This year marks the tenth anniversary of the recognition of the STAT proteins, named after their dual role as signal transducers and activators of transcription. These transcription factors are latent in the cytoplasm until they are activated by extracellular signalling proteins (mainly cytokines and growth factors, but also some peptides) that bind to specific cell-surface receptors. These extracellular-signalling proteins can activate various tyrosine kinases in the cell that phosphorylate STAT proteins. The activated STAT proteins accumulate in the nucleus to drive transcription. The duration and degree of gene activation are under strict regulation by a series of negatively acting proteins. This review concentrates on recent progress in studying these sets of proteins, highlighting important issues in which recent progress has been made or that still remain unresolved.

**Pathways leading to STAT activation**

As is the case for STAT proteins 1 and 2, which were discovered as targets of interferon activation, all the STAT proteins can be activated after one or more cytokines interact with their cognate receptors. The basic model for these cytokine pathways depends on a series of three tyrosine phosphorylations that are carried out by Janus kinase (JAK) proteins that are non-covaingly bound to specific receptors. Receptor dimerization or oligomerization leads to JAK (Box 1). Tyrosine phosphorylation on tyrosine residues, releasing their intrinsic catalytic activity. Tyrosine phosphorylation by activated JAKs of cytokine receptor cytoplasmic domains then provides binding sites for the Src-homology-2 (SH2) domain of the STAT proteins. The STAT proteins are then recruited to the JAKs, whereupon they are phosphorylated on a single tyrosine residue (around residue 700 of their 750-850-amino-acid sequence). Although the interactions and consequences of STAT binding to JAKs are the best studied, STAT1, STAT3, and STAT5 at least can also be activated by other receptor and tyrosine-kinase interactions (Box 1). Regardless of how STATs are tyrosine phosphorylated, it is clear that STAT–STAT interaction occurs immediately through reciprocal phosphotyrosine–SH2 interactions. STAT1, STAT3, STAT4, STAT5a, and STAT5b all form homodimers. STAT1 and STAT2, and STAT1 and STAT3 can also form heterodimers, depending on the nature and concentration of the activating ligand. In vitro tyrosine phosphorylation is accompanied by quantitative dimer formation, and there are no reports of monomeric tyrosine-phosphorylated STAT proteins. It seems possible, if not probable, that the dimeric (or higher-order) nature of the activating receptor is accompanied by near-simultaneous activation of two STAT molecules — one at each receptor in a complex — followed by STAT SH2–phosphotyrosine-mediated
INNATE IMMUNE RESPONSE

This is crucial during the early phase of host defence against infection by pathogens (such as bacteria and viruses), before the antigen-specific, adaptive immune response is induced.

CHAPERONE

Protein that mediates polypeptide folding or the assembly of another polypeptide-containing structure but does not form part of the completed structure or participate in its biological function.

HEPATOMA

Cancer of the liver.

JAKs are characterized by a carboxy-terminal catalytic domain and a related, but enzymatically inactive, adjacent pseudo-kinase or kinase-like domain. They also share five additional blocks of sequence similarity throughout the amino-terminal region. These sequence features define four members of the JAK family (Table 1). The draft sequences of the human and mouse genomes indicate that there are no other JAKs. Gene targeting in mice has shown that each enzyme has mostly non-redundant functions that can be largely — if not completely — explained as features of cytokine signalling (Table 1).

No JAK structure is yet available but the existing evidence indicates that this is a fruitful area for further study. Their kinase domains, in addition to an obvious role in catalysis, are also selective targets for inhibition by the suppressor of cytokine signalling (SOCS) proteins (see below). Moreover, the pseudo-kinase domains also seem to regulate catalysis negatively and might have a role in substrate recognition. Mutations in this region, at least in Drosophila melanogaster, can lead to leukaemia-like cell-proliferation diseases as a result of unopposed kinase activity. Sequence-based structure predictions have suggested that these proteins have an SH2 domain, although the function of this domain remains unclear. In fact, the conserved arginine that is found in all SH2 domains — which hydrogen bonds to phosphotyrosine residues in interacting proteins — is a histidine in tyrosine kinase 2 (TYK2) and has been mutated to an alanine in JAK1 without consequence. This raises the interesting possibility that this domain is a binding site for something other than phosphotyrosine that might regulate JAK activity.

The more proximal amino-terminal regions of JAK proteins are directly involved in selective interactions with cytokine receptors. In addition, as was first shown for the interferon-α (IFN-α) receptor, but has now been documented for several others, this interaction contributes a CHAPERONE function that facilitates the expression of the receptor–kinase complex at the cell surface. Interactions between some of these amino-terminal domains and the catalytic domain might also function in the regulation of enzymatic activity. Understanding how these long-distance interactions function at a mechanistic level, however, will require a JAK structure to be determined at atomic resolution.

Unphosphorylated cytoplasmic STAT proteins

The molecular state of the STATs in the cytoplasm before activation is not fully understood, and several possibilities have been described that need further exploration. There are clearly proteins that can bind STATs in the cytoplasm, although physiological roles for these interactions have yet to be detailed. STAT-interacting partners that might facilitate recruitment to receptors for enhanced signalling have been identified. One such protein is also a component of a RNA-polymerase II elongation complex and so might also facilitate STAT-dependent transcriptional activity after nuclear translocation. Furthermore, the dimerization of unphosphorylated STATs in the cytoplasm, possibly in association with additional cellular proteins, has also been reported and could be a common state for unphosphorylated STAT1 and STAT3. On the basis of the crystal structure of phosphorylated and unphosphorylated STAT1, these complexes might not have the same structure as the dimers that bind DNA (X. Chen et al., unpublished observations). STAT3 in HEPATOMA cells has been reported to exist mainly in association with large protein aggregates. Such complexes could provide a mechanism for sequestering STAT proteins before activation or be a platform that allows efficient access to plasma-membrane receptors.

Effects of crosstalk on STAT activation

The extent of STAT activation can be altered by prior exposure of cells to other stimuli that bind other receptors, commonly known as crosstalk. Because cells in the body often produce more than one cytokine or growth factor, crosstalk probably represents a physiologically important control for STAT activity.

Positive crosstalk. A well-defined example of positive crosstalk is the effect of type-I IFN (α and β) on signalling from the type-II IFN (γ) receptor and vice versa, and on receptors of the interleukin-6/glycoprotein-130 (IL-6/gp130) family. For example, pretreatment of cells with IFN-γ strongly augments a subsequent IFN-α response. This crosstalk can be at least partly explained by the increase in abundance of interferon regulatory factor 9 (IRF9; formerly called p48), the product of an
**Variations in mechanisms of STAT activation**

Tyrosine phosphorylation of signal transducers and activators of transcription (STAT) proteins at or around residue 700 occurs in response to cytokine receptors through Janus kinases (JAKs). However, at least several dozen receptors with intrinsic tyrosine kinase activity (RTKs) such as those for epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) seem to be able to mediate the activation of STAT proteins.46–121 Apparently, this activation can be direct (as in the case of STAT1 activation by PDGF receptor) or indirect. The latter case involves the recruitment of complexes of proteins to the phosphorylated RTK. Non-receptor tyrosine kinases (NRTKs), such as Src — the first tyrosine kinase to be discovered — are among the recruited proteins. STAT3 and Ssrc can interact independently and STAT3 probably becomes phosphorylated by Src on the EGF and PDGF receptors. Furthermore, it is clear that seven-transmembrane (7TM) receptors can, after binding their peptide or short polypeptide ligands, also activate STAT proteins.46–121,154 It has been proposed again that the tyrosine kinase involved is Src — or perhaps the JAKs become activated by associating with 7TM receptors155–157. STAT 1, STAT 3, STAT 4, STAT 5 and STAT 6 homodimerize. STAT1 and STAT2, and STAT1 and STAT3 can form heterodimers, and several STAT proteins can form tetramers (or potentially higher order complexes).

One final comment on STAT activation is needed: Direct recruitment of latent STAT proteins to activated cytokine receptors might be the most common and is certainly conceptually the simplest mechanism of cytokine activation, but pre-association of STAT1 and STAT2 with the interferon-α (IFN-α) receptor before ligand stimulation has also been described. Furthermore, it is known that, after IFN-α treatment, STAT 2 must be phosphorylated before STAT 1 at least in some cell types.121,155 So, in this case, the general model for phosphorysine-dependent recruitment might not apply. However, the exact role of pre-associated STAT 2 at the IFN-α receptor and that of induced receptor phosphorylation in the IFN-α pathway remains uncertain.125,126 STATpY, tyrosine-phosphorylated STAT; pS, serine phosphorylated.

Augmentation at the level of the receptor signalling complex has also been observed, possibly owing to the presence of multiple cytokine receptors within discrete plasma membrane domains known as LIPID RAFTS.25,26 Proximity might lead to signalling reinforcement when ligand-dependent activation of one receptor leads to the activation of a neighbouring receptor, possibly through cross-phosphorylation27 or down-modulation of the non-specific inhibitory effect on signalling by caveolin-1, a main constituent of lipid rafts.28

**Negative crosstalk.** A well-characterized example of negative crosstalk involves negative feedback by SOCS proteins (see below). These cytokine-inducible kinase inhibitors are fairly promiscuous, both in their effect on the induction of different cytokines and in their ability to inhibit distinct receptor–kinase complexes. Therefore, induction of SOCS proteins by one cytokine can inhibit other pathways owing to the already-enhanced abundance of an inhibitor. Negative crosstalk from non-STAT signalling pathways also occurs, although the underlying molecular mechanisms are less clear. For instance, stimulation of cyclic AMP, mitogen-activated protein kinases or glucocorticoid-dependent signalling pathways can lead to inhibition of subsequent cytokine-dependent stimulation of STAT phosphorylation.30–32 Such crosstalk presumably has a role in normal homeostasis and control of inflammation. Determining the extent and specificity of both positive and negative crosstalk will require further study.

**Structure of STAT proteins**

The modulatory nature of STAT proteins was first realized by sequence comparisons and mutagenesis studies, which showed that the carboxyl terminus was a TRANSACTIVATION DOMAIN (TAD), that an SH2 domain preceded the TAD and that the tyrosine residue that became phosphorylated lay between the two. Furthermore, the DNA-binding domain was in the centre of the molecule (Fig. 3a). These conclusions were confirmed and greatly extended by crystallographic studies of the core amino acids (residues ~130–710) of either dimeric STAT 1 or dimeric STAT 3 bound to DNA, which have very similar characteristics20,34 (Fig. 3a).

Beginning at residue 130, there is a four-stranded helical coiled coil that presents extensive possibilities for protein–protein interaction. Interactions with several proteins have been documented. A DNA-binding fold between residues 320 and 490 contains several β-sheets that are folded similarly to those found in the DNA-binding domains of the transcription factors nuclear factor κB (NF-κB) or p53. Contact with DNA is limited but occurs in both major and minor grooves. For example, in STAT1, N460 and K336 contact the major groove, and E421 contacts the minor groove. (N represents asparagine, K is lysine and E is glutamic acid.) A linker domain of highly conserved structure but unknown function follows from residues 490 to 580. Mutations within this domain affect the stability of DNA binding, which leads to a rapid off-rate and an inability to activate genes after IFN-γ induction.19 A classic SH2 structure
REVIEWS

The domains of JAK kinases (JAKs) 1–7, based primarily on sequence homology. The domains JH1–JH7 are regions of sequence similarity in the four known JAKs. Domain JH1 is the kinase domain and domain JH2 is the pseudo-kinase domain. The amino-terminal domains JH6 and JH7 contain sites that bind JAK to receptors. JH, JAK homology. Scale bar indicates amino acid residues.

Figure 2 | Janus kinase domain structure. The domains of JAKs 1–7, based primarily on sequence homology. The domains JH1–JH7 are regions of sequence similarity in the four known JAKs. Domain JH1 is the kinase domain and domain JH2 is the pseudo-kinase domain. The amino-terminal domains JH6 and JH7 contain sites that bind JAK to receptors. JH, JAK homology. Scale bar indicates amino acid residues.

Table 1 | Role of JAKs as revealed by gene-targeting in mice*

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Phenotype of null mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jak1</td>
<td>Perinatal lethality, probably caused by failure of cytokine signalling in neurogenesis; immunological impairments caused by failure of multiple haematopoietic cytokines</td>
</tr>
<tr>
<td>Jak2</td>
<td>Embryonic lethality caused by failure of erythropoiesis; additional immunological impairments caused by impaired cytokine signalling</td>
</tr>
<tr>
<td>Jak3</td>
<td>SCID caused by failure of cytokine signalling from γc-containing receptors</td>
</tr>
<tr>
<td>Tyk2</td>
<td>Increased pathogen susceptibility caused by impaired responses to interferon and IL-12</td>
</tr>
</tbody>
</table>

*Reviewed in REF 55 AND 6. II-12, Interleukin-12; Jak, JAK kinase; SCID, severe combined immunodeficiency; Tyk2, tyrosine kinase 2.

Nucleocyttoplasmic transport

Importing STAT proteins into the nucleus. Shortly after ligand-dependent tyrosine phosphorylation and dimerization, STATs accumulate in the nucleus. Such large protein complexes (~180 kDa for the STAT dimer) require facilitated transport into the nucleus. Binding of STAT1 to importin-α5, one of the subunits of the nucleocytoplasmic transport machinery, has been described and recent mutagenic studies established L407, K410 and K413 (within the DNA-binding domain) in STAT1 as crucial residues in the nuclear import of tyrosine phosphorylated STAT1 (REFS 45, 46). When the importin is bound, STAT dimers cannot bind DNA and, if the importin–STAT complex enters the nucleus, it might require DNA exploration to dislodge the importin. In at least some cells, some (unphosphorylated) STAT1 can enter the nucleus, and this is not affected by mutations of L407 (REF. 48). This is especially clear in genetically selected human cancer cells that lack STAT1 completely (U3A cells; box 2). Several messenger RNAs (mRNAs) that are required for STAT-directed apoptosis are not formed in these cells unless STAT1 is expressed. Even when STAT1 is expressed, however, the mRNAs in question are formed in the absence of STAT1 tyrosine phosphorylation; the cells then undergo induced apoptosis. Unphosphorylated STATs are hypothesized to interact with other transcription factors on DNA and to help stimulate transcription. The mechanistic details of how tyrosine phosphorylation and dimerization favor nuclear entry await crystallographic solution of the structures of unphosphorylated STAT proteins, and perhaps even crystallization of full-length unphosphorylated and tyrosine-phosphorylated STATs.

Nuclear export of STAT proteins. The export of STAT1 from the nucleus seems to depend on residues in the COILED-COIL DOMAIN and also in the DNA-binding domain. Whether residues in analogous positions are involved in nucleocytoplasmic transport of other STATs is not yet clear, nor is it known what structural state is assumed by STAT proteins in transit. Dephosphorylation of STATs occurs in the nucleus and is an important signal for export back to the cytoplasm. For STAT1 and STAT3, there is evidence that implicates TC45 (a nuclear tyrosine phosphatase) as a relevant STAT nuclear phosphatase. Cells that lack this enzyme retain tyrosine-phosphorylated STAT1 for much longer than normal cells, and overexpression of

COILED-COIL DOMAINS

A protein domain that forms a bundle of two or three α-helices. Short coiled-coil domains are involved in protein interactions, whereas long coiled-coil domains forming long rods occur in structural or motor proteins.

NF-κB

(Nuclear factor of κB). A widely expressed transcription factor that is activated by cellular stress and can induce the expression of numerous anti-apoptotic genes.
ubiquitin.

destruction by the addition of

that have been tagged for

degrading intracellular proteins

Protein complex responsible for

by the 26S proteasome.

target the protein for destruction

This often forms

covalently attached to

various proteins in various domains. The amino-terminal structure, the placement of which in the intact structure is undefined, also interacts with various partners, as does the carboxy-terminal transactivation domain, the structure of which is unknown. Modified with permission from REF 36 © 1998 American Association for the Advancement of Science, and from

Elsevier Science Ltd. CBP, CREB binding protein; IRF, interferon regulatory factor; Mcm, minichromosome maintenance; Nmi, N-Myc interactor; PIAS, protein inhibitor of activated STAT.

| STAT domain structure and protein binding sites. a | The core structure (amino acids ~130-712) shows binding of a STAT1 dimer to DNA and the location of binding sites of various proteins in various domains. The amino-terminal structure, the placement of which in the intact structure is undefined, also interacts with various partners, as does the carboxy-terminal transactivation domain, the structure of which is unknown. Modified with permission from REF 36 © 1998 American Association for the Advancement of Science, and from REF 36 © 1998 Elsevier Science Ltd. CBP, CREB binding protein; IRF, interferon regulatory factor; Mcm, minichromosome maintenance; Nmi, N-Myc interactor; PIAS, protein inhibitor of activated STAT.

b | STAT structure. STAT, signal transducer and activator of transcription. SH2, Src-homology-2 domain.

Figure 3

TC45 leads to dephosphorylation of STAT5 (REF 55).

However, TC45 has also been implicated in regulating cytoplasmic dephosphorylation events, such as the dephosphorylation of JAK1 and JAK3 (REF 56). Other phosphatases, such as SH2-containing phosphatase 1 (SH P1), SH P2 and protein-tyrosine phosphatase 1B (PTP1B) have also been implicated as cytoplasmic regulators of JAK or STAT phosphorylation55-59.

A general scheme, not yet extended to all STAT proteins, emerges for the nucleocytoplasmic transport of STAT1 after IFN treatment of cells: tyrosine phosphorylation and dimerisation favour importin binding, perhaps by conferring a change in conformation that exposes the region containing L407. Importin binding and GTP-dependent translocation then occurs44,48,50. Either because the nuclear tyrosine-phosphorylated STAT1 now has a conformation that is less favourable to proteins of the export machinery, such as chromosome region maintenance protein (Crm1 — a nuclear export protein)56, or because the phosphorylated STAT1 associates reversibly with DNA while exploring for favourable binding sites — whichocludes a necessary export signal45 — the tyrosine-phosphorylated STAT1 dimer remains nuclear until it loses its phosphorylated tyrosine residue, at which point it is returned to the cytoplasm. On the basis of staurosporin inhibition of further tyrosine phosphorylation, the cycle time (activation-inactivation) for an individual STAT1 molecule in EUPLOID fibroblasts is ~20 min (REF 52). This indicates that several cycles of STAT phosphorylation, nuclear migration, dephosphorylation and export, and re-phosphorylation and re-import occur during a full transcriptional response to cytokine stimulation. The duration of STAT phosphorylation (and therefore of transcriptional activity) is regulated by the balance of receptor-driven JAK catalytic activity and constitutive nuclear dephosphorylation40.

Much remains to be done to validate various points of this model, especially extending it to other cell types and to other STATs.

Negative regulators of STAT signalling

As mentioned above, and as will be made clear in the section on cell-biological integration of STAT activity, activation versus inactivation/inhibition of STAT proteins is crucial to their biological actions FIG 4.

Cyttoplasmic tyrosine phosphatases. There are several types of negative regulator of STAT proteins in the cell cytoplasm. Tyrosine dephosphorylation of receptor or kinase sites by SH P1, SH P2 or PTP1B limits further STAT tyrosine phosphorylation55-60. Humans with mutations in phosphatase recruitment sites on cytokine receptors and transgenic mice with mutations in these phosphatases confirm the importance of regulated phosphorylation55,60. For instance, absence of Ptp1B in mice leads to leptin hypersensitivity and enhanced Stat3 tyrosine phosphorylation; this is probably due to impaired jak1 dephosphorylation61.

SOC5 proteins block continued signalling. An entire subfield of STAT biology was ushered in when SOCS proteins were shown to prevent further receptor signalling by binding to receptor sites and/or JAK catalytic domains of STAT proteins (JABs), STAT-induced STAT inhibitors (SSI s)64-66 or cytokine-induced SH2 (CIS) protein67. The SOCS proteins, the genes for which are induced transcriptionally in response to cytokine stimulation, are recruited to active receptor complexes to cause inhibition. SOCS proteins can also cause protein turnover of the receptor through a ubiquitin-proteasome-mediated process59. The multiple functions of the large SOCS family provide a rich territory for studying the need for balanced STAT transcriptional output in development and in adult functions (TABLE 2).

For example, STAT4 induction by IL-12 is crucial to differentiation of T HELPER 2 (TH2) cells, whereas STAT6 induction by IL-4 is necessary for T HELPER 2 (TH2) cell induction. The levels of STAT activity in the two cell types is
REVIEWS

Box 2 | Transcriptional stimulation without phosphorylation

The first indication that signal transducers and activators of transcription (STAT) proteins might have a role in gene expression as unphosphorylated molecules came from experiments using genetically selected human cancer cells in culture that lacked STAT1 (U3A cells). Such cells did not undergo apoptosis upon challenge until STAT1 expression was restored. The U3A cells lacked a full component of CASPASES, which are present in the U3A cells that express STAT1. However, the apoptotic response did not require wild-type STAT1; a STAT1 Y701F mutant (in which Y is tyrosine and F is phenylalanine), which cannot be phosphorylated on tyrosine, still conferred the ability to respond to apoptotic signals. Subsequently, several other proteins involved in apoptosis were also found to be restored to U3A cells by the STAT1 Y701F mutant.

More recently, gene-array experiments using U3A cells (lacking STAT1) and cells with restored STAT1 were compared without IFN-γ stimulation. A set of genes was found in which the genes were not dependent on IFN-γ and whose mRNAs were present when STAT1 was added back to the U3A cells. One of these genes, termed low molecular weight polypeptide 2 (LMP-2), was studied in detail. STAT1 was constitutively bound to the promoter of this gene not stimulated by IFN-γ and in which no tyrosine phosphorylation of STAT1 could be found. The STAT1 was associated with the promoter, presumably because of its association with interferon regulatory factor-1 (IRF1) — a known binding partner of STAT1 — that was also present at the LMP-2 promoter, and can bind DNA on its own. It seems clear that unphosphorylated STAT proteins can have a role in transcription, even though it is evident that the main transcriptional stimulation by STATs follows tyrosine phosphorylation, dimerization and nuclear accumulation.

Apparently mediated by SOCS — SOCS1 is expressed at fivefold greater levels in T-1 cells than in T-2 cells, whereas T-2 cells express SOCS3 at a 23-fold higher concentration than T-1 cells. As another example, it is clear from studies of Socs1 knockout (Socs1−/−) mice — this mutation is perinatal lethal — that unregulated Stat1 activity is lethal, owing largely to the unopposed action of IFN-γ and IL-4. Animals that are null for both Socs1 and Stat1 or null for IFN-γ and Socs1 are resistant to the liver degeneration that is associated with a continuous high level of Stat1 activity in the Socs1-knockout animals. Similarly, deletion of the Socs2 gene leads to dysregulation of the insulin-like-growth-factor-1 (IGF-1) pathway, which causes gigantism. Loss of Socs3 results in embryonic lethality owing to impaired haemopoiesis and placental defects, further emphasizing the importance of negative regulation for proper cytokine action. Loss of Socs6 causes growth retardation owing to a requirement for this protein in the proper regulation downstream of insulin receptor substrate.

Nuclear regulators. The above-mentioned negative regulator functions operate in the cytoplasm. There are also at least two negative nuclear regulators. As mentioned previously, there is a short (~10–15 min) half-life for nuclear dephosphorylation of activated STAT proteins. Loss of this dephosphorylation would lead to prolonged STAT activation and possible untoward consequences that are mentioned below. Recently, the negative activity on STAT proteins of a group of proteins termed PIAS — proteins that inhibit activated STATs — has been discovered. PIAS1 and PIAS3 were first shown in cultured mammalian cells to interact only with tyrosine-phosphorylated STAT1 and STAT3, respectively, and to block DNA binding in vitro. Upon transfection and overexpression of the PIAS genes, transcriptional increases that are directed by demonstrably active STAT1 and STAT3 were also blocked.

Genetic interaction between the single Drosophila PIAS gene (dPIAS) and the single Drosophila STAT gene (STAT92E) indicates that PIAS might modulate STAT activity in vivo. The JAK–STAT pathway is required for development of the eye in the fly. PIAS overexpression decreases eye size and somatic-cell removal of PIAS prevents differentiation of the lens. Furthermore, a hyperactive JAK allele, tumorous lethal (tum), causes fly ‘leukaemia’, which increases in incidence with reduced PIAS levels and decreases in incidence with PIAS overexpression.

PIAS proteins have also been implicated in various processes that have no apparent connection to STAT proteins, including induction of apoptosis, modulation of ion channels, interaction with androgen receptors and interaction with RNA helicase. Recently, some PIAS proteins were shown to have E3-ligase-like activity for the small ubiquitin-related modifier SUMO. PIAS proteins mediate the conjugation of SUMO to several proteins, including p53 and c-Jun, and this represses their activities. However, the biological role of SUMO modification remains mysterious and so, the relationship between this enzymatic activity and the negative regulation of STATs by PIAS proteins remains to be elucidated.

Naturally occurring truncated STATs. In the first purification of STAT1, a carboxy-terminal truncated molecule (STAT1β) was found that could not drive IFN-γ-induced and STAT1-homodimer-dependent gene transcription. However, STAT1β could participate in IFN-γ-induced transcription as part of a three-protein interaction of STAT1β, STAT2 and IRF9. This was the first evidence of a carboxy-terminal TAD's. Subsequently, STAT3 and STAT5 were also found to have alternative carboxy-terminal truncated forms. These shortened proteins function as dominant-negative forms when overexpressed in cultured cells. In mice with a targeted knockout of Stat3β, which still have full-length Stat3α, the pattern of Stat3-induced transcription is distorted, which results in impaired recovery from endotoxic shock.

A naturally occurring truncated form of the Drosophila Stat92E has recently been uncovered. This protein lacks the first 130 amino acids and arises from an alternative transcriptional start site. The two different primary transcripts are spliced differentially to yield either a full-length 86-kDa form or a 71-kDa form. The short form acts as a negative regulator of the long form. For example, expression of full-length Stat92E is required for even-skipped (eve) expression in stripes 3 and 7 in the developing Drosophila embryo. By contrast, overexpression of the short form of Stat92E suppresses eve expression in stripes 3 and 7. So, it seems certain that balanced STAT-protein transcriptional activity is required at many points in development and in adults, and this balanced transcription requires the participation of negative regulators.

T HELPER 1/T HELPER 2 (T1/T2). Subsets of CD4+ T cells that are characterized by their cytokine-production profiles. T-1 cells primarily produce interferon-γ, and generally provide protection against intracellular pathogens, whereas T-2 cells mainly produce interleukin-4 (IL-4) and IL-5 and IL-13, and are important for immunity to helminth parasites.

CASpases. Cysteine proteases involved in apoptosis that cleave at specific aspartate residues.

HAEMATOPOIESIS. The commitment and differentiation processes that lead from a haematopoietic stem cell to the production of mature cells of all lineages — erythrocytes, myeloid cells (macrophages, mast cells, neutrophils and eosinophils), B and T cells, and natural killer cells.

RNA HELICASE. An ATP-dependent enzyme that catalyses the unwinding of RNA helices.

DOMINANT-NEGATIVE. A defective protein that retains interaction capabilities and so distorts or competes with normal proteins.

SEPTEMBER 2002 | VOLUME 3 www.nature.com/reviews/molcellbio
**ENDOTOXIC SHOCK**

Also known as septic shock. This is a serious, abnormal condition that occurs when an overwhelming infection leads to low blood pressure and low blood flow. Vital organs such as the brain, heart, kidneys and liver might not function properly or might fail.

**DNA MICROARRAY**

Array of polymerase chain reaction products (corresponding to either genomic or cDNA sequence) that is deposited onto solid glass slides.

**MACROPHAGE**

Any cell of the mononuclear phagocyte system that is characterized by its ability to phagocytose foreign particulate and colloidal material.

**STAT proteins in transcription**

The basic outlines of the biochemistry of STAT activation—inactivation, nuclear import—export and negative regulation have been referred to above. The ultimate biochemical effect of activated STATS, however, lies in their ability to increase—in a matter of minutes—the transcriptional activity of previously quiescent genes and/or to increase the transcription of less-active genes. Once STATs reach the nucleus, different STAT proteins activate different genes, owing, at least in part, to different binding affinities for natural sites and, in part, to the recruitment of distinct co-activators.

The number of genes activated by particular STAT pathways is a topic of much current research using DNA microarrays. For example, on the basis of STAT1 and STAT2 activation, IFN-α or IFN-γ increases the concentration of mRNAs from at least several dozen different genes over fourfold, and perhaps doubles the concentration of another 50–100 gene products. These include mRNAs of the IFN-α or IFN-γ genes that were shown some years ago to be transcribed much more rapidly after IFN treatment. The gene array experiments have been carried out using cells that lack STAT1 or that have been reconstituted with wild-type or mutant STAT1, which clearly shows an important role for phosphorylated STAT1, but also for unphosphorylated STAT1 in gene expression (Fig. 3).

It is known that STAT proteins bind DNA and that they associate with other transcription factors and co-activators (Fig. 3), but details of how transcriptional activation is increased are sparse.

**Post-translational modification.** Post-translational chemical modifications and tyrosine phosphorylations of STAT proteins have been described that affect STAT-driven transcription. Arginine methylation at the amino-terminal domain increased transcriptional effectiveness by blocking the association of STATS with PIAS proteins. Acetylation of STATS by associated co-activator proteins might also enhance their transcriptional activity.

**Table 2 | Role of SOCS proteins as shown by mouse genetics**

<table>
<thead>
<tr>
<th>SOCS protein</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cis</td>
<td>No phenotype of null, but enhanced T-cell signalling in mice that overexpress Cis</td>
</tr>
<tr>
<td>Socs1</td>
<td>Perinatal lethality owing to unopposed Ifn-γ-induced liver degeneration</td>
</tr>
<tr>
<td>Socs2</td>
<td>Giantism owing to unopposed signalling by growth hormone and Igf-1</td>
</tr>
<tr>
<td>Socs3</td>
<td>Embryonic lethality owing to multiple placental and haematopoietic defects</td>
</tr>
</tbody>
</table>

*See REFS 68-73. Cis, cytokine-induced SH2 protein; Ifn-γ, interferon-γ; Igf-1, insulin-like growth factor-1; Socs, suppressor of cytokine signalling.*

**Figure 4 | The negative regulators of STAT proteins.**

Phosphatases (a) and suppressors of cytokine signalling (SOCS proteins) (b) block further STAT activation in the cell cytoplasm. In the nucleus, nuclear phosphatases (c) can mediate STAT dephosphorylation, and interactions with proteins that inhibit activated STAT proteins (PIAS) (d) can also occur. In addition, naturally occurring short forms of STATS can potentially act as dominant-negative proteins by occupying DNA as non-functional protein or by binding to a wild-type STAT protein (e). AK, janus kinase; STAT, signal transducers and activators of transcription.
The mechanism underlying the enhanced transcriptional activity of serine-phosphorylated STAT proteins involves the selective recruitment of co-activators. Several proteins, including minichromosome mainte-
nance 5 (MCM5) and MCM3 (J. J. Zhang and J. E. D., unpublished observations), bind more avidly to the TAD of STAT1 when it is phosphorylated. The MCM proteins are known to function in DNA replication as helicases, but how they might participate in RNA synthesis is not known.

**STATs and HATs.** Like many eukaryotic transcription factors, the carboxy-terminal TADs of STAT1, STAT2, STAT3, STAT5 and STAT6 all interact with the co-activator HISTONE ACETYLTRANSFERASES (HATs), especially p300/CBP (CREB-binding protein)7,28. Whether STAT proteins use unusual transcriptional mechanisms remains unknown. However, at least STAT2 seems to deviate from the normal pathways of transcriptional activation. This protein is activated by IFN-α to mediate the defense against infections. Therefore, STAT2 must be capable of functioning under the stressful conditions of a viral infection. STAT2 recruits the acetyltransferase protein GCN5 (general control non-repressed; a protein originally discovered in yeast)28, which leads to acetylation of HISTONES in the promoters of IFN-α-repressed genes. Surprisingly, STAT2 can stimulate transcription through a complex that lacks some of the general proteins of transcriptional initiation, such as TATA-binding protein. Interestingly, STAT2 can stimulate transcription through a complex that lacks some of the general proteins of transcriptional initiation, such as TATA-binding protein. This is not known.

One of the most widely used co-activators in eukaryotic cells is the so-called mediator complex, a group of protein complexes that are conserved from yeast to humans. This complex has not yet been reported to interact with STAT proteins but it would be surprising if it did not.

**STATs and other DNA-binding proteins.** STAT proteins probably seldom act alone in transcriptional activation but, like most other mammalian activators, act in concert with other site-specific DNA-binding proteins (Fig. 3). Interactions necessary for maximal STAT-dependent transcription have been reported between: STAT1 and Sp1 and upstream stimulatory factor19,26; STATs and glucocorticoid receptor (GR)21; STAT6 and CCAAT/enhancer binding protein (C/EBP)56; several STAT proteins with N-Myc interactor (Nmi)97; and STAT3 with various other proteins, including c-Jun, GR, androgen receptor and similar to mothers against decapentaplegic (SMAD; ADP-ribosylation factor) (reviewed in refs 4,5,7). The first instance of this type of interaction to be noted was in the IFN-α-induced STAT1–STAT2 heterodimer with IRF9, which contacts STAT1 (REF. 4). This trimeric complex contacts a composite DNA element consisting of juxtaposed binding sites for both IRF9 and STAT1 (reviewed in refs 4,5,7).

**STAT proteins side-by-side.** STAT proteins also interact with each other on tandem DNA sites to achieve maximum transcriptional stimulation82. The amino terminus is required for these dimer–dimer interactions, which occur for all STATs except STAT2. No dimer–dimer interaction occurs between different STAT proteins, which is not surprising given that different STAT proteins are activated by different ligands and have different targets.

Although protein segments of interacting partners and even specific amino acids important for interaction have been located, we do not know precisely how such interactions increase initiation rates, this will be a fertile area for future investigation. It is not clear how STAT proteins are positioned precisely within gene-activating clusters of proteins, but it has been shown that STATs are physically present in chromatin at the time of transcriptional activation. CHROMATIN IMMUNOPRECIPITATION (ChIP) ASSAYS show the IFN-dependent presence of STAT2 on the ISG54 (interferon-stimulated-gene 54) promoter99, of STAT1 on the IRF1 (interferon-regulatory-factor-1) promoter (E. Yang and J. E. D., unpublished observations) and the class-II transactivator (CIITA) promoter102, and of STAT3 on the α2-macroglobulin promoter (L. Lerner et al., unpublished observations).

**Influence of STATs on biological functions.**

Transcription factors are, of course, crucial to biological outcomes in whole organisms. Many of the proteins discovered in the well-known Nusslein-Völlhard–Wieschaus screens for early developmental mutants in Drosophila proved to be transcription factors104. So, it is no surprise that STAT proteins are widely involved in developmental decisions.

---

**Table 3** | Role of STAT proteins as revealed by gene-targeting in mice*

<table>
<thead>
<tr>
<th>STAT protein</th>
<th>Phenotype of null mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stat1</td>
<td>Impaired responses to interferons; increased susceptibility to tumours; impaired growth control</td>
</tr>
<tr>
<td>Stat2</td>
<td>Impaired responses to interferons</td>
</tr>
<tr>
<td>Stat3</td>
<td>Embryonic lethality; multiple defects in adult tissues including impaired cell survival (both positive and negative) and impaired response to pathogens</td>
</tr>
<tr>
<td>Stat4</td>
<td>Impaired T₃,₁ differentiation owing to loss of IL-12 responsiveness</td>
</tr>
<tr>
<td>Stat5A</td>
<td>Impaired mammary gland development owing to loss of prolactin responsiveness</td>
</tr>
<tr>
<td>Stat5B</td>
<td>Impaired growth owing to loss of growth hormone responsiveness</td>
</tr>
<tr>
<td>Stat6</td>
<td>Impaired T₃,₁ differentiation owing to loss of IL-4 responsiveness</td>
</tr>
</tbody>
</table>

*Reviewed in REFS 4–7. IL, interleukin; T₃,₁, T helper 1 cell.

---

**HISTONE ACETYLTRANSFERASES (HAT).** An enzyme that adds acetyl groups to histones. Many HATs function as co-activators.

**HISTONE** A family of small, highly conserved basic proteins, found in the chromatin of all eukaryotic cells that associate with DNA to form a nucleosome.

**CHROMATIN IMMUNOPRECIPITATION (ChIP) ASSAYS** ChiP assays can be used to monitor the association of DNA-binding proteins with specific promoters in vivo. Briefly, live cells are treated with crosslinking agents to tether the protein to the DNA. The selected protein is then recovered by immunoprecipitation, the crosslinking is reversed and the co-purifying DNA is screened for the enrichment of specific promoter fragments using the polymerase chain reaction (PCR).

**References**
be assessed

individual. Multipotentiality can replace themselves (self-renewal)

haematopoietic cell

Cells that have the ability to

PROGENITORS

HAEMATOPOIETIC

against invading pathogens.

now recognized that they serve a

broader role in host defence.

were originally defined on the

(NK cells). Lymphocytes that

NATURAL KILLER CELLS

sequence.

of the Cre recombinase enzyme

(P1. Two short DNA sequences

P sites) are engineered to

bacteriophage

indicating new roles for Stat92E. First, stem cells in the

disc and eye development, as well as sex determina-

tion. In light of the fact that STAT3 and STAT5 are

Drosophila

overexpression of the ligand unpaired or the

(SPT) and one STAT (Stat92E)13. There is one

known ligand, outstretched (os) or unpaired (unp),

and, recently, receptors that have at least a distant

resemblance to cytokine receptors were reported to activate

the pathway105. The discovery of the developmental

pathways that are affected by STAT mutations began with experiments showing that a null allele was

lethal81,82,106. A role in early embryogenesis was uncov-

ered through localized gene expression — proteins or

mRNAs expressed in stripes 3.5 and 7 were decreased or

absent in hypomorphic alleles. Subsequently, the list of

tissues (structures) known to be affected extended to —

but is not limited to — wing veins, tracheal, monocytes,

erythrocyte, and eye development, as well as sex determina-

tion and germ cells106,107.

Of considerable interest are recent publications

indicating new roles for Stat92E. First, stem cells in the

Drosophila testis require JAK–STAT signalling for self-

renewal but not for differentiation107,108. In addition,

Stat92E activity was required for migration of the

border cells in the development of the ovary109 —

overexpression of the ligand unpaired or the

Drosophila JAK, hopscotch, caused increased migra-

tion110 (see below). How a single transcription factor

remains to be determined at a molecular level.

Role in infection. Several STAT proteins in mammals

have a crucial role in host defence (TABLE 4). Stat1 and

Stat2 are largely restricted to mediating the effects of

IFNs; Stat4 and Stat6 mediate the effects of IL-12 and IL-

4, respectively; and Stat3 mediates the effects of IL-6 and other gp130 ligands. Animals that lack either Stat1 or

Stat2 are exquisitely sensitive to microbial

infections120,121, and subtle mutations of STAT1 in

humans lead to decreased resistance to mycobacterial

infection122. The absence of Stat6 blocks the differentia-

tion of Th2 cells, and lack of Stat4 impairs IFN-γ produc-

tion by T cells and development of natural killer cells

during bacterial and viral infections (reviewed in REF. 6).

As mentioned above, study of the absence of Stat3

must be done by Cre-loxP-mediated removal in specific

tissue. The absence of Stat3 causes several biological

effects. For example in adult liver, Stat3 absence leads
to significantly impaired responses to acute phase activ-

ators commonly associated with bacterial infection123; in

the thymus, it impairs T lymphocyte survival124; in

macrophages, it disrupts resistance to intestinal microbes125 and, in haematopoietic progenitors, it leads to

increased accumulation of granulocytes126.

<table>
<thead>
<tr>
<th>Target tissue</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>Impaired second hair cycle, wound repair and keratinocyte migration</td>
</tr>
<tr>
<td>Thymic epithelium</td>
<td>Age-dependent thymic hypoplasia, hypersensitivity to stress</td>
</tr>
<tr>
<td>T lymphocytes</td>
<td>Impaired IL-6-dependent survival and IL-2r expression</td>
</tr>
<tr>
<td>Monocytes/neutrophils</td>
<td>Enhanced inflammatory responses and Th1 differentiation, chronic colitis</td>
</tr>
<tr>
<td>Granulocytes</td>
<td>Enhanced proliferation owing to impaired negative feedback</td>
</tr>
<tr>
<td>Mammary epithelium</td>
<td>Defective apoptosis, delayed mammary involution</td>
</tr>
<tr>
<td>Liver</td>
<td>Impaired acute phase response</td>
</tr>
<tr>
<td>Neurons</td>
<td>Impaired cell survival</td>
</tr>
</tbody>
</table>

*REFERENCES 123–126, IL-2r, interleukin-2 receptor-α, Th1, T helper 1 cell.

Stat92E in Drosophila. Drosophila has only one JAK

(hopscotch) and one STAT (Stat92E)13. There is one

known ligand, outstretched (os) or unpaired (unp),

and, recently, receptors that have at least a distant

resemblance to cytokine receptors were reported to activate

the pathway105. The discovery of the developmental

pathways that are affected by STAT mutations began with experiments showing that a null allele was

lethal81,82,106. A role in early embryogenesis was uncov-

ered through localized gene expression — proteins or

mRNAs expressed in stripes 3.5 and 7 were decreased or

absent in hypomorphic alleles. Subsequently, the list of

tissues (structures) known to be affected extended to —

but is not limited to — wing veins, tracheal, monocytes,

erythrocyte, and eye development, as well as sex determina-

tion and germ cells106,107.

Of considerable interest are recent publications

indicating new roles for Stat92E. First, stem cells in the

Drosophila testis require JAK–STAT signalling for self-

renewal but not for differentiation107,108. In addition,

Stat92E activity was required for migration of the

border cells in the development of the ovary109 —

overexpression of the ligand unpaired or the

Drosophila JAK, hopscotch, caused increased migra-

tion110 (see below). How a single transcription factor

remains to be determined at a molecular level.

Role in infection. Several STAT proteins in mammals

have a crucial role in host defence (TABLE 4). Stat1 and

Stat2 are largely restricted to mediating the effects of

IFNs; Stat4 and Stat6 mediate the effects of IL-12 and IL-

4, respectively; and Stat3 mediates the effects of IL-6 and other gp130 ligands. Animals that lack either Stat1 or

Stat2 are exquisitely sensitive to microbial

infections120,121, and subtle mutations of STAT1 in

humans lead to decreased resistance to mycobacterial

infection122. The absence of Stat6 blocks the differentia-

tion of Th2 cells, and lack of Stat4 impairs IFN-γ produc-

tion by T cells and development of natural killer cells

during bacterial and viral infections (reviewed in REF. 6).

As mentioned above, study of the absence of Stat3

must be done by Cre-loxP-mediated removal in specific

tissue. The absence of Stat3 causes several biological

effects. For example in adult liver, Stat3 absence leads
to significantly impaired responses to acute phase activ-

ators commonly associated with bacterial infection123; in

the thymus, it impairs T lymphocyte survival124; in

macrophages, it disrupts resistance to intestinal microbes125 and, in haematopoietic progenitors, it leads to

increased accumulation of granulocytes126.
Further evidence of the importance of STAT-mediated signals for resistance to infection can be seen from the multitude and variety of pathogen-encoded mechanisms that decrease STAT function and the cellular attempts to avoid these detrimental effects. Several viruses target STAT or JAK proteins for degradation or inhibit their activation in other ways. Vesicular stomatitis virus inhibits the nuclear translocation of proteins, but STAT1-mediated induction of the nuclear pore protein Nup98 overcomes this block. PICORNA VIRUSES inhibit cellular transcription by targeting transcription factors for degradation, but STAT2-dependent transcription somehow remains resistant to this effect.

**Role in growth control.** Signalling pathways that originate at the cell surface and send active transcriptional proteins to the nucleus are frequently dysfunctional in cancer cells. The STAT proteins are certainly no exception. Mice that lack Stat1 are more susceptible to chemically induced primary tumours and at the time of their activation, that can be readily transplanted and human cancer cells have often lost STAT responses to IFN, which normally imposes growth restraint. Of great current interest is persistently active STAT3, which is known to occur in a wide variety of human tumours. Furthermore, STAT3 can, by experimental mutation, be converted into an oncogene. The persistently active protein is required because introduction of a dominant-negative form of STAT3 into head and neck cancer cells or into multiple myeloma cells causes apoptosis of recipient cancer cells (reviewed in Reference 141). Persistent activation of STAT3 in head and neck cancer is associated with mutations in the epidermal growth factor (EGF) receptor or mutations that result in the production of excess ligand or normal receptor. In some multiple myelomas, excess production of IL-6 might be the underlying defect. Two recent reports of persistently active STAT3 highlight the importance of negative factors in STAT control. First, in hepatocellular carcinoma, silencing of the SOCS3 gene locus by methylation was associated with persistent STAT3 activation. Second, loss of PIA53 was found in a leukaemia in which STAT3 was persistently activated. So, mutations that cause continued signalling or ineffective negative upstream regulation of STAT3 seem to be important in promoting cancer. It is very likely that mutations in STAT3 itself are not the reason for its persistent activity in cancer.

**Conclusions and perspectives.** The enormous variety of experiments on the STAT genes and the proteins they encode in all animals reflect the widespread importance of these transcription factors and of the delicate balance normally exercised on the extent and time of their activation. Biochemistry of the mechanism of role of STAT proteins in transcription initiation has lagged behind the exploration of the biological decisions in which the STAT proteins participate. However, studies of the association of STAT proteins with various nuclear proteins have begun. With the reagents developed in the many studies on transcription both in vivo and in vitro, passage of another ten years should see answers to presently unanswered biochemical questions and show how integration into biological decisions is achieved by this important group of transcription factors.

Increasing our knowledge of how STAT proteins affect transcriptional regulation is important as basic information for understanding coordinate control in mammalian cells. Furthermore, as a practical consideration — if pharmacological intervention of specific transcriptional activity is ever to be achieved — it is necessary to know the most common interacting partners of specific activators. Therefore, many experiments have been and are being — done using mutated STAT proteins that fail to function at various steps in transcriptional activation, and experiments on STAT — STAT and STAT — protein interactions will very probably provide targets for drug development.

References 44–48 outline distinct mechanisms for constitutive and induced STAT nucleo-nuclear translocation.


REVIEWS


Acknowledgements

Work in the authors’ laboratories was supported by grants from the NIH.

Online links

DATABASES

The following terms in this article are linked online to:

Cell

Head and neck cancer | leukaemia

Nature

Cancer | cancer control

Oncogene

Gastric cancer

Access to this interactive links box is free online.