region [2474]). This Gal4 fragment is analogous to the so-called "positive control" (*pc*) mutants of lambda repressor. Those mutants, which bear amino acid changes on the surface of repressor that contacts RNA polymerase, bind DNA normally but do not activate transcription (2479).

Serendipity

Before touching upon some subsequent developments, we point out three bits of good fortune that allowed the success of the LexA-Gal4 domain swap experiment.

LexA turned out to be a good DNA-binding domain in eukaryotes. For unknown reasons (possibly small differences in protein level and/or operator affinity), lambda repressor, the protein we had studied most intensively, did not work in the initial yeast experiments. Had we not been studying LexA in previous years, we would not have had a prokaryotic DNA domain that we knew bound DNA in yeast.

The LexA fragment that was used to make LexA-Gal4 lacked the C-terminal dimerization domain required for efficient DNA binding of LexA. We believed that Gal4 also bound DNA as a dimer and, in those days before the widespread use of synthetic DNA, took advantage of a well-placed restriction site to generate a very large C-terminal fragment of Gal4 fused to the amino part of LexA. Luckily, that carboxyl Gal4 fragment supplied, in addition to an activating region, a dimerization function, but just barely: we now know that the first Gal4 residue in LexA-Gal4 is the first residue of the dimerization region. The ability of the Gal4 dimerization region to effect dimerization (and hence efficient DNA binding) of the LexA DNA binding domain defines another modular, swappable, functional element.

We subsequently learned that LexA and LexA-Gal4 lacked nuclear localization sequences and were not concentrated in the nucleus (2477). Both proteins were functional (even the >90-kiloDalton LexA-Gal4), which showed that such signals were not always required for nuclear entry, contrary to what we had believed up to that time.

Some Subsequent Developments

There were previous examples of protein modularity (for example, the DNA-binding and dimerization domains of lambda repressor). However, the domain swap experiment and the continuing flood of experiments to construct new multifunctional chimeric proteins vividly illustrate just how widespread is the modular principle of protein design.

Gal4, we soon learned, is a "universal" activator: it will work in any eukaryote, provided the test gene is modified by the introduction of Gal4 binding sites nearby and Gal4 is artificially expressed in that organism (2473, 2478). Gal4 is widely used to activate specific genes in specific tissues (e.g., in *Drosophila*; see 2499). Chimeras that contain other universal activation regions are used to turn genes on and off in response to treatment, for example, with tetracycline (2469).

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

Universal activating regions typically bear an excess of acidic residues, plus crucial hydrophobic residues. One can readily detect activating regions by fusing candidates to a DNA-binding domain, expressing the hybrid proteins, and assaying for activation of a reporter gene.

Tethering the activating region to the promoter does not require that the activating region be covalently linked to the DNA-binding domain. Bringing the activating region to DNA by a protein-protein interaction triggers activation of transcription. Consider, for example, the protein Gal80, which inhibits Gal4 by binding to and covering Gal4's activating region. Fusion of an activating region con- verts Gal80 to an activator when it binds to DNA-bound Gal4 (2476). Noting this, Fields and Song developed the two-hybrid system (952). That system detects interaction between proteins, one fused to a DNA-binding domain and the other fused to an activating region, by activation of transcription. Transcription by tethered activating regions (as in the two-hybrid system) is now used to map interactions among the protein complements of entire genomes.

Several lines of evidence show that the typical eukaryotic transcriptional activating region works by recruiting the transcriptional machinery to a nearby promoter. One finding consistent with this idea was that, when expressed in large amounts and not tethered to DNA, free activating regions squelched transcription (2467). The properties of activating regions also accord with this surmise. In particular, studies showing eukaryotic activating regions to be flexible in both size and sequence helped spell the end of the idea that the activation function might involve an enzymatic activity. Various lines of evidence suggest that these activating regions can touch an array of targets in the very complicated transcriptional machinery of eukaryotes (over fifty proteins), recruiting whatever is necessary to ensure transcription of a given gene. According to this general picture, the typical eukaryotic activator works by the same fundamental mechanism as does lambda repressor working as an activator: at the heart of each process lies cooperative binding to DNA of activator and target, a process that recruits the transcription machinery to the promoter.

The Legacy

Classical biochemistry taught us that substrate recognition by enzymes (those involved in intermediary metabolism, for example) was determined by highly specific interactions governed by the precise geometries of the interacting moieties. Each such enzyme typically recognizes one and only one substrate, usually a small molecule. Now we see that many enzymes, particularly those involved in regulatory processes, share a property with RNA polymerase: their specificities are not determined solely by their active sites. Rather, just as RNA polymerase has many possible genes it might work on, so too does the typical kinase, phosphatase, protease, etc., have many possible substrates. Which one is chosen, the regulatory choice, is determined in many cases by the kinds of mechanisms we describe here for transcriptional activation.

For example, modular regulatory proteins involved in proteolysis (e.g., F box proteins) direct a ubiquitylating enzyme to specific substrates by binding the enzyme to one surface, and the specific substrates to another. This strategy, imposing specificity on enzyme action by recruitment, figures prominently in the workings of the intracellular pathways that the cell uses to transmit and process information (2500).

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

The Authors

Mark Ptashne received his Ph.D. from Harvard University, where he was a Junior Fellow. He then became lecturer (1968-71), professor (1971-96), and chairman of the Department of Biochemistry and Molecular Biology (1980-83). Since 1997, he has occupied the Ludwig Chair of Molecular Biology at Sloan-Kettering Institute. He was scientific cofounder of the Genetics Institute in 1980. His many awards and honors include the Eli Lilly Award in Biological Chemistry (1975); le Prix Charles-Leopold Mayer, l'Academie des Sciences, Paris, France (with W. Gilbert) (1977); the U.S. Steel Foundation Award in Molecular Biology (1979); the General Motors Cancer Research Foundation Sloan Prize (1990); and the Lasker Award for Basic Research (1997). His research has concentrated on discerning general principles of gene regulation using phage lambda and yeast as model systems.

Grace Gill is currently an Assistant Professor in the Department of Pathology at Harvard Medical School. She received her Ph.D. from Harvard University where she studied the transcriptional activation function of GAL4 with Mark Ptashne. She was a postdoctoral fellow in the laboratory of Robert Tjian at the University of California, Berkeley, where she continued to investigate transcriptional activation mechanisms. Grace Gill was an Instructor of the Eukaryotic Gene Expression Course at Cold Spring Harbor Laboratory from 1997-2000 and a Visiting Scientist at Cell Press 2000-2001. Current research in her laboratory is directed towards understanding how the activity of specific promoters and transcription factors are regulated during development of the nervous system.

Mark Ptashne
Memorial Sloan Kettering Cancer Institute
1275 York Avenue, Box 595
New York, NY 10021
Phone: 212-639-5185
Fax: 212-717-3627
Email: m-ptashne@ski.mskcc.org

Grace Gill
Harvard Medical School
Pathology Department
200 Longwood Avenue, Boston, MA 02115
Phone: 432-0985
Fax: 432-1313
Email: grace_gill@hms.harvard.edu

Roger Brent
President and Research Director, The Molecular Sciences Institute
2168 Shattuck Avenue, Second Floor
Berkeley, CA 94704
Phone: 1-510-647-0690
Fax: 1-510-647-0699
Email: brent@molsci.org

10-21-2000

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195

REVIEWS

100. Ptashne, M. (1992) . *A Genetic Switch.* The Genetic Switch

2500. Ptashne, M. and Gann, A. (2002) . Genes and Signals

REFERENCES

952. Fields, S., and Song, O. (1989) . *A novel genetic system to detect protein-protein interactions.* Nature 340, 245 - 246. PubMed ID=89330544.

2464. Brent, R. and Ptashne, M. (1980) . *The lexA gene product represses transcription of its own promoter.* Proc. Nat. Acad. Sci. USA 77, 1932 - 1936. PubMed ID=80190159.

2465. Brent, R. and Ptashne, M. (1984) . *A bacterial repressor protein or a yeast transcriptional terminator can block upstream activation of a yeast gene.* Nature 312, 612 - 615. PubMed ID=85061617.

2466. Brent, R. and Ptashne, M. (1985) . *A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor.* Cell 43, 729 - 736. PubMed ID=86079556.

2467. Gill, G. and Ptashne, M. (1988) . *Negative effect of the transcriptional activator GAL4.* Nature 334, 721 - 724. PubMed ID=88318931.

2468. Giniger, E., Varnum, S. M., and Ptashne, M. (1985) . *Specific DNA binding of GAL4, a positive regulatory protein of yeast.* Cell 40, 767 - 774. PubMed ID=85176930.

2469. Gossen, M. and Bujard, H. (1992) . *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proc. Nat. Acad. Sci. USA 89, 5547 - 5551. PubMed ID=92302280.

2471. Guarente, L., Yocum, L., and Gifford, P. (1982) . *A GAL10-CYC1 hybrid yeast promoter identifies the GAL4 regulatory region as an upstream site.* Proc. Nat. Acad. Sci. USA 79, 7410 - 7414. PubMed ID=83117688.

2472. Hochschild, A., Irwin, N., and Ptashne, M. (1983) . *Repressor structure and the mechanism of positive control.* Cell 32, 319 - 325. PubMed ID=83129422.

2473. Kakidani, H. and Ptashne, M. (1988) . *GAL4 activates gene expression in mammalian cells.* Cell 52, 161 - 167. PubMed ID=88135748.

2474. Keegan, L., Gill, G., and Ptashne, M. (1986) . *Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein.* Science 231, 699 - 704. PubMed ID=86122884.

2475. Little, J. W., Mount, D. W., and Yanisch-Perron, C. R. (1981) . *Purified lexA protein is a repressor of the recA and lexA genes.* Proc. Nat. Acad. Sci. USA 78, 4199 - 4203. PubMed ID=82037806.

2476. Ma, J. and Ptashne, M. (1988) . *Converting a eukaryotic transcriptional inhibitor into an activator.* Cell 55, 443 - 446. PubMed ID=3180218.

2477. Silver, P. A., Brent, R., and Ptashne, M. (1986) . *DNA binding is not sufficient for nuclear localization of regulatory proteins in S. cerevisiae.* Mol. Cell Biol. 6, 4763 - 4766. PubMed ID=87089841.

2478. Webster, N., Jin, J. R., Green, S., Hollis, M., and Chambon, P. (1988) . *The yeast UASG is a transcriptional enhancer in human HeLa cells in the presence of the GAL4 trans-activator.* Cell 52, 169 - 178. PubMed ID=88135749.

2479. Brent, R. and Ptashne, M. (1981) . *Mechanism of action of the lexA gene product.* Proc. Nat. Acad. Sci. USA 78, 4204 - 4208. PubMed ID=82037807.

2480. Johnston, S. A, and Hopper, J. E. (1982) . *Isolation of the yeast regulatory gene GAL4 and analysis of its dosage effects on the galactose/melibiose region.* Proc. Nat. Acad. Sci. USA 79, 6971 - 6975. PubMed ID=83091077.

2481. Maniatis, T., and Ptashne, M. (1973) . *Multiple repressor binding at the operators of bacteriophage lambda.* Proc. Nat. Acad. Sci. USA 70, 1531 - 1535. PubMed ID=4514322.

2497. West, R. W., Yocum, R. R., and Ptashne, M. (1984) . *S. cerevisiae GAL1-GAL10 divergent promoter region: location and function of the upstream activating sequence UASG.* Mol. Cell Biol. 4, 2467 - 2478. PubMed ID=85085953.

2498. Yocum, R. R., Hanley, S., West, R., and Ptashne, M. (1984) . *Use of lacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in S. cerevisiae.* Mol. Cell Biol. 4, 1985 - 1998. PubMed ID=85061187.

2499. Brand, A. H., and Perrimon, N. (1993) . *Targeted gene expression as a means of altering cell fates and generating dominant phenotypes.* Development 118, 401 - 415. PubMed ID=94038684.

3048. Jeffreys, A. J., Tamaki, K., McCleod, A., Monckton, D. G., Neil, D. L. and Armour, J. A. L. Jeffreys,

A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. (1994) . *Complex gene conversion events in germline mutation at human minisatellites.* Nat. Genet. 6, 136 - 145. PubMed ID=94214492.

† Located at http://www.ergito.com. Prefixes in capital letters indicate the Ergito module.

http://dx.doi.org/ 10.6076/J7RN35SF
http://authors.fredhutch.org/1195