The EMBO Journal vol.15 no.20 pp.5616–5626, 1996

DNA binding of *in vitro* activated Stat1α, Stat1β and truncated Stat1: interaction between NH_2-terminal domains stabilizes binding of two dimers to tandem DNA sites

Uwe Vinkemeier, Steven L.Cohen[1], Ismail Moarefi[2], Brian T.Chait[1], John Kuriyan[2] and James E.Darnell,Jr

Laboratory of Molecular Cell Biology, [1]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry and [2]Howard Hughes Medical Insitute and Laboratory of Molecular Biophysics, The Rockefeller University, New York, NY 10021, USA

Stat1α, Stat1β and a proteolytically defined truncated Stat1 (132–713, Stat1tc) have been prepared from recombinant sources. All three proteins were specifically phosphorylated on Tyr701 *in vitro* and the phosphoprotein purified to homogeneity. This was achieved by employing a new isolation scheme that does not include DNA affinity steps and readily allows for the isolation of tens of milligrams of activated Stat protein. The purified phosphoprotein was free of traces of unphosphorylated polypeptide as detected by mass spectrometry. The phosphorylated Stat1 preparations bound to various DNA recognition sites with the same K_{eq} of ~1×10^{-9} M; distinction between 'weak' and 'strong' binding sites is determined by the very rapid dissociation (<30 s, $t_{1/2}$) from 'weak' sites compared with 'strong' sites (~3 min, $t_{1/2}$). Reports of 'weak' tandem binding sites in a natural gene caused us to examine binding to tandem sites leading to the finding that the Stat1α or β (38 amino acids shorter on the C terminus) bound to two tandem sites (but not two head-to-head sites) with a higher stability than to a single recognition site. The N-terminally truncated protein Stat1tc did not show this cooperative binding, thus implicating the N-terminal domain in promoting Stat1–Stat1 dimer interaction.

Keywords: cooperative DNA binding/N terminus/phosphoprotein/proteolysis/Stat1

Introduction

The Stat proteins are latent transcription factors that are activated in the cell cytoplasm by phosphorylation on a single tyrosine (Darnell *et al.*, 1994; Schindler and Darnell, 1995). The kinases responsible are either ligand-activated transmembrane receptors with intrinsic tyrosine kinase activity, such as EGF or PDGF receptors, or cytokine receptors that lack intrinsic kinase activity but have associated JAK kinases, such as those for interferons and interleukins (Ihle, 1995). After tyrosine phosphorylation, the Stat molecule either homo- or heterodimerizes, then translocates to the nucleus and participates in transcriptional activation (Shuai *et al.*, 1994; Qureshi *et al.*, 1995).

To conveniently study the biochemistry of activated Stat molecules, it is necessary not only to use recombinant DNA techniques to produce large amounts of protein, but also to phosphorylate the correct tyrosine residue and to separate the phosphorylated and non-phosphorylated proteins. In this paper we describe the production of human Stat1α and Stat1β, a shorter protein translated from a differently spliced mRNA, in insect cells infected with recombinant baculovirus, purification of large amounts of the proteins, and the protease sensitivity of Stat1α. A stable truncated form of Stat1 (Stat1tc) was then characterized and produced in bacteria. Both Stat1α and β and the Stat1tc were quantitatively phosphorylated *in vitro* with immunoprecipitated, activated EGF receptor kinase. The phosphoprotein was isolated in milligram quantities by a new chromatographic protocol, and the phosphorylation was shown by mass spectroscopy of Stat1 fragments to be on the correct tyrosine residue. Both the full-length and the truncated phosphorylated protein dimerize and bind to DNA.

With the purified activated DNA binding form of Stat1 available, we studied its DNA binding characteristics and found a K_{eq} of ~1×10^{-9} M for a variety of recognition sequences. By examining the stability of labelled preformed protein–DNA complexes when challenged with unlabelled DNA, we found a very short half-life of the protein–DNA complexes. For sites that showed the maximum binding stability, we determined a $t_{1/2}$ of 3 min. An extremely rapid exchange (half-life of <30 s) was observed for both Stat1α and Stat1tc bound to the sites that are natural 'weak' binding sites in genomic DNA. It was recently reported that Stat1 dimers (Guyer *et al.*, 1995) or dimers of *Drosophila* Stat protein (D-Stat) (Yan,R. *et al.*, 1996) may interact when two nearby Stat binding sites are both occupied. We therefore tested such a possibility with purified activated protein, and found evidence of interaction between bound dimeric molecules such that binding to adjacent sites was stabilized when both were occupied. Furthermore we found that this proposed Stat dimer interaction was dependent on the presence of the N-terminal 131 amino acids of Stat1.

Results

Production by recombinant techniques and purification of Stat1

The cDNA encoding Stat1α or β was inserted in baculovirus transfer vector (pAcSG2) and co-transfected with modified linearized AcPNV baculovirus DNA to produce virus particles. Insect cells (Sf9 cells) infected with the respective recombinant baculovirus produced the 91 or 84 kDa proteins, identified by Western blot as Stat1, which were purified (Figure 1A) through the steps indicated in Table I.

Stat1α is 750 amino acids long; Stat1β is a product of a differentially spliced mRNA which encodes a protein 712 amino acids long (Schindler *et al.*, 1992; Yan *et al.*,

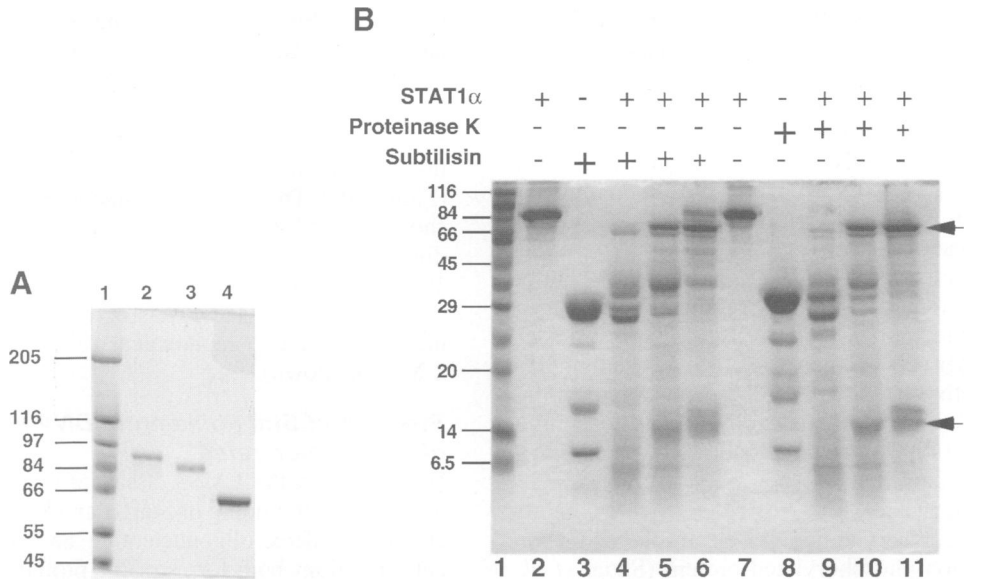

Fig. 1. (**A**) Polyacrylamide gel electrophoretic analysis of the purified non-phosphorylated proteins used in this study. Aliquots of Stat1α (lane 2, 2 μg), Stat1β (lane 3, 2 μg) and Stat1tc (lane 4, 4 μg) were run on a 7% SDS–PAGE gel and stained with Coomassie blue. Molecular weight standards were run in lane 1. M_r is given in kDa on the left. (**B**) Proteolysis of human Stat1α. 40 μg of purified Stat1α were digested with various amounts of subtilisin (lanes 4–6) or proteinase K (lanes 9–11) for 30 min on ice (see Materials and methods). The ratios (w/w) of protease to protein were 1:8 (lanes 4 and 9), 1:80 (lanes 5 and 10) and 1:800 (lanes 6 and 11). Aliquots of the reactions were resolved on a 16.5% SDS–polyacrylamide gel followed by Coomassie staining. Lane 1, standards (kDa); lanes 2 and 7, untreated Stat1α; lane 3, subtilisin (15 μg); lane 8, proteinase K (15 μg). Stable fragments of 65 kDa and 16 kDa (see text) are marked with arrows.

1995). It is known that both Stat1α and β are phosphorylated on a single tyrosine, residue 701; both dimerize, translocate to the nucleus and bind DNA (Shuai *et al.*, 1992, 1993a). We digested the purified Stat1α with several proteolytic enzymes to determine whether the protein could be divided into functional domains. Both subtilisin and proteinase K produced two major digestion products (Figure 1B), the largest of which migrated on SDS–PAGE with an estimated size of 65 kDa as compared with the full-length protein of 91 kDa. (Cleavage products of ~40 and 30 kDa were also seen.) The 65 kDa product had an N-terminal sequence of XTVMLDKQEKE indicating that it resulted from cleavage between residues 131 and 132 of the full-length protein. A single prominent smaller fragment of ~16 kDa was also observed. This fragment was the only one generated that retained reactivity with an antibody raised against the N terminus of Stat1 (not shown). The shorter 16 kDa fragment was therefore an N-terminal fragment of the molecule.

The longer major fragment beginning at residue 132 was poorly recognized by an antibody to the C-terminal 38 amino acids of Stat1α (not shown) suggesting additional cleavage near the C terminus. Since the fragment beginning at residue 132 was resistant to proteolysis and we knew that Stat1β terminating at residue 712 was phosphorylated on Tyr701 *in vivo*, and was able to dimerize and bind to DNA, we prepared a bacterial expression clone encoding residues 132–713. The inclusion of residue 713 was based on the accidental inclusion of the codon for amino acid 713 in the primers used for PCR amplification of the Stat1 cDNA. This product, Stat1(132–713) or Stat1tc, was expressed well in *Escherichia coli* with a major fraction of soluble protein; this was purified to homogeneity (Figure 1A and Materials and methods).

Table I. Purification of Stat1α/β

		Volume (ml)	Protein (mg)[a]
I	crude extract[b]	80	550
II	S-Sepharose	120	30
III	Q-Sepharose	30	12
IV	ammonium sulfate	1	8
V	alkylation	10	8
VI	phenyl-Sepharose	25	6[c]
VII	Superdex 200	3	5[c]

[a]Protein concentrations were determined by the method of Bradford (1976) using BSA as a standard (Bio-Rad).
[b]After pH 6.2 precipitation from 5×10^8 cells.
[c]Determined by UV spectroscopy (see Materials and methods).

Aggregation of native proteins

Because the purified Stat1α/β and Stat1tc both eluted during gel filtration as peaks with a broad leading shoulder, it appeared that aggregation of the protein had occurred. Thiol crosslinking was identified as the cause, since the preparation showed aggregates that migrated with an apparent molecular mass corresponding to dimers and higher order assemblies when run under oxidizing conditions on a denaturing polyacrylamide gel (not shown). Accordingly, to permanently block reactive thiols, we treated the cell extracts (from baculovirus-infected Sf9 cells for Stat1α and transformed *E.coli* for Stat1tc) with *N*-ethyl maleimide (NEM) during purification (Table I). The purified protein was cleaved with cyanogen bromide and endoprotease Asp-N. Mass spectrometric analysis of the resulting peptides showed that cysteines 155, 440 and 492 were alkylated by the NEM treatment, whereas two others were not (Cys552 and Cys577). The NEM treatment did not affect any of the subsequent experiments (e.g.

DNA binding, see Figure 3B) and was adopted as the standard preparation of a homogeneous protein.

Phosphorylation in vitro of Stat1α/β and truncated Stat1 by the EGF receptor

The *in vivo* activated DNA binding form of Stat1 is phosphorylated on Tyr701 when isolated from mammalian cells treated with ligands that activate either JAK kinases or transmembrane receptor kinases (Shuai *et al.*, 1992, 1993b). Good activity of EGF receptor kinase can be achieved with immunoprecipitates of membrane preparations from cultured human A431 cells that express 5×10^6 EGF receptors per cell (Yarden *et al.*, 1985; Quelle *et al.*, 1995). We therefore used such EGF receptor preparations as the source of enzyme to catalyze tyrosine phosphorylation of Stat1 and to test for the possibility of phosphorylation of the truncated Stat1.

Stat1α phosphorylated on tyrosine *in vivo* is slightly retarded during polyacrylamide gel electrophoresis compared with the non-phosphorylated protein (Shuai *et al.*, 1992). This same change of mobility was observed in a portion of the *in vitro* EGF receptor kinase-treated preparations of Stat1α; moreover, the retarded protein contained ^{32}P when the enzymatic reaction was carried out in the presence of [γ-^{32}P]ATP (Figure 2A). Similar results were obtained employing EGF receptor to phosphorylate Stat1tc. From Figure 2A it is clear that not all of the Stat1 protein was phosphorylated. Further experiments yielded somewhat higher amounts of phosphorylation, sometimes as high as ~75%. However, never was all of the substrate phosphorylated, and isolation of phosphorylated from non-phosphorylated material was required in many experiments. After several attempts with different chromatography substrates to separate the large quantities of phospho- and non-phosphoprotein, we chose heparin–agarose chromatography (Figure 2B). This chromatography step resulted in a separation into two Stat protein-containing peaks (eluted in steps of 150 mM and 400 mM KCl). The tyrosine phosphorylated protein (Figure 2B) and protein with DNA binding activity (not shown) were present in the second of these peaks.

To determine the purity of the isolated material and to analyse whether the correct tyrosine residue was phosphorylated, both purified, unphosphorylated (before reaction with EGF receptor) and phosphorylated protein (the phosphotyrosine-containing peak from the heparin–agarose column) were subjected to endoprotease Asp-N digestion and the resulting peptide fragments were analysed by mass spectrometry (Figure 2C). Phosphorylation increases the molecular mass of an unphosphorylated fragment by 80 Da. Comparison of the Asp-N fragments of phosphorylated versus unphosphorylated Stats showed an 80 Da shift on fragment 694–720 (Figure 2C), demonstrating that *in vitro* phosphorylation by EGF receptor kinase occurred exclusively on the single tyrosine residue that is phosphorylated in the cell. In addition, the bottom panel of Figure 2C demonstrates the absence of unphosphorylated Tyr701 in the purified EGF receptor kinase-treated protein.

Both in vitro phosphorylated Stat1α and Stat1tc bind specific DNA fragments

We used electrophoretic mobility shift assays (EMSA) (Fried and Crothers, 1981; Garner and Revzin, 1981) to test DNA binding of tyrosine phosphorylated Stat1α and Stat1tc. Both proteins bound to all tested labelled deoxyoligonucleotides known from earlier studies to bind Stat1 (the oligo *cfos*WT is illustrated in Figure 3A). The bound complexes were not affected by NEM indicating that alkylation of cysteine does not affect DNA binding (Figure 3B). This result is in line with earlier experiments showing that ISGF3α, now known to be a Stat1–2 heterodimer, is not affected by NEM treatment (Levy *et al.*, 1989). In addition, the DNA binding ability of homodimeric phosphorylated Stat1α or its truncated form was highly resistant to monovalent salt concentrations up to 2 M (not shown).

Strength of Stat1 binding to DNA and estimation of dissociation rates

We next used the EMSA assay to obtain an estimate of the strength of binding of Stat1α and Stat1tc. Using a fixed amount of deoxyoligonucleotide and increasing protein concentrations both forms of the protein behaved identically (Figure 4). From this data, comparing the bound and unbound fraction of DNA as a function of protein concentration, a K_{eq} of $\sim1\times10^{-9}$ M was estimated for both proteins. This is in the affinity range for transcription factors which have been reported to have K_{eq} between 10^{-9} and 10^{-12} M for the proteins with the highest affinity for their cognate DNA sites (Riggs *et al.*, 1970; Affolter *et al.*, 1990). The same results were obtained with several different oligonucleotides, the Ly6E and *cfos*WT Stat binding sites, which are 'weak' binding sites, and 'strong' sites, such as the selected optimum site, S1 (Horvath *et al.*, 1995) and a mutated *cfos* sequence (M67 site; Wagner *et al.*, 1990). ['Strong' and 'weak' in this context refer to experiments with cell extracts containing activated Stat1 which binds more of some oligonucleotides (strong) than others (weak).]

We then examined the stability of preformed DNA–protein complexes by the following widely used method: formation of a complex between protein and labelled oligonucleotides is allowed to occur and unlabelled oligonucleotides are added in vast molar excess after the reaction reaches equilibrium. At various times after the addition of unlabelled competitor DNA, aliquots are layered on a running native polyacrylamide gel to determine free and bound oligonucleotides. This type of experiment was carried out with both Stat1α and Stat1tc, and with two different labelled DNAs, the natural *cfos* site and the mutated *cfos*-promotor element (M67) as examples of 'weak' and 'strong' sites, respectively.

With the 'weak' site, the 'off' time was so short that the addition of unlabelled nucleotides for as little as 30 s removed all preformed protein–DNA complexes (Figure 5B; Stat1α shown in the left panel). With the 'strong' site, the preformed labelled complexes were displaced more slowly, the $t_{1/2}$ estimated at 3 min (Figure 5B, right panel employed Stat1tc). In these experiments there was no difference between Stat1α and Stat1tc.

Stat binding to tandem DNA sites: evidence for stabilized promotor occupancy through protein–protein interactions of Stat1α/β versus Stat1tc

Two recent reports on promoters of genes dependent for transcription on Stat proteins indicated that two neighbor-

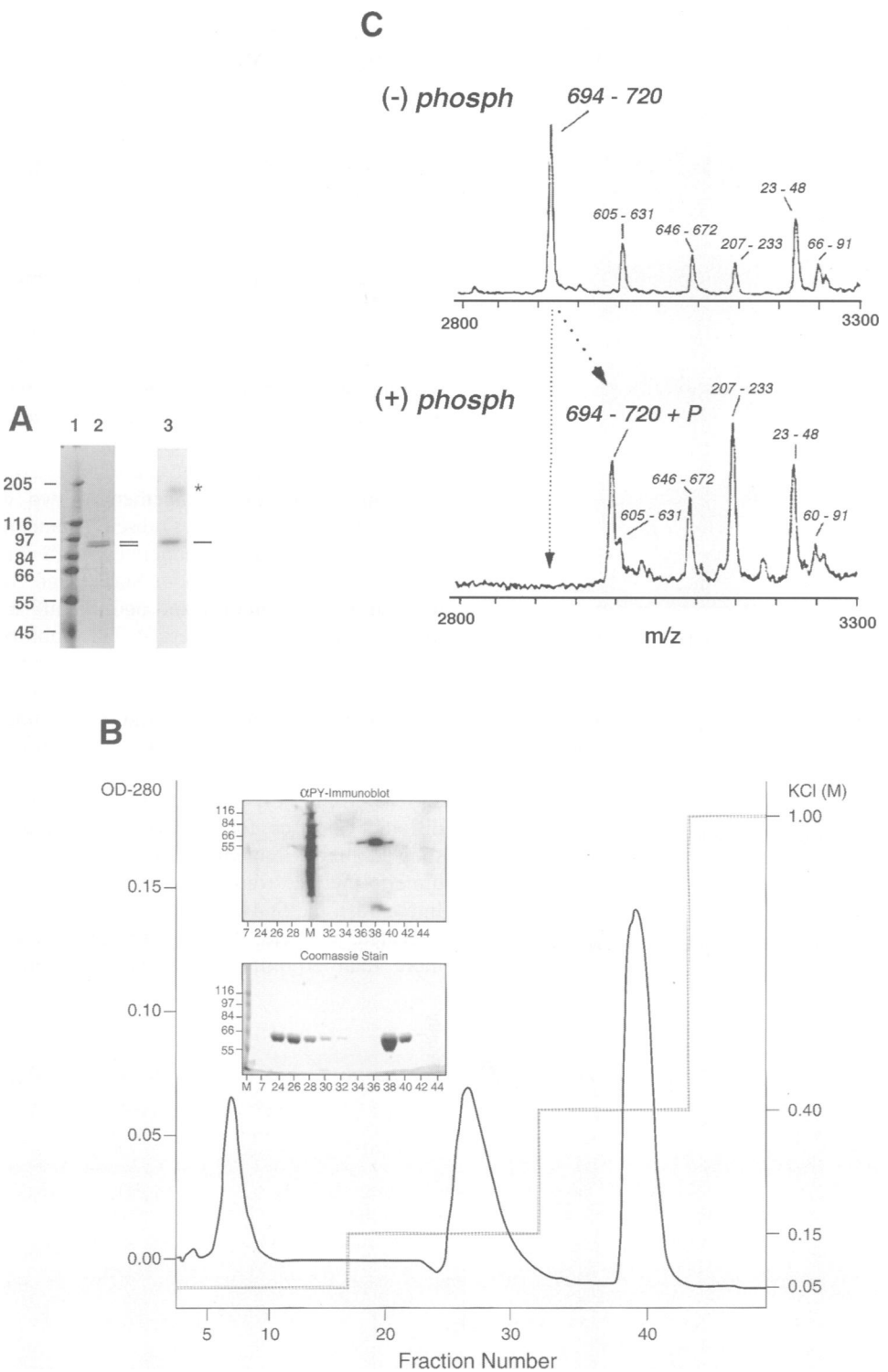

Fig. 2. (**A**) Phosphorylation of Stat1α with EGF receptor kinase *in vitro*. 2 μg of Stat protein was incubated with EGF receptor and 1 μCi of [γ-^{32}P]ATP for 6 h at 4°C. The reaction (20 μl volume) was stopped by the addition of SDS-sample buffer, resolved on a 7% SDS–PAGE, which was subsequently dried and exposed to an X-ray film. The typical 2 band pattern for phosphorylated Stat1 (Shuai *et al.*, 1992) is seen in the Coomassie stained gel in lane 2. Lane 3 shows the autoradiogram. Only the more slowly migrating band contains ^{32}P. (*) denotes the position of the phosphorylated EGF receptor. Fast- and slow-migrating Stat proteins are pointed out with lines. Lane 1, molecular weight markers (kDa). (**B**) Isolation of *in vitro* phosphorylated Stat1tc. A total of 25 mg of protein was loaded on a heparin–agarose column after an *in vitro* phoshorylation reaction and removal of EGF receptor (see Materials and methods). Shown is the profile of UV-absorptive material eluted with steps of 50 mM KCl, 150 mM KCl and 400 mM KCl. Five microliters out of the indicated fractions (2.5 ml) were run on a 7% SDS–PAGE and stained with Coomassie blue (lower insert) or blotted on a nitrocellulose membrane and probed with an anti-phosphotyrosine antibody [1:1500 diluted PY 20 (UBI); upper insert]. Molecular weights are in kDa. (**C**) Tyr701 is phosphorylated by EGF receptor. Endoproteinase AspN digests (15 min) were carried out on unphosphorylated (top spectrum) and chromatographically purified phosphorylated (bottom) alkylated Stat1α. The relevant portions of the matrix-assisted laser desorption/ionization mass spectrum are shown. Accurate molecular mass determinations allowed for unequivocal identification of the peptide fragments. Peaks are labelled according to the amino acid sequence of the corresponding peptides. P indicates a phosphate group.

ing Stat binding sites are both required for maximal transcriptional stimulation. In one of these reports the human *mig* gene promoter was found to have two weak Stat1 binding sites within 25 bp, neither of which alone conferred IFN-γ transcriptional activation while both sites together did so. Moreover the active element formed complexes with Stat1 protein that migrated more slowly than Stat1 dimers bound to DNA. The authors suggested that interaction between Stat homodimers might occur in the complexes (Guyer *et al.*, 1995). In addition, we recently reported that two D-Stat binding sites were found in the segment of the even-skipped promoter that directs stripe 3 formation in *Drosophila* embryos; both sites were required for maximum stripe 3 expression (Yan,R. *et al.*, 1996).

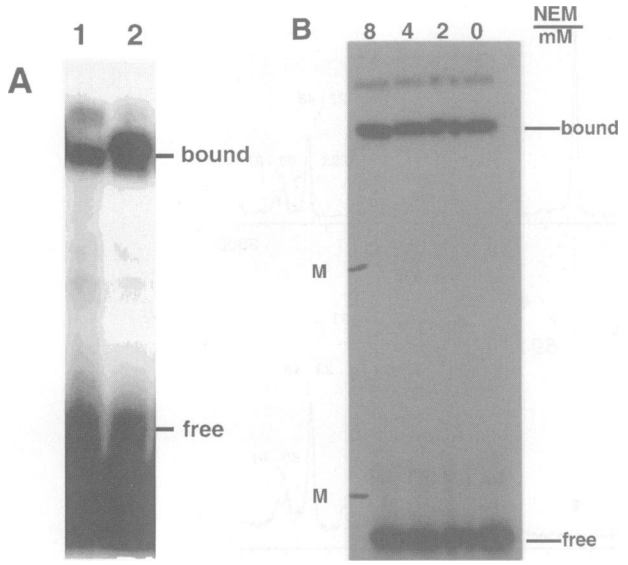

Fig. 3. (**A**) DNA binding of purified phosphorylated Stat1α (lane 1) and Stat1tc (lane 2) using as a probe the radioactively labelled c*fos*WT sequence. Binding reactions contained equimolar amounts of the respective protein. The position of migration of the free DNA probe (free) and the protein–DNA complex (bound) is indicated. Note the presence of a more slowly migrating band only with the full-length Stat1α (see also B). (**B**) Influence of cysteine alkylation on the DNA binding activity of Stat1α. A mixture of phosphorylated and unphosphorylated protein (0.23 μM final; ~15% phosphoprotein) was reacted in the presence of 0.8 mM DTT and the indicated concentrations of *N*-ethyl-maleimide (NEM) for 20 min at room temperature in a volume of 12.5 μl. The reaction was stopped with DTT to 10 mM followed by the addition of 1.5 pmol of labelled probe (c*fos*M67). Samples were resolved on a 4.5% native polyacrylamide gel. M denotes the position of bromophenol blue (lower) and xylene cyanol (upper) markers.

With the demonstration that Stat1α protein indeed does have such a rapid off-time, especially on natural 'weak' binding sites, we decided to explore the binding of activated protein to oligonucleotides containing two weak DNA binding sites. The experiments were carried out with both Stat1α and Stat1tc and a labelled oligonucleotide containing a variety of arrangements of two 'weak' binding sites. With two binding sites present in tandem on the same DNA fragment and at a moderately high concentration of protein (0.55×10^{-9} M), both Stat1α and Stat1tc formed the homodimer complex and additionally a more slowly moving complex [2×(dimeric)]. The mobility of this more slowly moving complex suggested occupation of both DNA binding sites, indicating one DNA molecule with two Stat dimers bound to it (Figure 6A, time zero). When such complexes were challenged for various times with an excess of unlabelled oligonucleotide, both the dimeric and [2×(dimeric)] complexes were dispelled but with different kinetics for Stat1α and Stat1tc. The Stat1tc showed almost immediate displacement (<1 min) of both dimeric and [2×(dimeric)] complexes (Figure 6A, left). In contrast, the Stat1α homodimer disappeared quickly as expected, whereas the [2×(dimeric)] complex required more than 30 min for partial displacement, clearly a

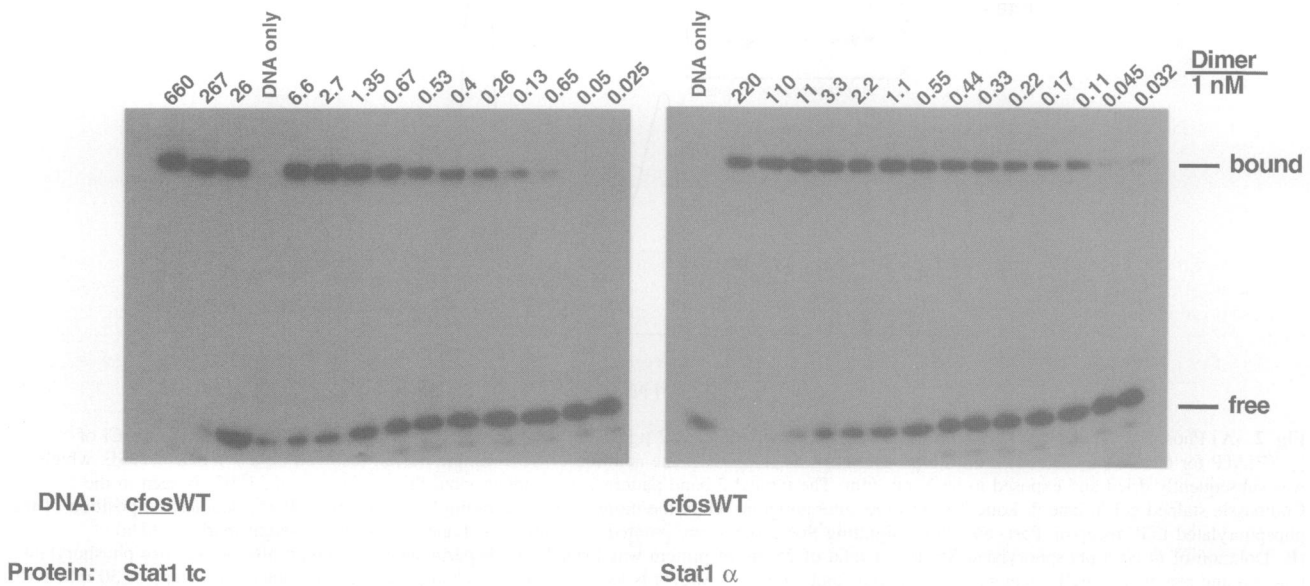

Fig. 4. Titration ^{32}P-labelled c*fos*WT oligonucleotide with phosphorylated Stat1tc (left panel) and full-length Stat1α (right panel). A fixed amount of ^{32}P-labelled c*fos*WT oligonucleotide (5.6×10^{-10} M) was incubated with Stat1 proteins in a 12.5 μl volume as described in Materials and methods. Numbers above the lanes indicate the concentrations of dimeric Stat1α and Stat1tc in each reaction. Protein-bound (bound) and free (free) DNA is identified. The concentration of free protein dimers at half saturation was determined to be ~1 nM in both cases and corresponds to the apparent equilibrium constant K_{eq}. In those lanes marked above 'DNA only' no Stat protein was included in the reaction.

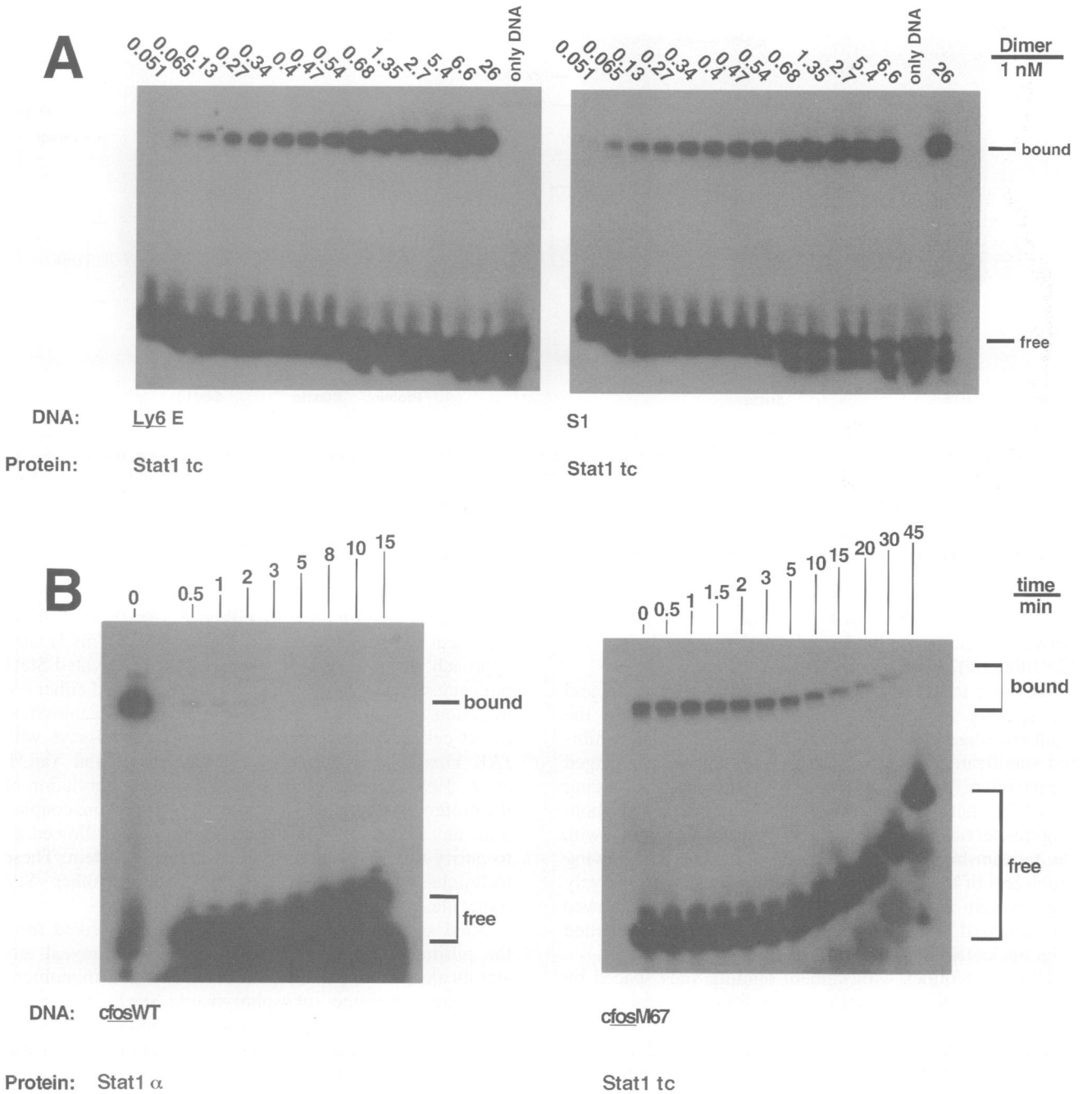

Fig. 5. (**A**) Titration of phosphorylated truncated Stat1 protein with ^{32}P-labelled oligonucleotides containing a 'low' (Ly6 E, left panel) or 'high' (S1, right panel) affinity binding site. The DNA concentration was fixed at 2.6×10^{-10} M and titrated in a 12.5 µl volume against a standard protein dilution series ranging from 5×10^{-11} to 2.6×10^{-8} M dimer final. Protein concentrations for the dimeric protein are given above each lane. The products were resolved on a native 4.5% polyacrylamide gel and quantified as described in Materials and methods. Bound, protein–DNA complex; free, free DNA. There was no Stat1tc included in reactions run on lanes denoted 'only DNA'. The dimer concentration at half saturation was determined from this autoradiograph to be ~1×10^{-9} M for both DNA sequences. (**B**) The complex of Stat1α with *cfos*WT DNA is less stable than the complex with *cfos*M67 DNA. Results are shown of off-rate experiments in which 0.55×10^{-9} M dimer was prebound with the radiolabelled DNA fragments (at 2×10^{-9} M) containing the *cfos*WT (0 min; left panel) or *cfos*M67 (0 min; right panel) sequences. Excess unlabelled DNA (100× molar excess) was added to the reaction at time zero, and aliquots were taken at the indicated intervals and loaded onto a running gel to visualize the amount of complex remaining. The half-life of the Stat1α–*cfos*WT complex is <0.5 min and that for the Stat1tc–*cfos*M67 complex in this titration is ~3 min. Because the electrophoresis was continous during the experiment the DNA fragments (free) and the complexes (bound) are located progressively higher on the gel with increasing time, because the later samples were electrophoresed for shorter periods of time than were the earlier ones.

significant increase in stability of this larger complex with the full-length proteins.

These results suggested that when Stat1α is bound at tandem binding sites, protein–protein interactions occur that require the presence of the N- and/or C-terminal domain of Stat1α to form the more stable DNA–protein complexes. To examine this question we compared Stat1tc in the chase assay with the Stat1β protein, which only lacks the C-terminal domain. As shown in Figure 6B, Stat1β exhibits the same behavior as the full-length protein,

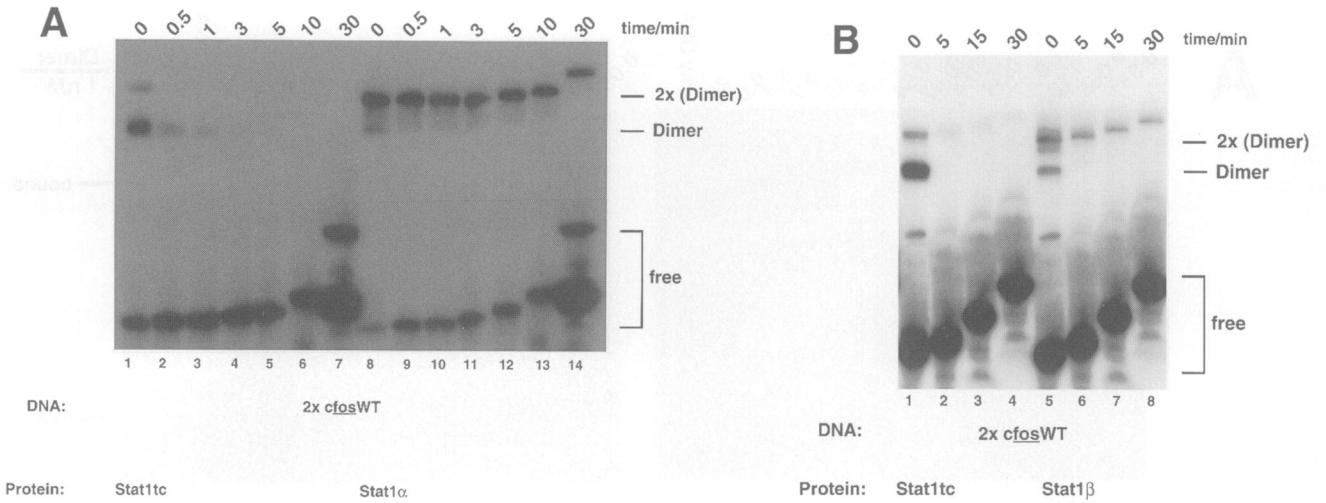

Fig. 6. (**A**) Comparison of the dissociation rates of complexes containing DNA fragments with two consecutive binding sites (2× c*fos*WT, 10 bp apart) and Stat1α (right) or Stat1tc (left). 0.5×10^{-9} M dimer was prebound with 0.7×10^{-9} M radiolabelled DNA for 5 min at room temperature (lanes 1 and 8). After the addition of a 100-fold molar excess of unlabelled DNA at time point zero the reaction was incubated further for the times indicated before aliquots were loaded on a running polyacrylamide gel. At time zero two differently migrating complexes are visible, denoted '[2×(Dimer)]' and 'Dimer'. Unbound (free) DNA runs at the bottom of the gel. (**B**) Identification of the N-terminal 131 amino acids as functional in [2×(Dimer)] stabilization on DNA. Comparison of stability of Stat1β (lanes 5–8) and Stat1tc (lanes 1–4) on DNA fragments contaning two consecutive binding sites (2× c*fos*WT, 10 bp apart). The experimental setup was the same as in (A).

indicating involvement solely of the N-terminal region between amino acids 1 and 131 in stabilizing the [2×(dimeric)] complexes.

We then tested the importance of the orientation and the spacing of the two Stat binding sites within the synthetic oligonucleotides. First the DNA sites that exhibited stabilization in [2×(dimeric)] binding were changed from tandem (→→) to inverted (→←), keeping the spacing at 10 bp between the two binding sites. While both oligonucleotides were capable of binding two dimers {with the tandem binding sites in inverted orientation showing much less of the [2×(dimeric)] complex even at relatively high protein to DNA ratio}, the inverted sites showed no increased stability when challenged with unlabelled oligonucleotide (Figure 7A).

Oligonucleotides with tandem binding sites spaced by 5 or 15 bp were prepared to test for comparison with the original oligonucleotide with 10 bp spacing. The oligonucleotide with a 15 bp spacing behaved indistinguishably from the one with 10 bp spacing, while the oligonucleotide with 5 bp spacing showed much less evidence of enhanced stability of the [2×(dimeric)] complex, suggesting that protein–protein interaction was less possible when the DNA spacer was of inadequate length (Figure 7B).

Discussion

The production of three purified Stat1 protein preparations from recombinant DNA constructs was achieved: Stat1α and Stat1β from baculovirus-infected insect cells, and a Stat1tc from *E.coli*. Digestion of purified Stat1α protein suggested a compact domain in the N terminus of 131 amino acids and a relatively protease-resistant large C-terminal fragment (132–712). Activated EGF receptor partially purified from membranes by immunoprecipitation was capable of catalyzing Stat1α, Stat1β and Stat1tc *in vitro* phosphorylation of Tyr701, the same tyrosine that is phosphorylated *in vivo* by either IFN-α, IFN-γ or EGF treatment of cells (Shuai *et al.*, 1992, 1993a). This *in vitro* approach was more efficient in generating activated Stat1 molecules than previous attempts that employed either co-infection of Stat1 and a JAK kinase in the baculovirus/insect cell system *in vivo*, or *in vitro* kinase assays with JAK kinases (our unpublished observations and Yan,H. *et al.*, 1996). These results on *in vitro* phosphorylation of the protein plus alkylation to prevent aggregation, coupled with an adequate chromatographic protocol, allowed us to purify milligram quantities of activated protein. These techniques should also be applicable to other Stat molecules.

Mass spectrometric analysis of peptides derived from the purified phosphorylated protein did not reveal any significant contamination with unreacted Stat monomers. All three tyrosine phosphorylated Stat1 derivatives dimerized and, as tested by EMSA, bound to the same DNA oligonucleotides previously shown to bind activated Stat1 in cell extracts.

The structure of the Stat molecule is expected to be complex considering the number of interactions these proteins must undergo. The region from residues 400–500 specifies DNA contacts (Horvath *et al.*, 1995), while the C-terminal half of the molecule contains the recognizable SH2 and putative SH3 domains (Fu, 1992; Schindler *et al.*, 1992), and the C terminus comprises the transactivation domain (Muller *et al.*, 1993; Wen *et al.*, 1995). From the digestion by proteases, which released an N-terminal and a C-terminal fragment, we suggest a compact structure for the N-terminal ~131 amino acids, a function for which we suggest later in the paper. In addition there is a large stable fragment beginning at 132 that can be phosphorylated on tyrosine and which dimerizes. Given the availability of activated Stat1α/β and Stat1tc, we showed that the Stat1 protein bound to various DNA fragments with a K_{eq} of 1×10^{-9} M. Compared with other regulatory proteins this is a relatively modest affinity. Despite having

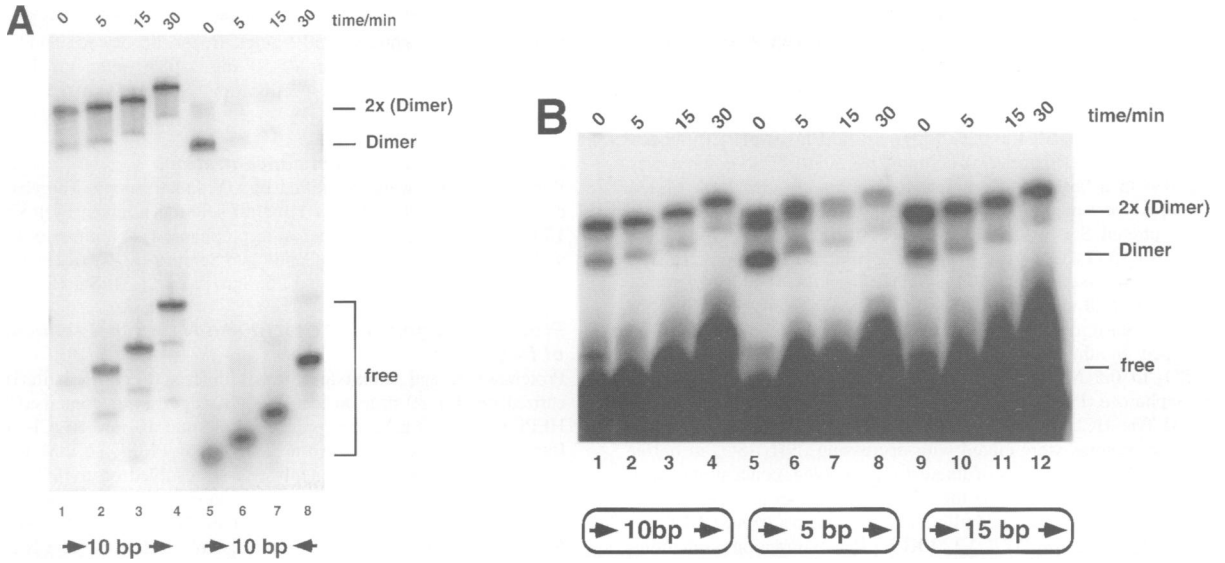

Fig. 7. (**A**) Influence of promotor orientation on protein–DNA complex formation and stability. 1.65×10^{-9} M Stat1α dimer were equilibrated with labelled DNA (0.7×10^{-9} M) with two consecutive binding sites (2× c*fos*WT) 10 bp apart in parallel (lanes 1–4) or antiparallel (lanes 5–8) orientation. The preformed complexes were chased with unlabelled competitor DNA as described for Figure 6A. (**B**) Stat1α binding to DNA fragments with two parallel binding sites (2× c*fos*WT) spaced 10 bp (lanes 1–4), 5 bp (lanes 5–8) or 15 bp (lanes 9–12) apart. The chase experiment was performed as described in the legend to Figure 6A.

similar apparent K_{eq} values, the binding with DNA may differ significantly in rates of association with and dissociation from the Stat protein. The Stat1 protein achieves equilibrium in DNA binding very rapidly, far quicker (<30 s) than the EMSA technique can determine. When the stability of Stat1 protein preparations to the various Stat1 binding sites was examined, measurable differences became apparent. Although the protein–DNA complex had a half-life of no more than 3 min for any of the sites tested, the 'off' times for different oligonucleotides varied by at least 6-fold. The difference between 'strong' and 'weak' oligonucleotide binding as detected in gel shift assays was found to be due to the exceedingly rapid 'off' time in competition assays with the displacement from 'weak' sites being essentially instantaneous. When we compared the DNA binding activities of the Stat dimer to a single recognition sequence, we could see no differences between the full-length Stat1α and the C- and N-terminally truncated Stat1tc.

The new finding of greatest potential biological relevance in these studies concerns the cooperative stabilization of Stat homodimers on neighboring binding sites This was observed when two tandem sites (separated by 10 or 15 bp) were both occupied by homodimers. A large complex was formed consisting presumably of two homodimers which was more stable to competition with unlabelled oligonucleotides than one dimer binding to a single site. This interaction required a minimum spacing (>5 bp) between adjacent sites and was strongly orientation-dependent, i.e. it occurred only if both recognition sequences were in tandem.

Additionally a domain in the Stat1 molecule required for this dimer–dimer interaction was determined. The Stat1β lacking the C-terminal 38 amino acids showed the same stabilization of the [2×(dimeric)] Stat complex on the DNA as the full-length protein. However, the truncated protein Stat1tc that lacks the N-terminal 131 amino acids (as well as the C-terminal sequence) formed the higher order complex less well, and this complex was not stabilized during oligonucleotide competition. Thus the N-terminal 131 amino acids of Stat1 defined by proteolysis as a stable domain, and which is dispensable for dimer formation and binding to single DNA sites, participates in Stat dimer–dimer interaction on tandem DNA sites. Interestingly, the isolated N-terminal domain dimerizes in solution (U.Vinkemeier and J.E.Darnell, unpublished observations). Because the N terminus of the Stats shows rather high sequence homology (Schindler and Darnell, 1995), it is conceivable that protein–protein interaction in this domain is of general importance in Stat function. Since there is evidence from the *mig* gene (Guyer *et al.*, 1995) that neighboring 'weak' Stat binding sites are required for an IFN-γ response, we think it likely that the interaction we describe has a biological role.

Materials and methods

Expression and purification of Stat1α and Stat1β

The sequences coding for human Stat1α and 1β were amplified by PCR (primers containing respective restriction sites in addition to homologous sequence; Vent polymerase; New England Biolabs) and the products cloned into the *Stu*I–*Bgl*II (Stat1α) or *Eco*RI–*Kpn*I (Stat1β) sites of the baculovirus transfer vector pAcSG2 (Pharmingen). Recombinant vectors were subsequently co-transfected with Baculogold baculovirus DNA (Pharmingen) into Sf9 insect cells as described (Gruenwald and Heitz, 1993) Recombinant viruses were identified by immunoblot of extracts of infected cells. For protein production Sf9 cells in suspension culture (0.8×10^6 cells/ml) were infected with recombinant viruses (mean of infection: 1.5) and harvested by centrifugation (1500 *g*, 15 min) 50 h post-infection.

The cells ($5–8\times10^8$) were lysed in 80 ml ice cold extraction buffer [20 mM MES, 100 mM KCl, 10 mM NaF, 10 mM Na_2HPO_4/NaH_2PO_4 pH 7.0, 10 mM $NaPP_i$, 0.02% NaN_3, 4 mM EDTA, 1 mM EGTA, 20 mM DTT, Complete™ protease inhibitors (Boehringer Mannheim), pH adjusted to 7.0 with 1 M Tris] with a Dounce homogenizer (2×10 strokes). All subsequent steps were performed at 4°C unless noted otherwise. Lysates were cleared by centrifugation at 20 000 *g* for 30 min. The supernatant was brought to pH 6.2 with 1 M MES and after the addition of 0.5 vol buffer 1 [20 mM MES, 0.02% NaN_3, 20 mM DTT, the pH was adjusted to 6.0 with 1 M Tris] was again centrifuged

for 20 min at 25 000 *g*. The resulting supernatant was loaded onto a S-Sepharose (Pharmacia) column (5×5.5 cm) and eluted with a linear salt gradient (50–300 mM KCl) and pH gradient (pH 6–7). Stat protein-containing fractions, located by immunoblot, were pooled, the pH adjusted to 8.0 with 1 M Tris and after the addition of 0.25 vol buffer 2 (20 mM Tris–HCl, 0.02% NaN_3, 10 mM DTT, pH 8.0) loaded onto a Q-Sepharose (Pharmacia) column (9×2 cm). This column was developed with a linear KCl gradient from 100–300 mM KCl. Eluted Stat1 proteins were precipitated with solid $(NH_4)_2SO_4$ to 60% saturation. The concentrated Stat proteins were dissolved in ~10 ml of buffer 3 [50 mM Na_2HPO_4/NaH_2PO_4 pH 7.2, 2 mM DTT, 1 mM EDTA, Complete™ protease inhibitors] and NEM (Sigma) was added to a final concentration of 20 mM. The alkylation reaction was incubated at room temperature for 10 min and then on ice for another 30 min. The reaction was stopped by the addition of β-mercaptoethanol to 50 mM and $(NH_4)_2SO_4$ to 0.5 M. The mixture was loaded on a low substituted phenyl-Sepharose (Pharmacia) column (15×2 cm) equilibrated in buffer 4 (20 mM Tris–HCl, 2 mM DTT, pH 7.4) + 0.5 M ammonium sulfate and the Stat proteins were eluted with decreasing $(NH_4)_2SO_4$ in buffer 4 at ~300 mM salt. Fractions of interest were pooled, concentrated with centriprep 50 (Amicon) to ~10 mg/ml and applied to a Superdex 200 column (XK 16, Pharmacia) equilibrated in buffer 5 (20 mM HEPES–HCl, 0.02% NaN_3, 2 mM DTT, 0.3 M KCl, pH 7.2). Fractions containing Stat1α or Stat1β, which both eluted very early, e.g. with a volume typical for globular proteins of M_r 350 kDa, were pooled, concentrated by ultrafiltration to ~20 mg/ml and quick frozen on dry ice. The purified proteins were stored at –70°C. All buffers used during protein purification were chilled, thoroughly degassed and flushed with N_2 before use.

Expression and purification of Stat1tc

The portion of the human Stat1 gene encoding residues 132–713 was amplified by PCR (Vent-Polymerase). The following primers were used: 5′-dGGGAATTCCATATGAGCACAGTGATGTTAGACAAAC and 5′-dCGGATCCTATTAGTGAACTTCAGACACAGAAATC (restriction sites underlined). The product was cloned into the *Nde*I–*Bam*HI sites of the pET20b expression vector (Novagen). N-terminal sequencing revealed the absence of the methionine residue introduced with the *Nde*I restriction site. Growth and induction of transformed *E.coli* [BL21DE3 (pLysS)] was as described (Studier and Moffatt, 1986). About 50% of the induced protein remained soluble and was subsequently isolated. Cells were collected by centrifugation (20 min; 4°C; 20 000 *g*) and resuspended in ice cold extraction buffer [100 ml/30 g cells; 20 mM HEPES–HCl, 0.1 M KCl, 10% glycerol, 1mM EDTA, 10 mM $MnCl_2$, 20 mM DTT, 100 U/ml DNase I (Boehringer Mannheim), Complete™ protease inhibitor, pH 7.6]. Cells were lysed by three cycles of freezing–thawing. Lysis was continued at 4°C while stirring slowly for 1 h. The lysate was centrifuged for 20 min at 22 000 *g* at 4°C. Polyethylenimine (0.1% final; Sigma) was added to the supernatent, the solution gently mixed and centrifuged for 15 min at 15 000 *g*. All subsequent steps were performed in the cold (4°C) unless stated otherwise. The supernatant containing soluble Stat1tc was precipitated with saturated ammonium sulfate solution (ultrapure; Gibco) in two steps (0–35%; 35–55% saturation final). The 35–55% pellet was redissolved in 20 ml of buffer 3 (see above) and alkylated as described above. The reaction was stopped by the addition of β-mercaptoethanol to 50 mM and solid ammonium sulfate to 0.9 M. The mixture was loaded onto a Fast Flow phenyl-Sepharose column (low substituted, 15×2 cm) that had been equilibrated in buffer A (50 mM Tris–HCl, 1 mM EDTA, 0.02% NaN_3, 2 mM DTT, pH 7.4) + 0.9 M ammonium sulfate. After washing the column a linear gradient from 0.9 to 0.05 M ammonium sulfate in buffer A was applied. Stat1tc eluted at ~0.5 M salt and the Stat1tc-containing fractions were pooled and dialyzed overnight against 2×4 l buffer B (40 mM MES–NaOH, 10% glycerol, 0.5 mM EDTA, 0.02% NaN_3, pH 6.5) + 140 mM KCl. This material was loaded onto a S-Sepharose column (5×5.5 cm) and the protein eluted in a linear 500 ml gradient of buffer B containing 140–300 mM KCl at ~220 mM. Fractions of interest were collected and dialyzed against 3 l buffer C (50 mM Tris–HCl, 10% glycerol, 2 mM DTT, pH 8) + 50 mM KCl with one change of buffer. The protein solution in buffer C + 50 mM KCl was applied to a Q-Sepharose column (9×2cm) and bound proteins were eluted with a linear gradient from 50 to 300 mM KCl in buffer C. Fractions with Stat1tc were combined and precipitated with solid ammonium sulfate to 55% saturation. At this stage the 95% pure preparation could be stored at –20°C until subjected to *in vitro* phosphorylation (see below) or was directly loaded onto a Superdex 200 gel filtration column (XK 16; Pharmacia). In this case the precipitated protein was dissolved in ~2 ml 10 mM HEPES–HCl, 100 mM KCl, 2 mM DTT, 0.5 mM EDTA, pH 7.4 and gel filtrated in this buffer. Stat1tc eluted in a symmetrical peak and was concentrated to ~20 mg/ml (Centriprep 50), quick frozen on dry ice and stored at –70°C. Typically yields of 40–50 mg >98% pure (as judged by Coomassie stain and mass spectroscopy) Stat1tc from 6 l of starting culture could be obtained.

Determination of protein concentrations

Purified proteins were quantified by UV spectroscopy. The absorbance ε of a 1 cm path length in a 1 mg/ml solution is given by $[(5700W + 1300Y)/M_r]$ with W = number of tryptophans; Y = number of tyrosines and M_r = molecular weight (Cantor and Schimmel, 1980). The following values were used: Stat1α: ε = 1.25; Stat1β: ε = 1.31; Stat1tc: ε = 1.27.

Proteolytic digestion of Stat1α and N-terminal sequencing of fragments

Proteinase K and subtilisin (Sigma) digests of purified Stat1α were carried out for 30 min on ice. The cleavage buffer contained 20 mM HEPES–HCl, 50 mM ammonium sulfate and 10 mM $MgCl_2$, pH 7.4. Reactions were stopped by the addition of PMSF (2 mM final) and SDS-sample buffer. The proteolysis was resolved on a 10% or 16.5% SDS–PAGE gel, which was either stained with Coomassie blue or electrotransferred onto a PVDF membrane (Immobilon P^{SQ}; Millipore). Sequencing of the 65 kDa protease-resistant Stat1α fragment N terminus was as described (LeGendre and Matsudaira, 1988). N-terminal sequence analysis was performed by the Protein/DNA facilities at Rockefeller University.

Cyanogen bromide and Endo AspN digests with mass spectrometric peptide analysis

Cyanogen bromide (Sigma) digests were performed on 90 pmol of recombinant protein in 50% formic acid at 25°C in the dark. Endoproteinase AspN (sequencing grade; Boehringer Mannheim) digests were carried out on 100–150 pmol of protein in either 25 mM Tris–HCl (pH 7.5) or 10 mM ammonium phosphate buffer (pH 8) with 150 mM KCl at 25°C. The protease:protein ratio was 1:50 w/w. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) was used to evaluate the peptide fragments. Aliquots (0.5 µl) of the digest were taken at various intervals (1 min–7 h), directly mixed into the MALDI-MS matrix solution (Cohen and Chait, 1996), and subject to MALDI-MS analysis in a procedure reported earlier (Cohen *et al.*, 1995).

Preparation of EGF receptor kinase and in vitro phosphorylation of Stat proteins

Human carcinoma A431 cells were grown to 90% confluency in 150 mm diameter plates in Dulbecco's modified Eagle's medium supplemented with 10% bovine calf serum (Hyclone). Cells were washed once with chilled PBS and lysates were prepared in 1 ml ice cold lysis buffer (10 mM HEPES–HCl, 150 mM NaCl, 0.5% Triton X-100, 10% glycerol, 1 mM Na_3VO_4, 10 mM EDTA, Complete™ protease inhibitors, pH 7.5). After 10 min on ice, the cells were scraped, vortexed and Dounce homogenized (five strokes). The lysates were cleared by centrifugation at 4°C for 20 min at top speed in an Eppendorf microfuge and stored at –70°C until needed. Immediately before use the lysates were diluted 1+4 with lysis buffer ('diluted lysate').

EGF receptor precipitates were obtained by incubating 5 ml of diluted lysate with 50 µg of anti-EGF receptor monoclonal antibody 108 (a kind gift of Dr J.Schlessinger, New York University) directed against the extracellular domain. After 2 h of rotating the sample at 4°C, 750 µl protein A–agarose (50% slurry; Oncogene Science) was added, and the incubation was allowed to proceed, while rotating, for another 1 h. Agarose beads containing the EGF receptor immunoprecipitates were then washed five times with lysis buffer and finally twice with storage buffer (20% glycerol, 20 mM HEPES–HCl, 100 mM NaCl, 0.1 mM Na_3VO_4). Precipitates from 5 ml diluted lysate were dissolved in 0.5 ml storage buffer, flash frozen on dry ice and stored at –70°C. Immediately before an *in vitro* kinase reaction the protein A–agarose bound EGF receptor from 5 ml dilute lysate was washed once with 1× kinase buffer (20 mM Tris–HCl, 50 mM KCl, 0.3 mM Na_3VO_4, 2 mM DTT, pH 8.0) and then dissolved in 0.4 ml (total volume) of this buffer. Afterwards the washed EGF receptor precipitate was incubated on ice for 10 min in the presence of a final concentration of mouse EGF (a kind gift of Dr S.Cohen, Vanderbilt University) of 0.15 ng/µl. Phosphorylation reactions were carried out in Eppendorf tubes in a final volume of 1 ml. To the pre-incubated kinase preparation the following was added: 60 µl 10× kinase buffer, 20 µl 0.1 M DTT, 50 µl 0.1 M ATP, 4 mg Stat protein (Superdex 200 eluate for Stat1α/β; ammonium sulfate pellets, dissolved in 20 mM Tris–HCl pH 8.0, for Stat1tc), 10 µl 1 M $MnCl_2$

and dH_2O to 1 ml. The reaction was allowed to proceed for 15 h at 4°C. After 3 h an additional 15 µl of 0.1 M ATP was added.

Separation of phosphorylated from unphosphorylated Stat proteins

The *in vitro* kinase reaction mixture (see above) was freed from EGF receptor bound to agarose beads by spinning through a plug of siliconized glass wool at the bottom of a pierced Eppendorf tube. The glass wool was washed with 0.5 ml HA-buffer (20 mM Tris–HCl, 1 mM EDTA, 2 mM DTT, pH 8.0) and the pooled volumes loaded onto a heparin–agarose (Bio-Rad) column (7×1.5 cm). The column was washed with 50 ml HA-buffer, and then elution of the bound Stat proteins with two consecutive 50 ml volumes of HA-buffer +150 mM KCl and HA-buffer +400 mM KCl followed. Unphosphorylated proteins (eluted with 150 mM KCl) were concentrated by ultrafiltration to ~10 mg/ml, flash frozen on dry ice and stored at –70°C. Phosphorylated Stat1α/β was concentrated to 1 mg/ml, glycerol was added to 50% (v/v) and the material was stored at –20°C. Phosphorylated Stat1tc was brought to a concentration of ~15 mg/ml and run on a Superdex 200 columns under the conditions described above for the native protein. The gel filtered phosphorylated Stat1tc was pooled, concentrated to ~20 mg/ml, flash frozen on dry ice and stored at –70°C.

EMSA

A 12.5 µl reaction volume contained DNA binding buffer (20 mM HEPES–HCl, 4% Ficoll, 40 mM KCl, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 1 mM DTT) and was mixed with radiolabelled DNA (see below) at a final concentration of 1×10^{-10} M unless stated otherwise, 50 ng dIdC, 2.5 µg BSA (Boehringer Mannheim), and the indicated amount of purified phosphorylated Stat1. Incubation followed at room temperature. The time necessary to reach equilibrium was assessed by EMSA (not shown; Stone *et al.*, 1991). For all DNA fragments tested equilibrium turned out to be fully established at the earliest timepoints that can be determined by this technique (30 s). Therefore incubation periods of 5–15 min were chosen. Reaction products were loaded onto a 4% polyacrylamide gel (1.5 mm thick) containing 0.25× Tris-borate-EDTA which had been pre-run at 20 V/cm for 2 h at 4°C. Electrophoresis was continued for 60 min at 4°C. Gels were dried and exposed to X-ray film and quantified by a Molecular Dynamics PhosphorImager.

Binding site oligonucleotides

Single-stranded oligonucleotides that were purified on the basis of trityl affinity were obtained from The Great American Gene Company (Ransom Hill). Oligonucleotides longer than 30 nt were further purified on 6% sequencing gels and DNA recovered by soak elution and ethanol precipitation. Nucleic acid concentrations were determined by absorbance at 260 nm using the calculated molar extinction coefficient for each oligonucleotide (corrected for the hyperchromic effect). Complementary oligonucleotides at a concentration of 1 pmol/µl were hybridized for 3 h after thermal denaturation in 5 mM Tris–HCl, 50 mM KCl, 10 mM $MgCl_2$, pH 8.0. One pmol of synthetic duplex molecule was labelled to high specific activity by the Klenow fill-in reaction [0.5 mM dATP (and 0.5 mM dCTP for S1), 100 µCi [α-^{32}P]dGTP (3000 Ci/mmol; 10 mCi/ml and [α-^{32}P]dTTP for S1; Du Pont), 5 U Exo$^-$Klenow enzyme (New England Biolabs)] and rendered completely double-stranded with a 0.5 mM dGTP (and 0.5 mM dTTP for S1) cold chase. Unincorporated nucleotides were removed by gel filtration (spin quant columns; Pharmacia) in 10 mM Tris–HCl, 100 mM NaCl, 1 mM EDTA, pH 8.0. Labelled oligonucleotides were stored at 4°C.

The following duplex DNA fragments with protruding 5′-TCC (except for S1 which has 5′-GATC) were used (the core recognition sequence is underlined): *cfos*WT 5′-dGTATTCCCGTCAATGCA-3′; *Lv6* E 5′-dGTATTCCTGTAAGATCT-3′; *cfos*M67 5′-dGATTTCCCGTAAAT-CAT-3′; S1 5′-dGTTGTTCCGGGAAAAGG-3′; 2× *cfos*WT (10 bp spacing) 5′-dAGTCAGTTCCCGTCAATGCATCAGGTTCCCGTCA-ATGCAT-3′; 2× *cfos*WT (5 bp spacing) 5′-dAGTCAGTTCCCGT-CAATGAGTTCCCGTCAATGCA-3′; 2× *cfos*WT (15 bp spacing) 5′-dAGTCAGTTCCCGTCAATGATCGCTACAGAGTTCCCGTCAA-GCA-3′; 2× *cfos*WT (inverted repeat) 5′-dAGTCATTTCCCGTCAAT-GCATCAGTTG ACGGGAAAGTAGT-3′.

Dissociation rate determination

Under the reaction conditions described above, each oligonucleotide (at 2×10^{-9} M or stated otherwise) was mixed with 0.55×10^{-9} M dimer of purified phosphorylated Stat1 protein. The reaction volume was scaled up to 100 µl. The reaction was incubated for 5–15 min at room temperature and for time zero, an aliquot (10 µl) was removed and loaded directly onto a pre-run polyacrylamide gel (see EMSA). Afterwards, a 100× molar excess of homologous unlabelled DNA (in <1% of the reaction volume) was added. At subsequent time points (indicated in Figures 5B, 6 and 7) 10 µl aliquots were withdrawn and also loaded onto the running gel (at 10 V/cm). After entering the final time point (after 30–45 min), electrophoresis was continued at 20 V/cm until the unbound labelled DNA fraction reached the bottom of the gel. Gels were dried, exposed to X-ray film and labelled protein–DNA complexes and unbound labelled DNA were quantified as described above. The half-life was determined from a semi-log plot of the numerical data (shifted radioactivity over shifted radioactivity at time zero versus time). For many sequences studied the half-life was too short (>30 s) to be determined by EMSA. All experiments were performed at least twice with the different oligonucleotides.

Determination of apparent equilibrium constants for protein–DNA interactions

A fixed quantity of ^{32}P-labelled oligonucleotide varied between 1×10^{-10} M and 5.6×10^{-10} M in three separate experiments, was titrated against a standard protein dilution series (common to all oligonucleotides tested) in a volume of 12.5 µl under the reaction conditions described above. Numerical data were used to construct a standard binding curve from which the free dimer concentration, when 50% of the probe is shifted, could be determined.

Acknowledgements

We thank Raymond C.N.Chiu, Steven Jacques and Huguette Viguet for expert technical assistance, and Lois Cousseau for preparing the manuscript. Protein sequence analysis was provided by The Rockefeller University/DNA Technology Center, which is supported in part by NIH shared instrumentation grants and by funds provided by the US Army and Navy for purchase of equipment. This work was supported by NIH grants AI32489 and AI34420 to J.E.D, and NIH grant RR00862 to B.T.C. U.V. was the recipient of an EMBO long-term fellowship.

References

Affolter,M., Percival-Smith,A., Muller,M., Leupin,W. and Gehring,W.J. (1990) DNA binding properties of the purified Antennapedia homeodomain. *Proc. Natl Acad. Sci. USA*, **87**, 4093–4097.

Bradford,M.M. (1976) A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–251.

Cantor,C.R. and Schimmel,P.R. (1980) In *Biophysical Chemistry. Part II Techniques for the Study of Biological Structure and Function*. W.H.Freeman and Company, San Francisco, CA, pp. 380–381.

Cohen,S.L. and Chait,B.T. (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal. Chem.*, **68**, 31–37.

Cohen,S.L., Ferre-D'Amare,A.R., Burley,S.K. and Chait,B.T. (1995) Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Sci.*, **4**, 1088–1099.

Darnell,J.E.,Jr, Kerr,I.M. and Stark,G.M. (1994) Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, **264**, 1415–1421.

Fried,M. and Crothers,D.M. (1981) Equilibria and kinetics of lac repressor–operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.*, **9**, 6505–6525.

Fu,X.-Y. (1992) A transcription factor with SH2 and SH3 domains is directly activated by an interferon α-induced cytoplasmic protein tyrosine kinase(s). *Cell*, **70**, 323–335.

Garner,M.M. and Revzin,A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.*, **9**, 3047–3060.

Gruenwald,S. and Heitz,J. (1993) *Baculovirus Expression Vector System: Procedures and Methods Manual*. 2nd Edn. PharMingen, San Diego, CA.

Guyer,N.B., Severns,C.W., Wong,P., Feghali,C.A. and Wright,T.M. (1995) IFN-γ induces a p91/Stat1α-related transcription factor with distinct activation and binding properties. *J. Immunol.*, **155**, 3472–3480.

Horvath,C.M., Wen,Z. and Darnell,J.E.,Jr (1995) A STAT protein domain that determines DNA sequence recognition suggests a novel DNA-binding domain. *Genes Dev.*, **9**, 984–994.

Ihle,J.N. (1995) Cytokine receptor signalling. *Nature*, **377**, 591–594.

LeGendre,N. and Matsudaira,P. (1988) Direct protein microsequencing from Immobilon-P transfer membrane. *Biotechniques*, **6**, 154–159.

Levy,D.E., Kessler,D.S., Pine,R.I. and Darnell,J.E.,Jr (1989) Cytoplasmic activation of ISGF3, the positive regulator of interferon-α stimulated transcription, reconstituted *in vitro*. *Genes Dev.*, **3**, 1362–1372.

Muller,M., Laxton,C., Briscoe,J., Schindler,C., Improta,T., Darnell, J.E.,Jr, Stark,G.R. and Kerr,I.M. (1993) Complementation of a mutant cell line: Central role of the 91-kDa polypeptide of ISGF3 in the interferon-α and -γ signal transduction pathway. *EMBO J.*, **12**, 4221–4228.

Quelle,F.W., Thierfelder,W., Witthuhn,B.A., Tang,B., Cohen,S. and Ihle,J.N. (1995) Phosphorylation and activation of the DNA binding activity of purified Stat1 by the Janus protein tyrosine kinases and epidermal growth factor receptor. *J. Biol. Chem.*, **270**, 20775–20780.

Qureshi,S.A., Salditt-Georgieff,M. and Darnell,J.E.,Jr (1995) Tyrosine phosphorylated Stat1 and Stat2 plus a 48 kDa protein all contact DNA in forming the interferon stimulated gene factor 3 (ISGF3) *Proc. Natl Acad. Sci. USA*, **92**, 3929–3833.

Riggs,A.D., Suzuki,H. and Bourgeois,S. (1970) *lac* repressor-operator interaction. I. Equilibrium studies. *J. Mol. Biol.*, **48**, 67–83.

Schindler,C. and Darnell,J.E.,Jr (1995) Transcriptional responses to polypeptide ligands: The JAK-STAT pathway. *Annu. Rev. Biochem.*, **64**, 621–51.

Schindler,C., Fu,X.-Y., Improta,T., Aebersold,R. and Darnell,J.E.,Jr (1992) Proteins of transcription factor ISGF-3: One gene encodes the 91 and 84 kDa ISGF-3 proteins that are activated by inteferon-α. *Proc. Natl Acad. Sci. USA*, **89**, 7836–7839.

Shuai,K., Schindler,C., Prezioso,V.R. and Darnell,J.E.,Jr (1992) Activation of transcription by IFN-γ: tyrosine phosphorylation of a 91 kDa DNA binding protein. *Science*, **259**, 1808–1812.

Shuai,K., Stark,G.R., Kerr,I.M. and Darnell,J.E.,Jr (1993a) A single phosphotyrosine residue of Stat91 required for gene activation by interferon-γ. *Science*, **261**, 1744–1746.

Shuai,K., Ziemiecki,A., Wilks,A.F., Harpur,A.G., Sadowski,H.B., Gilman,M.Z. and Darnell,J.E.,Jr (1993b) Polypeptide signaling to the nucleus through tyrosine phosphorylation of JAK and STAT proteins. *Nature*, **366**, 580–583.

Shuai,K., Horvath,C.M., Tsai-Huang,L.H., Qureshi,S., Cowburn,D. and Darnell,J.E.,Jr (1994) Interferon activation of the transcription factor Stat91 involves dimerization through SH2-phosphotyrosyl peptide interactions. *Cell*, **76**, 821–828.

Stone,S.R., Hughes,M.J. and Jost,J.P. (1991) Qualitative and quantitative studies of protein–DNA interactions by gel mobility-shift assay. In Jost,J.P. and Saluz,H.P. (eds), *A Laboratory Guide to In Vitro Studies of Protein-DNA Interactions. BioMethods.* Vol. 5. pp. 163–195.

Studier,F.W. and Moffatt,B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.*, **189**, 113–130.

Wagner,B.J., Hayes,T.E., Hoban,C.J. and Cochran,B.H. (1990) The SIF binding element confers sis/PDGF inducibility onto the c-*fos* promoter. *EMBO J.*, **9**, 4477–4484.

Wen,Z., Zhong,Z. and Darnell,J.E.,Jr (1995) Maximal activation of transcription of Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell*, **82**, 241–250.

Yan,H., Krishnan,K., Greenlund,A.C., Gupta,S., Lim,J.T.E., Schreiber, R.D., Schindler,C.W. and Krolewski,J.J. (1996) Phosphorylated interferon-α receptor 1 subunit (IFNaR1) acts as a docking site for the latent form of the 113 kDa STAT2 protein. *EMBO J.*, **15**, 1064–1074.

Yan,R., Qureshi,S., Zhong,Z., Wen,Z. and Darnell,J.E.,Jr (1995) The genomic structure of the STAT genes: Multiple exons in coincident sites in Stat1 and Stat2. *Nucleic Acids Res.*, **23**, 459–463.

Yan,R., Small,S., Desplan,C., Dearolf,C.R. and Darnell,J.E.,Jr (1996) Identification of a Stat gene that functions in *Drosophila* development. *Cell*, **84**, 421–430.

Yarden,Y., Harari,I. and Schlessinger,J. (1985) Purification of an active EGF receptor kinase with monoclonal antireceptor antibodies. *J. Biol. Chem.*, **260**, 315–319.

Received on May 28, 1996; revised on July 1, 1996