23. $Af\rho$ is defined as the product of the dust albedo A, the filling factor f of the dust within the aperture, and the projected radius ρ of the aperture; it provides an aperture-independent measure of the dust production for a comet with a canonical spatial distribution of dust.
24. C. Arpigny *et al.*, *Atlas of Cometary Spectra* (Kluwer, Dordrecht, Netherlands, 1997).
25. M. F. A'Hearn, R. L. Millis, D. G. Schleicher, D. J. Osip, P. v. Birch, *Icarus* **118**, 223 (1995).
26. D. Schleicher *et al.*, *IAU Circular 6372* (1996).
27. Depending on the characteristic length of small-scale intensity variations, the distance Δ of the comet, the angular resolution of the instrument, and the photon statistics, $\tilde{s}_x$ may be an average over more extended regions and thus different from $\hat{s}_x$. As a compromise between high spatial resolution and sufficient signal to noise ratio, we determined $\tilde{s}_x$ by binning the photons into 30″ by 30″ pixels and smoothing the images with a gaussian filter with σ = 2′.
28. According to Cravens (*17*), we used the following parameters: cross section, $\sigma = 3.0 \times 10^{-15}$ cm^2; expansion velocity in the coma, v_c = 1.0 km s^{-1}; velocity of the solar wind, v_{sw} = 400 km s^{-1}; density of solar wind ions, $n_{sw} = 10$ cm^{-3}; fraction of heavy ions, f_h = 0.001; x-ray yield, f_y = 0.1; energy released per heavy ion, E_{ion} = 0.2 keV. The neutral gas density profile in the coma was assumed to be $n_c(r_c) = Q_a/4\pi v_c r_c^2$ where r_c is the cometocentric distance and $Q_a = 1.5 \cdot Q_{H_2O}$. The gas production rate $Q_{H_2O} = 2 \times 10^{29}$ s^{-1} adopted by Cravens for obs. no. 8 referred to a measurement on March 24.5 UT, when Q was systematically higher than during the ROSAT observations from March 26.5 to 28.4 (*26*).
29. This simplified model has several interesting implications: (i) The volume emissivity scales directly with the heavy ion flux, and so does the integrated column density. Thus, a change in the heavy ion flux (that is, in v_{sw}, n_{sw}, or f_h) causes each image pixel to change in intensity by the same factor, leaving the HWHM radius r_x unaffected. Uncertainties in f_y or E_{ion} have similar consequences. There is, however, a limitation in this simplified model: for high gas production rates, r_x becomes so large that the assumption of a $n_c(r_c) \sim r_c^{-2}$ density profile is no longer valid. Formally, a gas production rate $Q = 10^{29}$ s^{-1} would yield $r_x = 1.8 \times 10^5$ km. For Q/Δ below 10^{28} s^{-1} AU^{-1}, r_x becomes smaller than 1′ and thus difficult to resolve. (ii) A change in the gas production rate Q causes each dimension of the image to scale in size by the same amount, leaving the peak surface flux unaffected. Uncertainties in the cross section σ or the coma expansion velocity v_c have the same effect. This is a consequence of the $n_c \sim r_c^{-2}$ density profile in the coma: for $\bar{n}_c = f \cdot n_c$, $\bar{r}_c = f \cdot r_c$, $\bar{z} = f \cdot z$, it follows that $\bar{n}_c(\bar{r}_c)\, d\bar{z} = n_c(r_c)\, dz$. Thus, the angular size of a comet in x-rays is mainly determined by the ratio Q/Δ. (iii) The coma is thin (transparent) for x-rays but in its inner part usually thick with respect to charge transfer: $\hat{s}_x$ is independent of Q down to $Q_{min} = 2.8 \times 10^{25}$ s$^{-1} \cdot R_n$, for a nucleus with radius R_n (in km). Thus, for $Q > Q_{min}$ [that is, for all comets in our sample (Table 1)], $\hat{s}_x$ is limited by the energy supply of the highly charged heavy ions in the solar wind; it scales with the product of the density and velocity of these ions, which should exhibit an r_h^{-2} dependence on average. (iv) The peak surface flux at $\varphi = 0°$ and r_h = 1 AU is: $\hat{s}_x = v_{sw} n_{sw} f_h f_y E_{ion} = 1.3 \times 10^{-5}$ erg cm^{-2} s^{-1} = 8.6 × 10^{-14} erg cm^{-2} s^{-1} arc min^{-2}. (v) Because the emission volume is not spherically symmetric and self-absorption of x-rays in the coma is negligible, $\hat{s}_x$ increases with φ, reaching a maximum at $\varphi = 90°$ which exceeds $\hat{s}_x$ at $\varphi = 0°$ by a factor of about 2. We define $\hat{s}_x$ as the peak surface flux observed at $\varphi = 0°$.
30. We note that "resolved" in this context is different from the usual definition, which compares the observed brightness distribution with the point spread function of the imaging device.
31. The $\tilde{s}_x/\hat{s}_x$ correction does not depend on the solar wind parameters n_{sw}, v_{sw}, f_h or on f_y or E_{ion}. The only parameters that affect the correction curves (Fig. 6A) are the expansion velocity $v_c(r_c)$ [and thus the density profile $n_c(r_c)$] and the cross section σ. The fact that in all cases the correction leads to a comparable $\hat{s}_x$ (when normalized to the same r_h) indicates that the assumed values of $v_c(r_c)$ and σ are not unreasonable (*37*). Because the peak surface fluxes are also close to the value expected for the chosen solar wind parameters, all assumptions are self-consistent.
32. N. C. Wickramasinghe and F. Hoyle, *Astrophys. Space Sci.* **239**, 121 (1996).
33. W. H. Ip and V. Chow, *Icarus*, in press.
34. R. Bingham, J. M. Dawson, V. D. Shapiro, D. A. Mendis, B. J. Kellet, *Science* **275**, 49 (1997).
35. T. G. Northrop, C. M. Lisse, M. J. Mumma, M. D. Desch, *Icarus* **127**, 246 (1997).
36. J. Brandt, C. Lisse, Y. Yi, *Bull. Am. Astron. Soc.* **189**, 25.05 (1996).
37. The coma expansion velocity v_c depends on the heliocentric distance r_h. The dependence, however, is too weak to alter the $\tilde{s}_x/\hat{s}x$ correction significantly: from r_h = 1 AU to r_h = 2 AU (the range of our sample), v_c decreases by less than 10% [for example, H. L. F. Houpis and D. A. Mendis, *Moon Planets* **25**, 95 (1995)].
38. H. U. Zimmermann *et al.*, *EXSAS User's Guide*, MPE Report 257 (ROSAT Scientific Data Center, Garching, Germany, ed. 4, October, 1994).
39. *The International Comet Quarterly*, D. Green, C. Morris, S. Nakano, Eds. (Smithsonian Astrophysical Observatory, Cambridge, MA, 1989–91), vols. 11–13.
40. For the obs. no. 1 through 6, the gas production rates were determined from log $Q_{H_2O(q)}$, the water production rates at perihelion, and the dependence of Q_{OH} on r_h (*25*) (for obs. no. 6, the average slope −2.7 was used). For obs. no. 8, Q_{r_h} was taken from (*24*), assuming $Q_{H_2O} \approx Q_{OH}$. This value was extrapolated to obs. no. 9 with the average −2.7 slope and compared with the log (Q_{r_h}/Δ) determined from the m_1 values of C/1996 B2 (see below); the weighted average of both values was taken. For the obs. no. 7 and 10, Q_{r_h} had to be determined from m_1, as no other information was available. This was done by assuming that the optical flux scales with $Q\ \Delta^{-2}\ r_h^{-2}$: the average value log $Q + 0.4 \cdot m_1 - 2 \log \Delta - 2 \log r_h$ was calculated from obs. no. 1 through 6 and 8, and used for the estimate log $Q \approx 31.6 - 0.4\, m_1 + 2 \log \Delta + 2 \log r_h$.
41. H. Kosai, *IAU Circular 5157* (1991).
42. We thank C. Lisse, B. Aschenbach, F. Favata, M. Freyberg, W.-H. Ip, K. Jockers, E. Rieger, H.-U. Schmidt, S. Snowden, R. Treumann, and R. Wegmann for interesting and stimulating discussions. We are grateful to P. Camilleri for supplying the optical image of comet C/1990 K1. The ROSAT project is supported by the German Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF/DARA) and the Max-Planck-Gesellschaft.

STATs and Gene Regulation

James E. Darnell Jr.

STATs (signal transducers and activators of transcription) are a family of latent cytoplasmic proteins that are activated to participate in gene control when cells encounter various extracellular polypeptides. Biochemical and molecular genetic explorations have defined a single tyrosine phosphorylation site and, in a dimeric partner molecule, an Src homology 2 (SH2) phosphotyrosine-binding domain, a DNA interaction domain, and a number of protein-protein interaction domains (with receptors, other transcription factors, the transcription machinery, and perhaps a tyrosine phosphatase). Mouse genetics experiments have defined crucial roles for each known mammalian STAT. The discovery of a STAT in *Drosophila*, and most recently in *Dictyostelium discoideum*, implies an ancient evolutionary origin for this dual-function set of proteins.

A large number of extracellular signaling polypeptides (>35) interact with specific cell surface receptors that trigger the activation of latent cytoplasmic transcription factors termed STATs. The STATs become phosphorylated on tyrosine, then dimerize by reciprocal SH2 phosphotyrosine interaction and enter the nucleus to regulate transcription of many different genes. The STATs were recognized as ligand-induced transcription factors in interferon (IFN)-treated cells and then in cells and tissues exposed to many other signaling polypeptides (*1*–*3*). A great deal of new information has accumulated about specificity in the activation of the STATs, about the functional domains of the proteins, and about the variety of their biologic functions both in development and in adults. Also, knowledge is beginning to accumulate about how STATs effect transcriptional changes.

The author is in the Laboratory of Molecular Cell Biology, Rockefeller University, New York, NY 10021, USA.

Mechanism of Activation

Seven mammalian STAT genes have been identified in three chromosomal clusters (*4*). The genes encoding Stats 1 and 4 map to a region of mouse chromosome 1 (equivalent to human chromosome 2, bands q12 to q33); Stats 3, 5A, and 5B map to a region of mouse chromosome 11 (human chromosome 12, bands q13 to q14-1); and Stats 2 and 6 map to a region of mouse chromosome 10 (human chromosome 17, bands q11-1 to q22). Stats 1, 3, 4, 5A, and 5B are between 750 and 795 amino acids long, whereas Stats 2 and 6 are ~850 amino acids long (*2*, *5*) (Fig. 1). Differential splicing leads to the production of a number of additional proteins, but the extent of variable splicing has not been widely explored. An important unresolved issue is whether

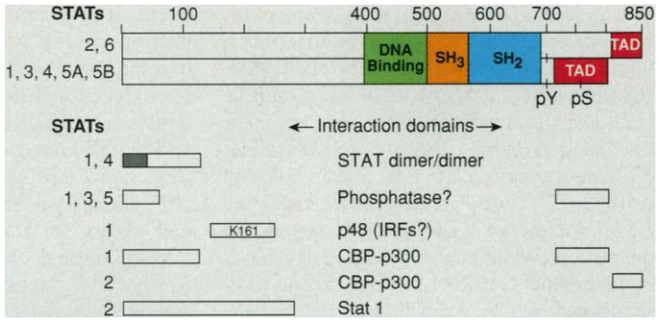

Fig. 1. Properties of STAT molecules. The STAT molecules are either ~850 (Stats 2 and 6) or 750 to 795 amino acids long (Stats 1, 3, 4, 5A, and 5B). Upper panel: The universally shared regions and their boundaries are indicated in color. Phosphotyrosine (pY) is present in all activated STATs; phosphoserine (pS) is present in activated Stats 1, 3, 4, 5A, and 5B. Transactivation domains (TAD) are shown in red. Lower panel: Protein interaction domains in the STATs listed at the left. The NH_2-terminal (leftmost) domain of Stats 1 and 4 is divided; the dark box indicates that removal of 40 residues of Stat4 destabilizes dimer-dimer interactions in that molecule.

any more family members exist.

Phosphorylation on a single tyrosine located around residue 700 in each protein is obligatory for STAT activation (*1*, *2*) (Fig. 1). Ligand-activated receptors that catalyze this phosphorylation include receptors with intrinsic tyrosine kinase activity [for example, epidermal growth factor (EGF), platelet-derived growth factor (PDGF), and colony-stimulating factor–1] as well as receptors that lack intrinsic tyrosine kinase activity but to which Janus kinases (JAKs) are noncovalently associated (*3*, *6*). The mammalian JAKs include four large tyrosine kinases (~1200 amino acids) characterized by a COOH-located kinase and a neighboring domain that resembles but is not an active kinase. Receptors to which JAKs are bound are often referred to as cytokine receptors. Their ligands include IFN-α, -β, and -γ; interleukins (IL) 2 to 7, 10 to 13, and 15; and erythropoietin, growth hormone, prolactin, thrombopoietin, and other polypeptides.

Ligand-mediated dimerization of either type of receptor is believed to result in reciprocal tyrosine phosphorylation, and consequent activation, of the intrinsic or the attached kinase (*6*). Phosphorylation of the kinase is the first of three tyrosine phosphorylations culminating in STAT activation (Fig. 2). The activated JAKs phosphorylate tyrosine sites on the cytoplasmic tail of the receptor that serve as docking sites for the SH2 domains that occur in all the STATs. The receptor-bound STAT is then phosphorylated on tyrosine. This chain of events was first shown for the IFN-γ receptor (*7*) and has since been shown for a variety of other receptors (*2*, *3*, *6*, *7*).

The activated JAK protein kinases do not seem to have specificity for a particular STAT substrate. For example, different receptors can activate the same STAT molecule through phosphorylation of the same tyrosine site, even though they activate distinct JAKs (*6*, *8*). Recombinant receptor molecules with different JAK binding sites but with the same STAT docking sites activate the same STAT (*8*, *9*). Also, STAT docking sites can be added to heterologous receptors, allowing docking and STAT activation by the ligand specific to the extracellular domain of the recombinant receptor (*10*). Finally, exchange of SH2 domains between Stats 1 and 2 reversed the specificity of receptor activation of the recombinant molecules (*11*). Thus, the initial specificity for most STAT activation appears to be determined by specific interactions between STATs and receptors (Fig. 2A).

This relatively simple picture for most STAT activations may understate the complexity of STAT-receptor-kinase interactions because in several cases more complicated interactions occur (Fig. 2, B and C). At least some STAT binding to receptors that have not bound ligand has been reported (*12*). This implies binding to receptor sites other than phosphotyrosine residues. The role of this binding in STAT activation is not yet clear (*3*). At the IFN-α receptor both Stat1 and Stat2 become activated, but Stat1 requires the presence of a tyrosine-phosphorylated Stat2 to become phosphorylated, whereas Stat2 can dock and be phosphorylated on its own (*13*). The NH_2-terminal region of Stat2 as well as the phosphotyrosine of Stat2 are required to activate Stat1 (*3*, *12*) (Fig. 2B). Stat5 can be phosphorylated at the IL-6 receptor apparently without receptor tyrosine docking sites (*14*); it appears to interact directly with a domain of JAK2 (Fig. 2C). Further work on STAT activation may well reveal other such interactions.

Stats 1, 3, and 5 can be activated by various receptor tyrosine kinases (RTKs), but much less is known about the specificity of RTK activation of STATs. The enzymatic activity of the receptor is required (*15*), and tyrosine residues outside the catalytic center of the EGF receptor are known to be required for activation of Stats 1 and 3 (*16*, *17*). Association of the STAT molecules with these receptor phosphotyrosine residues is implied but has not been directly demonstrated. The activated EGF receptor kinase can phosphorylate the activating sites on STATs in vitro (*18*). The EGF receptor is composed of members of a large protein family (*19*). Differential activation of Stats 1, 3, and 5 by EGF in various cell types might reflect the use of different members of this receptor family with different affinities for

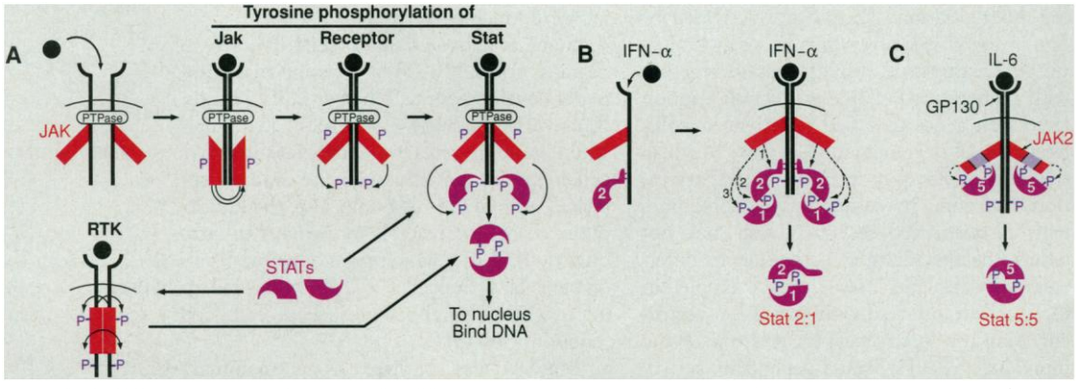

Fig. 2. STAT activation mechanisms. (**A**) Generalized mechanism of STAT tyrosine phosphorylation by either a JAK or an RTK. (**B**) Stat2 interaction with nonliganded IFN-α receptor chain [IFNAR-2 chain (*12*)] followed by ligand-mediated receptor assembly. Three phosphorylations catalyzed by the JAK then ensue (*1*–*3*): receptor tyrosine phosphorylation, Stat2 phosphorylation, and finally Stat1 phosphorylation. (**C**) Recruitment of Stat5 to JH2 domain (purple stippled) of JAK2 receptor associated with GP130 transmembrane receptor chain. JH2 is one of seven conserved domains within the JAKs and is a pseudo-kinase domain (*3*, *6*).

the various STATs. JAK activation also occurs after treatment of cells with growth factor (3). However, this JAK activation is not required for activation of Stat1 by EGF or PDGF (16).

Very soon after becoming phosphorylated on tyrosine, STAT molecules form homo- or heterodimers dependent on intermolecular SH2-phosphotyrosine interactions (2). Phosphopeptides containing the Stat1 or Stat2 tyrosine phosphorylation sites can prevent Stat1 dimerization (20). Dimers or heterodimers, but not monomers, are competent to bind DNA. The known DNA binding heterodimers are Stat1:2 (strong binding requires the joint presence of another protein, p48) and Stat1:3 (2). STATs that form homodimers that bind DNA include Stats 1, 3, 4, 5 (Stat5A and 5B interact in a manner equivalent to a heterodimer), and 6 (2, 3, 6). Stat2:2 dimers form sparingly in the absence of Stat1 and bind DNA weakly (21), as do Stat2:3 heterodimers (22). It is not known whether STATs dimerize on the surface of the receptor kinase complex or after release.

STAT Functions and Origin

Because more than 35 different polypeptide ligands are known to activate one or more different STATs, this group of proteins is implicated in a wide variety of biologic events. Three of the known STATs have a narrow activation profile: Stat2 is activated only by IFN-α; Stat4 in lymphocytes is activated by IL-12 and IFN-α (23); and Stat6 is activated only by IL-4 and IL-13, which share a receptor chain (2, 5, 6). In contrast, Stats 1, 3, 5A, and 5B are activated by many different ligands, which raises the question of how these STATs participate in specific biological responses.

Several broad answers to this question can be given. Multiple receptors that activate the same STAT are usually not present on the same cell. For example, receptors for IL-2, erythropoietin (Epo), thrombopoietin, prolactin, and growth hormone—all of which activate Stat5—are present on different cells during differentiation. Thus, Stat5 activated by Epo in red blood cell precursors, or by prolactin in breast tissue, would operate together with a different set of preexisting transcription factors on genes that had experienced a different set of developmental events. Much of the early exploratory work on STAT activation was done in cultured cells, sometimes with reconstructed receptors, and may not reflect the specificity of activation in differentiated cells and tissues. For example, in COS cells transfected with expression vectors encoding the leptin receptor and one of the individual STATs, ligand-dependent activation of Stats 3, 5, and 6 was observed (24). However, in the hypothalamus of mice (where leptin is thought to act physiologically), only Stat3 was activated after injection of leptin, although growth hormone elicited activation of Stat5 (25, 26). Thus, there must be cell specificity in the assembly of receptors or associated proteins that leads to differential STAT activation relative to that in cultured cells. Even if two ligands activate the same STAT or set of STATs, they may do so quantitatively differently for different periods of time, thus contributing to different transcriptional outcomes. Such quantitative variations can control physiologic decisions. For example, evidence from genetic studies of invertebrate development shows that relatively small changes in enzymes of the Ras-MAP (mitogen-activated protein) kinase pathway (and, presumably, in transcription affected by that pathway) can result in severe developmental defects (27, 28). Finally, differential splicing of the STAT genes is known to occur, although its frequency is not known (2, 3). At least a dozen proteins arise from differently spliced messenger RNA (mRNA) from the seven genes. Differential protein processing also produces modified Stat5 proteins (29). These different proteins could obviously have distinct transcriptional effects.

Selection of cell lines resistant to IFN action (30) aided in the proof that Stat1 was required for the action of IFN-α and IFN-γ and that Stat2 was required for the action of IFN-α but not IFN-γ (1–3). All of the STAT genes have now been deleted by homologous recombination in mouse embryonic stem cells, and the effects on mice developed from the mutant cells have been studied (Table 1). Each of the STATs has a crucial function in mice. Mice with no Stat1 have no innate response to either viral or bacterial infection, presumably because the first line of defense against potential pathogens requires the IFN response (31, 32). However, $Stat1^{-/-}$ animals kept in a pathogen-free environment are apparently normal and capable of reproducing. Thus, despite results in cell culture and some tissues that show at least modest Stat1 activation by a number of growth factors, there appears to be no obligatory role for Stat1 in mouse development (2, 3).

Stat4 is activated in T cells in vitro in response to IL-12 (33), a cytokine that stimulates development of T helper 1 (T_H1) cells. Removal of the Stat4 gene resulted in animals whose immunocytes were deficient in T_H1 cell function, but otherwise the animals appeared normal (34, 35). Likewise, the knockout of Stat6—activatable in vivo and in vitro only by IL-4 (5), which promotes the development of T helper 2 (T_H2) cells—led to mice whose lymphocytes lacked T_H2 cell function (36–38).

Stat5A was purified as a mammary growth factor (39) that binds to the β-casein gene promoter in cells stimulated with prolactin. It apparently has a role in milk production. Stat5B is more than 90% identical in amino acid sequence to Stat5A (40) and is produced at least as widely as Stat5A in tissues, including breast tissues. Stat5A (41) and Stat5B knockouts (42) have been examined, and each mutant shows a highly specific phenotype. $Stat5A^{-/-}$ animals are normal except for the inability of females to develop normal breast tissue and to lactate. In $Stat5A^{-/-}$ animals, Stat5B is not well activated in defective breast tissue, which may relate to the inability of Stat5B to complement the absence of Stat5A. The $Stat5B^{-/-}$ animals have a different phenotype related to activation of Stat5 by growth hormone, which is known to cause sexually dimorphic responses in the liver. $Stat5B^{-/-}$ male mice grow slowly and have serum levels of liver-produced proteins that are characteristic of female mice. However, $Stat5B^{-/-}$ females can lactate (42).

Thus, five of the seven known STATs have a demonstrated role in adult cell types that is sharply circumscribed and related to a role identified by the in vitro tests that led to their discovery. In contrast, Stat3 knockouts (43) die before gastrulation, probably never forming mesoderm. This outcome is in accord with the expression of Stat3 in the visceral endoderm, the function of which is required for gastrulation (44). Stat2 knockouts also do not reach term, but the stage at which embryos die has not been determined (29).

Overall, the findings with STAT deletions in mice are compatible with quite different interpretations: (i) The STATs were duplicated (only to a limited extent) in the evolution leading to mammals and now serve a relatively small number of specialized functions. (ii) Combinations of STATs may contribute to many regulatory events, and this possibility is not revealed in single knockouts. (iii) Additional STAT molecules remain to be discovered that serve equally specialized roles in cell types other than immunocytes, breast epithelium, and liver.

Table 1. Results of targeted disruptions of mouse STAT genes.

Targeted gene	Phenotype	Reference
Stat1	No innate response to viral or bacterial infection	(31, 32)
Stat2	Early embryonic lethal	(29)
Stat3	Early embryonic lethal; no gastrulation	(38)
Stat4	No T_H1 cell function	(34, 35)
Stat5A	No breast development or lactation	(41)
Stat5B	No breast development or lactation	(42)
Stat6	No T_H2 cell function	(37–39)

Identification and preliminary characterization of the function of the first invertebrate STAT from *Drosophila*, Stat92E (*45*, *46*), could lend weight to any of these conjectures. An early embryonic role for this protein is revealed by the expression pattern of the mRNA for Stat92E, its participation in specific early embryonic gene expression (stripe 3 expression of *even-skipped*), and the fact that the Stat92E null allele *marelle* is an embryonic lethal (*45*–*47*). Thus, Stat92E could have an early role similar to that of Stats 2 or 3, or it could have a function in flies for which an equivalent in mammals exists but has not yet been detected. Two other nuclear factors in *Drosophila* cells that bind the Stat92E DNA binding site were observed, and these factors contain phosphotyrosine (*45*, *48*).

If early evolutionary development of a gene or pathway hints at a possibly wider range of such mammalian proteins, then the recent findings of Williams and colleagues are of crucial importance (*49*). These workers have cloned the gene encoding the DNA binding protein required in the induction of pre-stalk cell differentiation by DIF (differentiation-inducing factor) in *Dictyostelium discoideum*. DIF is a lipophilic molecule (a chlorinated hexaphenone) that is membrane soluble, presumably enters the cell to act, and thus resembles the activators of the steroid superfamily more than it resembles polypeptides as gene activators. The newly cloned DIF-induced DNA binding protein is ~700 residues long, contains phosphotyrosine in its active form, and exhibits a substantial homology to STATs in the SH2 region in particular, as well as a more than chance homology to STATs in the DNA binding domain. If this *Dictyostelium* protein truly represents a STAT molecule, then the evolutionary path of early STAT genes to mammals is very long and could easily include the development of a great many STATs and STAT-like molecules.

STAT DNA Binding and Transcriptional Activation

Many of the DNA binding studies of the STATs have used oligonucleotides chosen to give maximum binding. Selection of optimum binding sites for Stats 1, 3, 4, and 92E all resulted in the recovery of the consensus oligonucleotide sequence, TTCC(C or G)GGAA (or generically TTN5AA, where N represents any nucleotide) (*50*, *51*). Only Stat6, which binds optimally to TTN6AA sequences, prefers different optimal binding sites (*52*–*54*). However, the natural sites from genes that are regulated in response to particular ligands show clear preferential binding affinities to the different STATs (*2*, *39*, *50*–*53*). Thus, selective gene activation by the various STATs could be attributable to differential STAT dimer binding to DNA.

In a study that involved swapping various amino acid segments between the closely related Stat1 and Stat3 sequences and testing binding to DNA sites that are differentially recognized by Stats 1 and 3, the region between residues 400 and 500 was found to confer DNA binding specificity (*50*) (Fig. 1). Moreover, mutations of highly conserved residues (Glu-Glu to Ala-Ala or Val-Val-Val to Ala-Ala-Ala) within the same region resulted in molecules that were inducibly tyrosine-phosphorylated, dimerized, and translocated to the nucleus, but failed to bind DNA. Mutations in homologous amino acids of Stat6 (*52*) gave similar results, which strongly suggested that residues 400 to 500 confer DNA binding specificity (*52*) in all STATs. However, these experiments did not define the boundaries of the residues involved in DNA binding, nor did they prove that residues in this region actually contact DNA. STAT molecules that are competent to become phosphorylated and dimerize but fail to bind DNA should act as dominant negative mutations and have in fact been reported to do so (*52*, *55*).

The DNA binding discussed so far used single DNA binding sites. Cooperative binding to neighboring sites of two (or more) STAT dimers has also been established. For example, closely spaced Stat1 or Stat5 DNA binding sites exist in natural promoters for a chemokine (MIG) gene (*56*) and a gene encoding an hepatic serine protease inhibitor (*57*). STAT-DNA complexes with these oligonucleotides migrated more slowly than did dimeric STAT-DNA complexes, and these higher order complexes are most likely composed of two interacting STAT dimers. Partially purified and tyrosine-phosphorylated Stat 1, 4, or 6 obtained from baculovirus-infected cells coinfected with a JAK produced footprints embracing neighboring Stat1, Stat4, or Stat6 sites (*51*). Many of these individual sites bound STAT proteins weakly or not at all. Deletion of the first 40 amino acids of Stat1 abolished this cooperative interaction.

The same phenomenon was also uncovered in a quite different set of experiments (*58*). Purified Stats 1α or 1β or a proteolytically stable Stat1 fragment (residues 132 to 713) could be correctly phosphorylated in vitro, and all three of these forms bound about equally to single DNA sites. In a test of the stability of DNA-protein interaction, the difference between strong binding sites and weak binding sites was attributed to the "off time" of the complex (the half-time of disappearance of an already formed complex with labeled DNA when challenged with an unlabeled oligonucleotide). When closely spaced double sites were used, two distinct differences between the full-length protein and the truncated protein were uncovered. First, Stat1α or 1β formed the dimer-dimer complex, whereas the NH_2-terminally truncated protein did not (Fig. 1). Second, the stability of the dimer-dimer interaction even on weak binding sites exceeded that for binding to single strong sites. Thus, dimer-dimer interactions on weak STAT DNA binding sites that are mediated by an NH_2-terminal domain may be of considerable physiological importance and should be searched for in promoter studies of STAT-activated genes.

Dimers of STATs bound to DNA also interact with several other classes of DNA binding proteins. The IFN-α–induced interferon-stimulated gene factor–3 (ISGF-3) is composed of a Stat1:2 heterodimer and a protein termed p48 that was found to be a member of the interferon regulatory factor–1 (IRF-1) family of proteins, of which at least six members are now known (*59*). The DNA contact sites for p48 and Stat1 are only five or six base pairs apart (*60*) (Fig. 1). A protein-protein contact region between p48 and Stat1 that is functionally important has been established (*61*). Mutation of a single residue ($Lys^{161} \rightarrow Ala$) in Stat1 in the p48 contact domain renders Stat1 incapable of participating in the IFN-α response but leaves it unaffected in the IFN-γ response, which does not require p48.

The transcription factor c-Jun was found to interact with activated Stat3, and Stat3 supplemented the transcriptional activation capacity of c-Jun in a transfection assay (*62*). In the human immunoglobulin heavy chain ε promoter, neighboring binding sites for the transcription factor C/EBPα and Stat6 are required for transcription (*63*), and interaction between these proteins on this promoter has been described (*52*).

A direct and apparently physiologically functional interaction has been reported for Stat5 and the glucocorticoid receptor (GR) (*64*). Maximal β-casein production by lactating mammals requires both prolactin and glucocorticoid induction. The promoter of the β-casein gene has both Stat5 sites and GR binding sites. Tyrosine-phosphorylated Stat5 and GR can be coprecipitated from prolactin- and glucocorticoid-treated cells, making a functional protein-protein interaction between these two proteins highly likely.

Still another case of STAT interaction with other DNA binding proteins occurs in the I-CAM promoter (*65*). (I-CAM is a cell adhesion molecule that incidentally serves as the major rhinovirus receptor.) This promoter has adjacent DNA binding sites for activated Stat1 and the transcription factor SP1. Incubation of this promoter with cell extracts containing activated Stat1 and SP1 formed DNA-protein complexes that contained both proteins.

It seems likely that interactions between other transcription factors and STATs will be found. For example, the c-Fos promoter contains neighboring binding sites for multiple transcription factors including SRF, TCF, AP1, CREB [cyclic adenosine monophosphate (cAMP) response element–binding protein], and a STAT, most likely Stat3 (66–68). This promoter in transgenic animals requires the simultaneous presence of at least four of these sites to give a successful transcriptional response to natural inducers in the animal, or to EGF and PDGF in culture. Interaction of such a cluster of proteins in a complex has been termed an enhanceosome (69).

Activation-Inactivation Cycle

STAT activation, as a rule, is transient, which suggests that STATs are either dephosphorylated by a protein tyrosine phosphatase or destroyed. Many transcriptionally active proteins—for example, c-Fos, c-Jun, and p53—in fact have short half-lives (70–72). Prolongation or non–ligand-dependent activation of STATs can be observed in cells treated with pervanadate (peroxide and vanadate), indicating a role for phosphotyrosine phosphatases in the activation-inactivation cycle (73, 74). However, it is not known whether one or all of the three ligand-dependent tyrosine phosphorylations (the kinase, the receptor, and the STAT) in the JAK-STAT pathway (Fig. 2) is prolonged by treatment of cells with pervanadate.

Inhibitors of proteosome activity (75) also prolong the activation of Stat1 in response to IFN-γ (76, 77). However, PDGF receptors and T cell receptors are removed from the cell surface in a proteosome-dependent manner (78, 79). The prolonged Stat1 activation occasioned by the proteosome inhibitor depends on a continued signaling from the receptor, because treatment of cells with staurosporine (an inhibitor of tyrosine phosphorylation) caused loss of the proteosome-supported DNA binding activity (77). Furthermore, the times required for maximum IFN-γ–stimulated activation and translocation of nuclear Stat1 and for staurosporine-induced depletion of DNA binding were similar (20 to 30 min), suggesting a total activation-inactivation cycle time of ~20 min for any given STAT molecule. Continued receptor signaling leads to a balance of accumulation and removal within this time (20 to 30 min), and the prolonged presence (2 to 3 hours) of activated Stat1 represents a state in which activation and nuclear entry are in equilibrium with removal by dephosphorylation. When signaling from the surface gradually dies out, the activated nuclear molecules disappear. When newly synthesized ^{35}S-labeled Stat1 was followed throughout an IFN-γ treatment cycle for 4 hours, little or no Stat1 protein was lost (±10%) even though as much as 25 to 30% was in the nucleus at 20 min, implying that the Stat1 inactivation cycle occurs by dephosphorylation rather than destruction of activated molecules (77).

Various regions of Stats 1, 3, and 5 appear to be required for the activation-inactivation cycle. The removal of the COOH-terminus (~50 amino acids) of Stats 3 and 5 results in a protein that in vivo can be activated, dimerize, and bind DNA specifically (80–82). Although wild-type Stat3 or Stat5 remains activated for 3 to 4 hours or less, the truncated molecule remains phosphorylated and capable of binding DNA for 12 hours or more. A similar result was found for Stat1, but the prolongation of activation was caused by deletion of a different section of the protein. Removal of the COOH-terminus of Stat1 does not influence the time course of this activation-inactivation cycle. However, removal of the NH_2-terminal 60 amino acids produced a molecule that remains tyrosine-phosphorylated for 12 to 24 hours, much longer than the wild-type protein (2 to 3 hours) (83). Because Stats 1, 3, and 5 are otherwise so similar in their construction, these disparate results are at the moment confusing. However, it appears that the rate of dephosphorylation of already activated molecules will have an important effect on the final transcriptional output from activated STATs.

Trans-Activation by the STATs: Domains That Associate with Other Nuclear Proteins

A consistent theme has emerged from studies of the functional domains of transcription factors that, like the STATs, bind to DNA sites distant from the RNA polymerase II initiation sites (84): The transcription activation domains of each of these distant binding factors require contact with one or more proteins that integrate the activating potential of the distant binding factor with the RNA polymerase II and the general transcription factors. These integrating proteins are called coactivators, mediators, or TAFs (TATA binding protein–associated factors). When properly contacted by a distant DNA binding factor, the integrating proteins enhance the probability that RNA polymerase II will begin transcription at a particular site.

Stat1 is required for both IFN-α and IFN-γ transcriptional responses. U3 cells completely lack Stat1 and in U3 cells Stat1α (750 amino acids) restored both IFN-α and IFN-γ responsiveness. Stat1β, a natural splice variant that lacks the terminal 38 amino acids of Stat1α, restored IFN-α but not IFN-γ responsiveness (85). Thus, transcription activation by IFN-γ requires at least the terminal 38 amino acids of Stat1. Stats 1α, 3, 4, 5A, and 5B are all about the same length (~750 to 795 amino acids), and the COOH-terminus is required for full gene activation by all of these proteins (80–82, 86). The COOH-terminal amino acid sequences in these proteins are not highly conserved except for a four–amino acid segment, Pro-Met-Ser-Pro. Stat1 and Stat3 are phosphorylated on a single serine (Ser^{727} in both) within this segment (87, 88). Ser^{727} is required for maximal gene activation by Stats 1 and 3 in cells stimulated with IFN-γ or EGF, respectively (Fig. 1). Serine phosphorylation of Stat5 is also implied by a ligand-dependent change in gel migration consistent with serine phosphorylation (89). For transcriptional activation of Stats 1, 3, and 5, tyrosine phosphorylation is obligatory, and serine phosphorylation—which is also inducible through as yet unidentified pathways—is supplementary.

Independence of tyrosine and serine phosphorylation has been shown for Stat1 (90). The mutant $Ser^{727} \rightarrow$ Ala can be phosphorylated on tyrosine, and the $Tyr^{701} \rightarrow$ Phe mutant can be phosphorylated on serine. The IFN-γ–induced tyrosine phosphorylation occurs within 5 min or less, whereas the serine phosphorylation takes more than 10 min (90). Both phosphorylations occur on cytoplasmic Stat1 before it moves to the nucleus. The tyrosine phosphorylation occurs at the plasma membrane, but the serine phosphorylation most likely occurs in the cytoplasm of the receptor. Thus, although these phosphorylation events are both important for full transcriptional activity, they are controlled by distinct signaling pathways.

The IFN-α–induced transcription factor ISGF-3, composed of Stat1:2 heterodimer and p48 protein, can function with Stat1β, but the COOH-terminus of Stat2 is required for an active ISGF-3 (91). Stat6, which is ~850 amino acids long, also requires its COOH-terminal residues to activate transcription (52). The COOH-terminal residues in Stat2, but not Stat6, are particularly rich in acidic amino acids and might therefore interact with different proteins than the COOH-terminal residues of Stats 1, 3, 4, and 5. There is no detectable serine phosphorylation of Stat2 during its function as part of ISGF-3 (92).

The STATs interact with one of the large group of proteins that serve as bridges between transcription factors and RNA polymerase II (93). The COOH-terminus of Stat2, a region containing the trans-activation domain, interacts with a large nuclear protein (~2500 residues) called p300 (Fig. 1). The p300 protein binds viral oncoproteins from polyoma and adenoviruses (94). Another large protein with a similar sequence, termed CBP (CREB binding protein), binds serine-phosphorylated transcription factors such as CREB (95). The

CBP and p300 proteins appear to behave similarly in all reported binding assays, and their role may be to relax chromatin near transcription start sites (96). The COOH-terminal segment of Stat2 bound to the same region of p300 as did the CREB protein. Moreover, the Stat2 segment cooperated with p300 in a cotransfection analysis to increase transcription rates.

Stat1 also interacts with CBP or p300 in at least two different sites on each molecule (97). The NH_2-terminus (within 130 amino acids) of Stat1 binds to the CREB site, and the COOH-terminus of Stat1, which contains the transactivation domain, binds to a domain of CBP or p300 that also binds to the E1A oncoprotein of adenovirus. The Stat1 interactions with CBP or p300 offer a possible explanation for two observed physiologic effects. First, the antiviral effect of IFN decreases early adenovirus mRNA formation that is dependent on adenovirus E1A protein. This is possibly explained because activated Stat1 would compete with E1A for binding to CBP (97). Second, transcription of the gene encoding the macrophage scavenger receptor is dependent on several transcription factors including ETS and AP1 (98). This gene is activated in cells treated with polypeptides that lead to serine phosphorylation of AP1 and ETS. These factors in turn are thought to require CBP, p300, or both for integrating their transcriptional stimulus. IFN-γ, which activates large amounts of Stat1, decreases the AP1- and ETS-mediated transcriptional increase, possibly by competing for CBP, p300, or both.

Conclusion

Since the original description of the STAT molecules less than 5 years ago, much has been accomplished to reveal the biologic and biochemical potential of this protein family. Nevertheless, a great many unanswered questions remain. For example, most studies have been concerned with the action of single extracellular signaling polypeptides, but cells in tissues are often exposed to multiple agents simultaneously. Increased cAMP can blunt the activation of STATs, and prior treatment of cells with granulocyte-macrophage colony-stimulating factor (GM-CSF) can prevent IL-6–induced activation of Stat3 (99). The physiologic outcomes of such interrelations are obviously of great importance. An appealing aspect of the STAT family of transcription regulators remains their relative simplicity and directness of action on transcription. Detailed studies on activated proteins that come and go within an hour or two and drastically change transcription rates will undoubtedly help to illuminate how transcriptional initiation is regulated.

REFERENCES AND NOTES

1. J. E. Darnell Jr., I. M. Kerr, G. R. Stark, *Science* **264**, 1415 (1994).
2. C. Schindler and J. E. Darnell Jr., *Annu. Rev. Biochem.* **64**, 621 (1995).
3. D. W. Leaman, S. Leung, X. Li, G. R. Stark, *FASEB J.* **10**, 1578 (1996).
4. N. G. Copeland *et al.*, *Genomics* **29**, 225 (1995).
5. J. Hou *et al.*, *Science* **265**, 1701 (1994).
6. J. N. Ihle, B. A. Witthuhn, F. W. Quelle, K. Yamamoto, O. Silvennoinen, *Annu. Rev. Immunol.* **13**, 369 (1995).
7. A. L. Greenlund, M. A. Farrar, B. L. Viviano, R. D. Schreiber, *EMBO J.* **13**, 1591 (1994).
8. F. Kohlhuber *et al.*, *Mol. Cell. Biol.* **17**, 695 (1997).
9. S. V. Kotenko *et al.*, *J. Biol. Chem.* **271**, 17174 (1996).
10. N. Stahl *et al.*, *Science* **267**, 1349 (1995).
11. M. H. Heim, I. M. Kerr, G. R. Stark, J. E. Darnell Jr., *ibid.*, p. 1347.
12. X. Li, S. Leung, I. M. Kerr, G. R. Stark, *Mol. Cell. Biol.* **17**, 2048 (1997).
13. S. Leung, S. A. Qureshi, I. M. Kerr, J. E. Darnell Jr., G. R. Stark, *ibid.* **15**, 1312 (1995).
14. Y. Fujitani *et al.*, *Oncogene* **14**, 751 (1997).
15. O. Silvennoinen, C. Schindler, J. Schlessinger, D. E. Levy, *Science* **261**, 1736 (1993).
16. D. W. Leaman *et al.*, *Mol. Cell. Biol.* **16**, 369 (1996).
17. P. J. Coffer and W. Kruijer, *Biochem. Biophys. Res. Commun.* **210**, 74 (1995).
18. F. W. Quelle *et al.*, *J. Biol. Chem.* **270**, 20775 (1995).
19. M. A. Lemmon and J. Schlessinger, *Trends Biochem. Sci.* **18**, 459 (1993).
20. K. Shuai *et al.*, *Cell* **76**, 821 (1994).
21. H. A. R. Bluyssen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5645 (1995).
22. C. M. Horvath and J. E. Darnell Jr., unpublished data.
23. S. S. Cho *et al.*, *J. Immunol.* **157**, 4781 (1996).
24. N. Ghilardi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6231 (1996).
25. C. Vaisse *et al.*, *Nature Genet.* **14**, 95 (1996).
26. C. Vaisse, unpublished data.
27. D. A. Wasserman, M. Therriery, G. M. Rubin, *Curr. Opin. Genet. Dev.* **5**, 44 (1995).
28. R. S. Kayne and P. W. Sternberg, *ibid.*, p. 38.
29. C. Schindler, personal communication.
30. S. Pellegrini, J. John, M. Shearer, I. M. Kerr, G. R. Stark, *Mol. Cell. Biol.* **9**, 4605 (1989).
31. M. A. Meraz *et al.*, *Cell* **84**, 431 (1996).
32. J. E. Durbin, R. Hackenmiller, M. C. Simon, D. E. Levy, *ibid.*, p. 443.
33. N. G. Jacobson *et al.*, *J. Exp. Med.* **181**, 1755 (1995).
34. M. H. Kaplan, Y. L. Sun, T. Hoey, M. J. Grusby, *Nature* **382**, 174 (1996).
35. W. E. Thierfelder *et al.*, *ibid.*, p. 171.
36. K. Shimoda *et al.*, *ibid.* **380**, 630 (1996).
37. M. H. Kaplan, U. Schindler, S. T. Smiley, M. J. Grusby, *Immunity* **4**, 313 (1996).
38. K. Takeda *et al.*, *Nature* **380**, 627 (1996).
39. H. Wakao, F. Gouilleux, B. Groner, *EMBO J.* **13**, 2182 (1994).
40. A. L.-F. Mui, H. Wakao, A. M. O'Farrell, N. Harada, A. Miyajima, *ibid.* **14**, 1166 (1995).
41. X. Liu *et al.*, *Genes Dev.* **11**, 179 (1997).
42. G. B. Udy *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7239 (1997).
43. K. Takeda *et al.*, *ibid.*, p. 3801.
44. S. A. Duncan, Z. Zhong, Z. Wen, J. E. Darnell Jr., *Dev. Dyn.* **208**, 190 (1997).
45. R. Yan, S. Small, C. Desplan, C. R. Dearolf, J. E. Darnell Jr., *Cell* **84**, 421 (1996).
46. X. S. Hou, M. B. Melnick, N. Perrimon, *ibid.*, p. 411.
47. R. Yan, H. Lou, J. E. Darnell Jr., C. R. Dearolf, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5842 (1996).
48. R. Yan and J. E. Darnell Jr., unpublished data.
49. T. Kawata *et al.*, *Cell* **89**, 909 (1997).
50. C. M. Horvath, Z. Wen, J. E. Darnell Jr., *Genes Dev.* **9**, 984 (1995).
51. X. Xu, Y.-L. Sun, T. Hoey, *Science* **273**, 794 (1996).
52. T. Mikita, D. Campbell, P. Wu, K. Williamson, U. Schindler, *Mol. Cell. Biol.* **16**, 5811 (1996).
53. H. M. Seidel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3041 (1995).
54. U. Schindler, P. Wu, M. Rothe, M. Brasseua, S. McKnight, *Immunity* **2**, 689 (1995).
55. K. Nakajima *et al.*, *EMBO J.* **15**, 3651 (1996).
56. N. B. Guyer, C. W. Severns, P. Wong, C. A. Feghali, T. M. Wright, *J. Immunol.* **155**, 3472 (1995).
57. P. L. Bergad, H.-M. Shih, H. C. Towle, S. J. Schwarzenberg, S. A. Berry, *J. Biol. Chem.* **270**, 24903 (1995).
58. U. Vinkemeier *et al.*, *EMBO J.* **15**, 5616 (1996).
59. Y. Kanno *et al.*, *Mol. Cell. Biol.* **13**, 3951 (1993).
60. S. A. Qureshi, M. Salditt-Georgieff, J. E. Darnell Jr., *Proc. Natl. Acad. Sci. U.S.A.* **92**, 3929 (1995).
61. C. M. Horvath, G. R. Stark, I. M. Kerr, J. E. Darnell Jr., *Mol. Cell. Biol.* **16**, 6957 (1996).
62. T. S. Schaefer, L. K. Sanders, D. Nathans, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9097 (1995).
63. S. Delphin and J. Stavenezer, *J. Exp. Med.* **181**, 181 (1995).
64. E. Stocklin, M. Wissler, F. Gouilleux, B. Groner, *Nature* **383**, 726 (1996).
65. D. C. Look, M. R. Pelletier, R. M. Tidwell, W. T. Roswit, M. J. Holtzman, *J. Biol. Chem.* **270**, 30264 (1995).
66. L. M. Robertson *et al.*, *Neuron* **14**, 241 (1995).
67. H. B. Sadowski, K. Shuai, J. E. Darnell Jr., M. Z. Gilman, *Science* **261**, 1739 (1993).
68. Z. Zhong, Z. Wen, J. E. Darnell Jr., *ibid.* **264**, 95 (1994).
69. D. Thanos and T. Maniatis, *Cell* **83**, 1091 (1995).
70. M. Trier, L. M. Staszewski, D. Bohmann, *ibid.* **78**, 787 (1994).
71. I. Stancovski, H. Gonen, A. Orian, A. L. Schwartz, A. Ciechanover, *Mol. Cell. Biol.* **15**, 7106 (1995).
72. M. Oren, W. Maltzman, A. J. Levine, *ibid.* **1**, 101 (1981).
73. M. David, P. M. Grimley, D. S. Finbloom, A. C. Larner, *ibid.* **13**, 7515 (1993).
74. S. J. Haque, V. Flati, A. Deb, B. R. G. Williams, *J. Biol. Chem.* **270**, 25709 (1995).
75. K. L. Rock *et al.*, *Cell* **78**, 761 (1994).
76. T. K. Kim and T. Maniatis, *Science* **273**, 1717 (1996).
77. R. L. Haspel, M. Salditt-Georgieff, J. E. Darnell Jr., *EMBO J.* **15**, 6262 (1996).
78. C. Cenciarelli *et al.*, *Science* **257**, 795 (1992).
79. S. Mori, K. Tanaka, S. Omura, Y. Saito, *J. Biol. Chem.* **270**, 29447 (1995).
80. R. M. Moriggl *et al.*, *Mol. Cell. Biol.* **16**, 5691 (1996).
81. C. Caldenhoven *et al.*, *J. Biol. Chem.* **271**, 31221 (1996).
82. D. Wang, D. Stravopodis, S. Teglund, J. Kitazawa, J. N. Ihle, *Mol. Cell. Biol.* **16**, 6141 (1996).
83. K. Shuai, J. Y. Liao, M. M. Song, *ibid.*, p. 4932.
84. RNA Polymerase II Transcriptional Machinery special issue, *Trends Biochem. Sci.* (1996).
85. M. Muller *et al.*, *EMBO J.* **12**, 4221 (1993).
86. A. L.-F. Mui, H. Wakao, T. Kinoshita, T. Kitamura, A. Miyajima, *ibid.* **15**, 2425 (1996).
87. Z. Wen, Z. Zhong, J. E. Darnell Jr., *Cell* **82**, 241 (1995).
88. Z. Wen and J. E. Darnell Jr., *Nucleic Acids Res.* **25**, 2062 (1997).
89. C. Beadling, J. Ng, J. W. Babbage, D. C. Cantrell, *EMBO J.* **15**, 1902 (1996).
90. X. Zhu, Z. Wen, J. E. Darnell Jr., unpublished data.
91. S. A. Qureshi, S. Leung, I. M. Kerr, G. R. Stark, J. E. Darnell Jr., *Mol. Cell. Biol.* **16**, 288 (1996).
92. C. Schindler, K. Shuai, V. R. Prezioso, J. E. Darnell Jr., *Science* **257**, 809 (1992).
93. S. Bhattacharya *et al.*, *Nature* **383**, 344 (1996).
94. R. Eckner *et al.*, *Genes Dev.* **8**, 869 (1994).
95. J. C. Chrivia *et al.*, *Nature* **365**, 855 (1993).
96. E. Pennisi, *Science* **275**, 155 (1997).
97. J. J. Zhang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15092 (1996).
98. A. E. Horvai *et al.*, *ibid.* **94**, 1074 (1997).
99. T. K. Sengupta, E. N. Schmitt, L. Ivashkiv, *ibid.* **93**, 9499 (1996).

100. I thank C. M. Horvath and U. Vinkemeier for critical reading of the manuscript. Supported in part by NIH grants AI32489 and AI34420.