THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 9, Issue of March 2, pp. 6675–6688, 2001
Printed in U.S.A.

DNA Binding Specificity of Different STAT Proteins

COMPARISON OF *IN VITRO* SPECIFICITY WITH NATURAL TARGET SITES*[S]

Received for publication, March 2, 2000, and in revised form, October 23, 2000
Published, JBC Papers in Press, October 26, 2000, DOI 10.1074/jbc.M001748200

Georg B. Ehret‡§¶, Patrick Reichenbach‡, Ulrike Schindler‖, Curt M. Horvath, Stefan Fritz‡‡§§, Markus Nabholz‡, and Philipp Bucher‡**

*From the ‡Swiss Institute for Experimental Cancer Research (ISREC) 1066 Epalinges, Switzerland, ‖Tularik Inc., South San Francisco, California 94080, **Mount Sinai School of Medicine, Immunobiology Center, New York, New York 10029-6574, and the ‡‡Institute for Microbiology, Biochemistry and Genetics, University of Erlangen-Nürnberg, D-91058 Erlangen, Germany*

STAT transcription factors are expressed in many cell types and bind to similar sequences. However, different STAT gene knock-outs show very distinct phenotypes. To determine whether differences between the binding specificities of STAT proteins account for these effects, we compared the sequences bound by STAT1, STAT5A, STAT5B, and STAT6. One sequence set was selected from random oligonucleotides by recombinant STAT1, STAT5A, or STAT6. For another set including many weak binding sites, we quantified the relative affinities to STAT1, STAT5A, STAT5B, and STAT6. We compared the results to the binding sites in natural STAT target genes identified by others. The experiments confirmed the similar specificity of different STAT proteins. Detailed analysis indicated that STAT5A specificity is more similar to that of STAT6 than that of STAT1, as expected from the evolutionary relationships. The preference of STAT6 for sites in which the half-palindromes (TTC) are separated by four nucleotides (N_4) was confirmed, but analysis of weak binding sites showed that STAT6 binds fairly well to N_3 sites. As previously reported, STAT1 and STAT5 prefer N_3 sites; however, STAT5A, but not STAT1, weakly binds N_4 sites. None of the STATs bound to half-palindromes. There were no specificity differences between STAT5A and STAT5B.

STAT[1] proteins were discovered in the course of the analysis of interferon signaling pathways (1, 2). Up to now seven mammalian genes coding for members of this family of intracellular signaling proteins have been found. STAT molecules are present as latent transcription factors in the cytoplasm. Signaling by a large number of cytokine, growth factor, and hormone receptors leads to activation of one or more STATs by JAK kinases or intrinsic receptor tyrosine kinase activity. The interaction between the SH2 domain present in all STATs and the STAT phosphotyrosine is the basis for the formation of STAT dimers with 2-fold symmetry. The dimers are transported into the nucleus where they bind to a palindromic DNA motif present in cytokine-inducible genes (see Refs. 3 and 4 for review). The same palindromic core motif, $TTCN_{2-4}GAA$, has been found in sequences recognized by all STATs.

Initially, the STAT signaling pathway was proposed to account for specificity of cellular responses to different ligands, because STAT molecules directly link the receptor to the regulatory elements in responsive genes. However, it was soon found that the same STAT molecule can be activated by different extracellular signals and that many cytokine receptors can activate more than one STAT protein. In view of this apparent promiscuity of receptor-STAT interactions, it was surprising to find that targeted disruption of STAT genes in mice resulted in very specific and distinct phenotypes. These results suggest that *in vivo*, in contrast to cell culture systems, an unknown number of receptors acts primarily through a single STAT species. They leave open, however, the question of how STATs select their target genes. The observation that the same STAT protein can induce distinct genes in different cell types indicates that the induction of these genes does not solely depend on the presence of active STAT but may require the cooperation of other proteins (5) or may be controlled by the state of the chromatin of a potential STAT target gene (6). In some cell types different receptors activate distinct STATs and induce expression of exclusive genes (7). In these and in other cases, target gene selection may also be determined by the capacity of different STATs to cooperate with distinct transcription factors that bind to regulatory elements of particular genes. Finally, target gene selection by STATs may be determined by preferences in the sequence specificity of different STAT dimers. Although the same palindromic core is found in sequences bound by all STATs, previous work (164) has shown that STATs can be distinguished on the basis of their sequence specificity. The most striking finding was that STAT6 differs from the other members of the family in that it prefers sites in which the two halves of the palindrome are separated by four rather than three nucleotides (8) ($\underline{TTC}N_4\underline{GAA}$ and $\underline{TTC}N_3\underline{GAA}$ hereafter referred to as N_4 and N_3 sites, respectively. We refer to this distance N_x as spacer.) The observation that the fine specificity for STAT1 and STAT3 is different when sites of low affinity are taken into account (9) gave rise to the suggestion that natural STAT responsive cis-acting elements may be weak affinity binding sites that preferentially bind specific STATs. In some cases the specificity of such sites may be increased by

* This work was supported in part by research grants from the Swiss National Science Foundation (to M. N.) and from the Deutsche Forschungsgemeinschaft Grant DFG SFB 473, project B6 (to G. Fey at the University Erlangen-Nürnberg). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at *http://www.jbc.org*) contains sequences selected by STAT1, STAT5A, and STAT6, and Fig. 5.

§ Present address Universität Heidelberg, Im Neuenheimer Feld 346, D-69120 Heidelberg, Germany.

¶ To whom correspondence should be addressed. Tel.: 41-21-692-58-92; Fax: 41-21-652-6933; E-mail: markus.nabholz@isrec.unil.ch.

§§ Present address: NanoTools GmbH, Tscheulinstrasse 21, D-79331 Teningen, Germany.

[1] The abbreviations used are: STAT, signal transducer and activator of transcription; PCR, polymerase chain reaction; IFN, interferon; IL, interleukin; oligos, oligonucleotides; HMM, hidden Markov model.

the arrangement in pairs that allow the cooperative binding of two STAT dimers (10–13).

Until now, there have been relatively few studies (9, 14, 15) comparing the sequence specificity of different STAT proteins, and in particular, there are few data on the relative affinity of weak binding sites that may be the physiological STAT targets. In the course of our previous work (6), we identified two STAT-binding sites in the interleukin-2 responsive enhancer of the *IL-2Rα* gene, and we showed that this enhancer responds to STAT5 but does not seem to respond to STAT1 although the sites do bind STAT1 *in vitro*. This led us to compare the fine specificities of STAT5 and STAT1. We included STAT6 in this study, since the domain responsible for sequence specificity is more similar to that of STAT5 than to the one of STAT1. First we analyzed base frequency matrices obtained from the sequences selected by STAT1, STAT5A, and STAT6 in binding site selection experiments. The results revealed that apart from its preference for N_4 sites, STAT6 selected sites that resembled more the ones selected by STAT5 than by STAT1. We compared these matrices to the results from experiments in which we quantified the relative affinity of a series of weak binding sites for STAT1, STAT5A and -B, and STAT6. We found that the principles that govern STAT binding to high affinity sequences equally apply to weak binding sites. Our results are in reasonable agreement with the sequence pattern among the putative target sites for various STAT proteins in cytokine-responsive genes. Thus, our results may be useful in predicting which STAT regulates expression of a gene that includes a STAT consensus site in its regulatory elements.

EXPERIMENTAL PROCEDURES

Recombinant Proteins and Cell Extracts—Preparation of histidine-tagged recombinant human STAT1 and human STAT6 proteins from insect have been described previously (8). The baculovirus expression vector pVL1393 (PharMingen) was used. Recombinant human STAT5A was prepared by the same method. Fig. 3*A* indicates that these preparations contain no major impurities. Recombinant STAT5B was also prepared using a baculovirus expression system. Briefly, rat STAT5B cDNA (without a histidine tag) was cloned into pFastBac1, and virus was produced in SF-21 cells. Subsequently, SF-21 cells were coinfected with STAT5B and murine JAK2 encoding viruses. 60 h post-infection protein extracts were purified by DNA affinity chromatography, using a biotinylated oligo with tandem STAT-binding sites and magnetic beads, according to Dreier *et al.* (16). SDS-polyacrylamide gel electrophoresis analysis (not shown) indicates that, as expected, this preparation is much less pure than that of the histidine-tagged proteins. Extracts containing endogenous STAT1 were produced as described previously (9) from BUD 8 fibroblasts.

DNA-binding Site Selection—The STAT1 and STAT6-binding site selection experiments that produced the data used in Fig. 2 have been described previously (8, 9). The selection for STAT1-binding sites was carried out on a pool of 76-base oligomers (26 random bases sandwiched between two constant 25-base regions containing primer recognition sites) with nuclear extract from interferon-γ-treated fibroblasts, according to Ref. 17. Bound oligonucleotides were precipitated with an anti-STAT1 antibody and protein A-agarose and subsequently amplified by PCR. Sequences were cloned after five rounds of selection amplification, and 52 clones were sequenced. The sequences of STAT5A- and STAT6-binding sites were selected from a pool of 50-base oligonucleotides with 14 randomized positions as follows: 5′-GTCTGTCTGAGGTGAGATCT-ATN_{14}ACAAGCTTGTCTAGCGACGTCGCG-3′. The constant regions 5′ and 3′ of the randomized positions contained a *Sac*I or a *Hin*dIII site, respectively, which was used for subsequent cloning. The binding procedure was carried out essentially as described (18), using recombinant STAT5 and STAT6. Three rounds of selection by bandshift were performed. Bound sequences were eluted, and 10 cycles of PCR amplification were used to amplify the selected oligos for the second and third round using the following conditions: 1 min at 94 °C, 1 min at 55 °C, and 30 s at 72 °C. PCRs were carried out in 50 μl according to the specifications of the manufacturer (PerkinElmer Life Sciences), except that dCTP was replaced by 3×10^{-5} μmol (3,000 Ci/mmol) of [α-^{32}P]dCTP. After the third mobility shift, the bound DNA was eluted, and a PCR with 20 cycles was carried out in the absence of radioactive dCTP. The resulting fragments were cloned and subjected to automated DNA sequencing from which 39 sequences (STAT5A) and 45 sequences (STAT6) were recovered.

Analysis of in Vitro Selected Binding Sites—A hidden Markov model (HMM) was chosen for a formal description of the binding specificity of individual STAT proteins in the site-selection experiments. The parameters of the STAT1-, STAT5-, and STAT6-specific binding site models shown in Fig. 2 were obtained by training the same initial model shown in Fig. 1 with the respective sequence sets. The Baum-Welch algorithm as implemented in SAM software release 1.3.3 was chosen as training method for this purpose. All oligonucleotides were presented in both orientations to the algorithm taking into account the perfectly symmetric nature of the STAT dimer.

Competitors and Probes—Oligonucleotides (see Table II for upper strands) were purchased from Microsynth (Balgach, Switzerland) or MWG Biotech (Ebersberg, Germany). Complementary oligonucleotides were annealed by heating and slow cooling in 10 mM Tris (pH 7.4), 1 mM EDTA, 50 mM NaCl and kept at −80 °C. To control for completeness of annealing, 100 ng of double-stranded oligonucleotide was fractionated on a nondenaturing 12% polyacrylamide gel that was stained with ethidium bromide after the run. Oligonucleotide concentrations were verified by running aliquots on a 12% polyacrylamide gel and comparing the ethidium bromide signal with that of a titration series of the reference competitor loaded on the same gel. Adjustments for the differences in composition of the sequences were made. Serial 2-fold dilutions of the stock solutions were used in the competition assays.

Competition Assays—Electrophoretic mobility shift assays were carried out with the STAT1-responsive element of the *Fc-γRI* gene (referred to as GRR (19)) as a probe, which was end-labeled according to standard procedures (20) with [γ-^{32}P]ATP (Amersham Pharmacia Biotech) and purified on a 12% polyacrylamide gel. After precipitation, the probe was resuspended in 10 mM Tris (pH 7.4), 1 mM EDTA, 50 mM NaCl and kept at 4 °C. 30,000 cpm/sample were used per bandshift reaction.

The binding buffer for the bandshift reactions contained 10 mM Tris (pH 7.4), 1 mM EDTA, 140 mM KCl, 10% glycerol, 1 mM dithiothreitol, 10 mM sodium hydrogen phosphate, 4 mM urea, 0.1 mM ATP, 20 mM spermine, 1 mg/ml bovine serum albumin, 5 mg/ml poly(dI-dC).

To optimize binding conditions for recombinant STAT proteins, we assessed several substances for their capacity to influence STAT-DNA interactions. Among the compounds tested, the naturally occurring polyanion spermine considerably increased recombinant STAT binding. Maximal effects were obtained at the concentration specified above. Note that spermine had no effect on the formation of STAT5 or STAT6 complexes using crude nuclear extracts from IL-2- or IL-4-stimulated cells. Salt concentration was chosen to equal approximately the intracellular concentrations.

For the competition experiments a constant amount of probe was mixed with graded amounts of competitor, before addition of STAT protein. In each experiment reactions with different concentrations of unlabeled reference GRR oligonucleotide were included.

The following amounts of recombinant proteins were used: STAT5A, 15 ng/sample; STAT5B, 150 ng/sample; STAT6, 50 ng/sample; STAT1, 10 ng/sample. Binding reactions (22 μl) were incubated for 150 min at 29 °C and then loaded onto a running 4% nondenaturing polyacrylamide gel, which had been prerun at 4 °C for 30 min. The binding reaction had reached equilibrium at the time of loading on the gel.

RESULTS

In Vitro Selection Experiments—Data from binding site selection experiments with human STAT1 and STAT6 have been partly published previously, and we used the sequences derived from these experiments for the analysis described here. For human STAT5A we carried out a similar procedure, starting with randomized 14-mers placed between PCR-priming sites. Three cycles of selection were carried out by isolation of bound fragments from a bandshift procedure, and the resulting sequences were cloned. 39 clones were subjected to sequencing (for details see "Experimental Procedures").

These three oligonucleotide sequence sets constitute the raw data from which the binding site models specific for the individual STAT family member were generated and they are included as Supplemental Material.

Initial Assumptions, Choice of the Binding Site Model—The first critical step in the characterization of the binding speci-

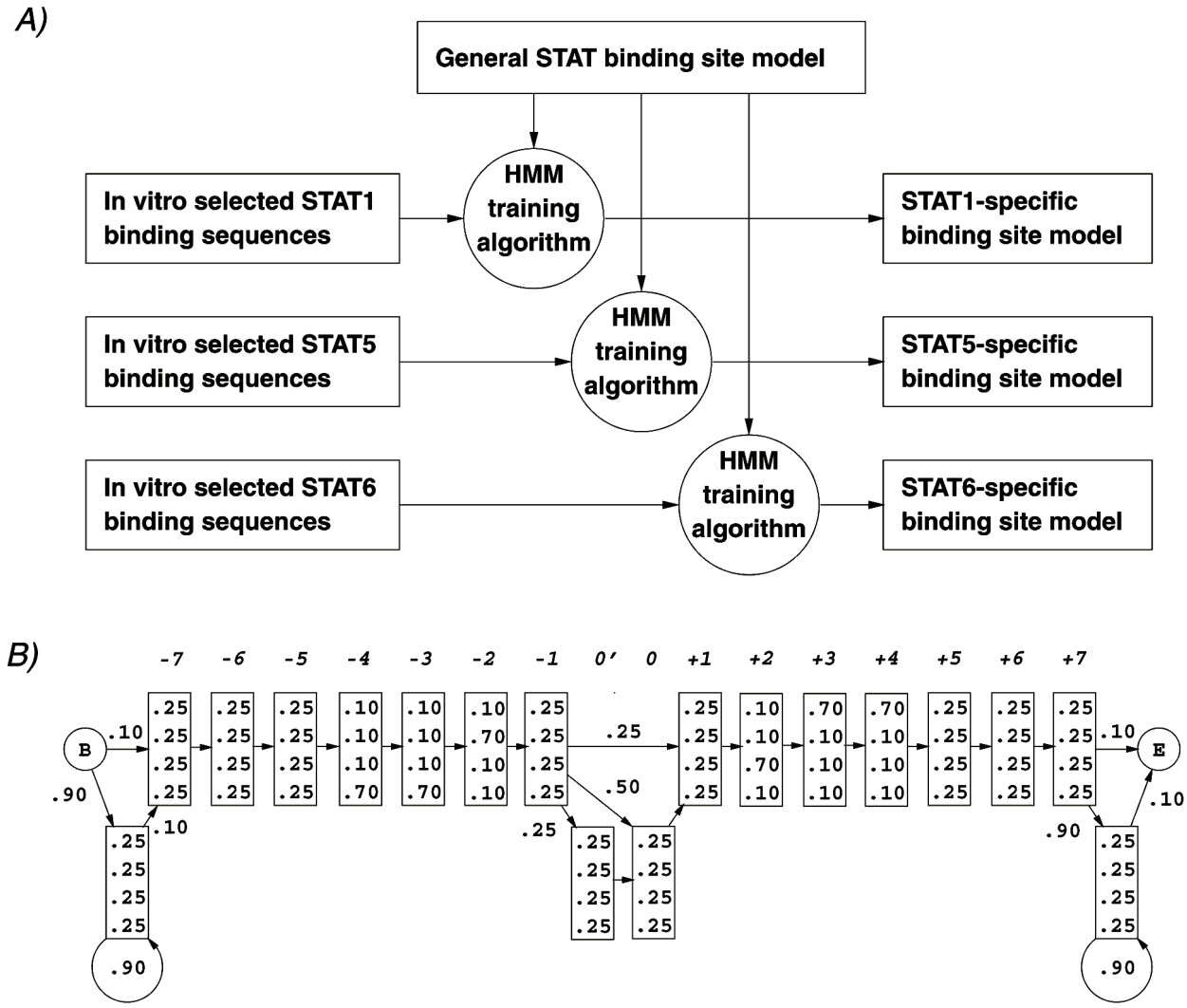

FIG. 1. **Analysis of *in vitro* selected STAT1, STAT5A, and STAT6 sequences.** *A,* principle of the method and data flow. The individual protein-specific models were obtained with an HMM training algorithm starting from the same general STAT binding site model. The data flow diagram serves to make clear that all three sequence sets were subjected to the same data analysis procedure. We therefore claim that the differences between the resulting STAT1-, STAT5-, and STAT6-specific models shown in Fig. 2 are entirely data driven. *B,* initial model used in the model building process. This HMM expresses our assumptions about a general STAT protein-binding oligonucleotide. HMMs are probabilistic constructs that define families of related sequences (23). The *rectangular boxes* define the base frequencies at particular positions in the model (the order is A, C, G, T from *top* to *bottom*). The *boxes* numbered -7 to -1 and $+1$ to $+7$ correspond to individual base positions. The half-sites are connected by three alternative paths containing 0-, 1-, and 2-base positions, respectively. The *arrows pointing to the right* from model position -1 define the initial frequencies of the three spacer classes. Note that the half-site regions are seeded with a moderate preference for TTC and GAA (70% for the consensus base, 10% for each mismatch). The *boxes* with *loops* at the beginning and at the end of the model represent random sequences of the oligonucleotides outside the binding regions. The trained models have the same architecture as the input model but different parameter values that are shown in Fig. 2 (numerical values are shown in Table I).

ficity of a transcription factor is the choice of the binding site model. As mentioned, all STAT proteins are thought to recognize the same palindromic core motif, $TTCN_{2-4}GAA$ (where N_x is referred to as spacer). However, the specificities of individual STAT proteins may differ in the following aspects: (i) differential tolerance of specific mismatches in the core motif; (ii) preference of certain bases at flanking positions; and (iii) preferences for different spacer lengths between the core motif half-sites. In our model, we included three flanking positions upstream and two flanking positions downstream of the TTC half-site consensus triplet. The choice of the length of the sequence analyzed was based on the most distal contact observed in the co-crystal structure, the G-C base pair located three positions upstream of the consensus triplet (21). The number of two downstream positions is dictated by the maximal spacer length. Also on the basis of crystallographic data, we assumed perfect dyad symmetry of the binding site, and we used each selected sequence in both orientations. Numerically, we express the position-specific base preferences and the spacer length preferences as percentages. Our model thus contains 35 parameters, 4×8 relating to the different bases at different half-site positions, and 3 to alternative spacer lengths.

Derivation of STAT1-, STAT5-, and STAT6-specific High Affinity Binding Site Models from in Vitro Selected Oligonucleotides—Estimating the parameters of a binding site model of the kind described above from *in vitro* selected oligonucleotide sequences is not trivial because the oligonucleotides are longer than the actual binding site, and the location of the binding site within an oligonucleotide is not often not easy to determine. A common approach in the analysis of such data is to manually align putative binding regions according to some prior knowledge and to tabulate the base frequencies observed in consecutive columns of the multiple alignment. The drawbacks of such a method are its lack of reproducibility and circularity. It will be unclear to what extent such a manual alignment rep-

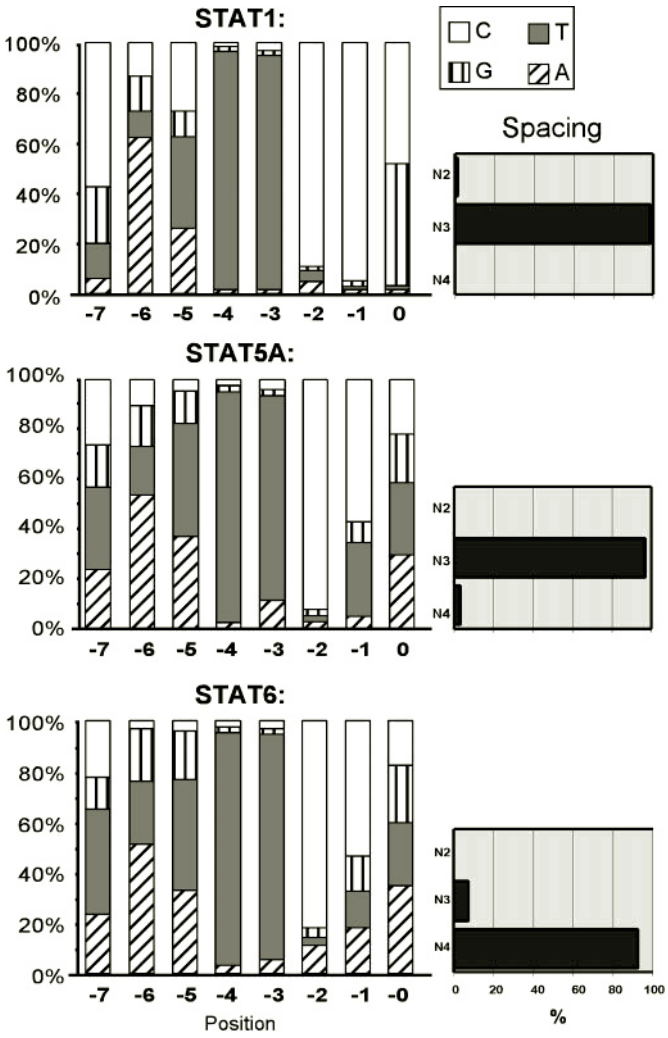

FIG. 2. **High affinity binding site models for STAT1, STAT5, and STAT6.** The models were derived from primary data (*in vitro* selected oligonucleotide sequences) using an HMM training algorithm as described in the text and in Fig. 1. Only one half-site is shown since we submitted each oligonucleotide in both orientations to the data analysis procedure, and the result was entirely symmetric. The position notation is shown in Fig. 1. The spacer refers to the number x of nucleotides between the canonical half-sites ($TTCN_xGAA$). Spacer length frequencies are the numbers associated with the *three arrows pointing to the right* from box −1 of the trained models in Fig. 1.

resents initial assumptions or the true binding specificity of the protein used in the *in vitro* selection experiment. This type of procedure is particularly inappropriate in a case like ours where the goal of the analysis is to reveal subtle differences in the binding preferences of members of the same transcription factor family. Circular reasoning is a major problem when dealing with spacer length variation. If, for instance, we align STAT5-binding sites by making the assumption that STAT5 can only bind to N_3 sites, we would not find any N_4 site in the resulting multiple alignment, but this would reflect our initial assumptions rather than an empirical finding. To make claims about differences between binding specificities, especially with regard to spacer length preferences of individual STAT family members, we have to subject the *in vitro* selected oligonucleotide sets to exactly the same procedure. This could be achieved by applying the same *ab initio* method to all three data sets or by using a model refinement procedure starting from the same general STAT-binding site model. Since we do not know of any robust algorithm capable of finding the optimal binding site model without using an initial assumption for training sets of

TABLE I
Base frequency matrices of STAT half-sites derived from selection experiments (see Fig. 2)

	POSITION:[a]								Spacer[b]	
STAT1	-7	-6	-5	-4	-3	-2	-1	0	N_x	%
A	6	62	26	*2*	*2*	*5*	2	2	N_2	1.5
T	14	11	36	*95*	*93*	*4*	2	2	N_3	98.5
G	23	14	10	*2*	*2*	*2*	2	49	N_4	0
C	57	13	27	*2*	*3*	*89*	95	48		
STAT5A	-7	-6	-5	-4	-3	-2	-1	0		
A	23	54	37	*2*	*11*	*3*	5	29	N_2	0
T	33	19	46	*93*	*82*	*2*	30	29	N_3	97
G	17	17	13	*2*	*2*	*2*	8	20	N_4	3
C	26	11	5	*2*	*4*	*93*	57	22		
STAT6	-7	-6	-5	-4	-3	-2	-1	-0		
A	23	51	33	*3*	*5*	*11*	18	35	N_2	0
T	42	25	44	*92*	*89*	*3*	14	25	N_3	7.5
G	13	21	19	*2*	*2*	*4*	14	23	N_4	92.5
C	22	3	4	*3*	*3*	*82*	53	18		

[a] Position 0, or ±0 for STAT6, is the center of the palindromic binding site that can be described as $TTCN_xGAA$ ($2 \leq x \leq 4$). The palindromic nucleotides are shaded.
[b] See Footnote a.

the size available to us, we chose the latter approach. The initial model used as common starting point is shown in Fig. 1. We call it a hidden Markov model (HMM), because we use an HMM training algorithm for parameter estimation. Note that the initial model contains some assumptions about a general STAT-binding site as indicated above but leaves a great degree of flexibility. We seeded the model only with a week preference for N_3 spacers (50%) because we were concerned that a stronger preference could mislead the refinement process in the case of STAT6 that is expected to prefer N_4 spacer sites. We chose the same initial preference for N_2 spacers as for N_4 spacers, primarily to control the method.

For model refinement we used the Baum-Welch HMM training algorithm. This method, in conjunction with the specific model architecture chosen, is equivalent to the Expectation Maximization algorithm introduced by Cardon and Stormo (22) for the analysis *E. coli* promoters. The method was chosen because promoter recognition by *Escherichia coli* RNA polymerase can be viewed as a prototype of a protein-DNA interaction governed by a sequence motif consisting of two conserved blocks separated by a spacer of variable length. In simplified terms, the Baum-Welch algorithm is an iterative alignment procedure that works as follows. The initial model is used as a guide to identify and align putative binding sites during the first iteration. The base and spacer length frequencies of the resulting multiple alignment define a new model that guides the alignment process during the second iteration. One expects that each iteration results in a model that gives a better fit to the training sequences. With sufficiently large training sets, such an algorithm has a remarkable capability of overcoming errors in the initial model. In our work this is exemplified by the increase of the frequency of the N_4 spacer class from 25 to 93% during the training with STAT6-binding oligonucleotides. The simplified algorithm described above corresponds to the "Viterbi training" variant of the Baum-Welch algorithm which is more sophisticated in its original form (for details see Ref. 23).

Comparison of STAT1, STAT5, and STAT6 High Affinity Binding Site Models—Like all other STATs, STAT5A shows a very strong preference for sequences containing two palindromic half-sites TTC . . . GAA (24–26) (Fig. 2 and numerical values in Table I). We use the position notation proposed by Chen *et al.* (21) and Becker *et al.* (27), *i.e.* position −2 corresponds to the C in TTC and +2 to the G in GAA. Unless otherwise specified, we only refer to one half-site, implying that the complementary is true for the other half-site. With regard to the central and the flanking nucleotides, the matrix derived from the STAT5-selected sites resembles very much the STAT6

matrix, whereas the STAT1 matrix is more loosely related. Two differences are most striking. (i) Almost all STAT1-selected sites have a C in position −1, whereas about 40% of STAT5A- and STAT6-selected sites contain other nucleotides at this position. Indeed, for STAT1 the requirement for a C at this position appears to be even more stringent than that for C at position −2. (ii) STAT1 shows a strong preference for G or C at position 0, *i.e.* the center of the palindrome, whereas neither STAT5A nor STAT6 strongly select specific nucleotides at this position. Another interesting difference concerns nucleotide −7. STAT1 has a preference for C at this position (57%), which translates into a preference for G at position +7. This is in line with the crystal structure of STAT1 bound to DNA that shows a contact of glutamate 421 with a guanosine at position +7 and the suggestion that this contact would be impossible with a thymidine at position +7 (21). It would be interesting to know whether this contact could be observed in crystals of STAT5A or STAT6 that show no preference for C at this position. Note further that at position −5 C is strongly avoided by both STAT5A and STAT6. STAT6, and to a lesser extent STAT5A, also select against C at position −6. These results are in agreement with the recently published data of Soldaini *et al.* (28).

Some of the variation between the matrices may be due to differences between the conditions of the STAT5A and STAT6 selection experiments on the one hand and the STAT1 selection on the other hand. However, the results of the competition experiments described below reveal the same preference patterns, which strongly suggest that they are not an experimental artifact.

The hidden Markov modeling confirms the strong preference of STAT6 for sites in which there are four nucleotides between the two halves of the palindrome (TTCN$_4$GAA), as is shown in Fig. 2 and Table I. Nevertheless, STAT6 selected 7% N_3 sites (TTCN$_3$GAA), and below we provide specific evidence that STAT6 can indeed bind to N_3 sites. Like STAT1, STAT5A seems to select preferentially N_3 sites, but the preference for such sites appears less absolute. Unlike STAT5A or STAT6, STAT1 appears also to select some N_2 sites. Among the oligonucleotides selected *in vitro* by STAT1 (see Supplemental Material), we found at least one candidate for such a binding mode (CATCAC TTCCGGAA ATGGCGT).

Comparison of Low Affinity Binding Sites—We wanted to investigate whether the differences between the sequence specificities of the various STAT proteins revealed by the comparison of the selected high affinity binding sites also apply to low affinity binding sites. For this purpose we tested the effect of base changes on the affinities for STAT1, STAT5A and -B, and STAT6. We used recombinant STAT proteins to avoid effects of unknown cellular components on DNA binding. To determine relative affinities, we carried out competition assays in which graded amounts of accurately quantified double-stranded synthetic oligonucleotides were added as competitors to a standard binding reaction between a STAT protein and a radioactive reference probe (for an example see Fig. 3*B*). For the latter we used the GRR element of the *Fc-γRI* gene, as it binds strongly to STAT1, -5, and -6 (12, 29, 30). Binding reactions were carried out under conditions that had been optimized in preliminary experiments (see "Experimental Procedures"). They were fractionated on nondenaturing polyacrylamide gels. Autoradiographs of the gels were used to determine the quantities of test and reference competitor that reduced the amount of bound labeled probe to the same level (see Fig. 3*B*). For example the affinity of STAT1 for sequence II, spacer N3 was calculated in the following way (the roman number in front of the sequence name designates the class of the element in Table II for better orientation). By using 5880 fmol of spacer N3 sequence and 35 fmol of unlabeled GRR sequence, the shifted bands had the same intensity. The adjacent bands served as control. We divided the amount of test competitor by the amount of the reference competitor (5580 fmol/35 fmol = 168), and we rounded the final result to 170 times reduction of affinity compared with the affinity of the reference probe. In some instances, gels were also scanned with a PhosphorImager to verify the accuracy of the results obtained by visual comparison and to ascertain that the amount of probe added to each reaction and loaded on the gel was constant. The complete results obtained with 40 sequences are reported in Table II, where they are expressed as the fold reduction of affinity compared with that for the reference probe.

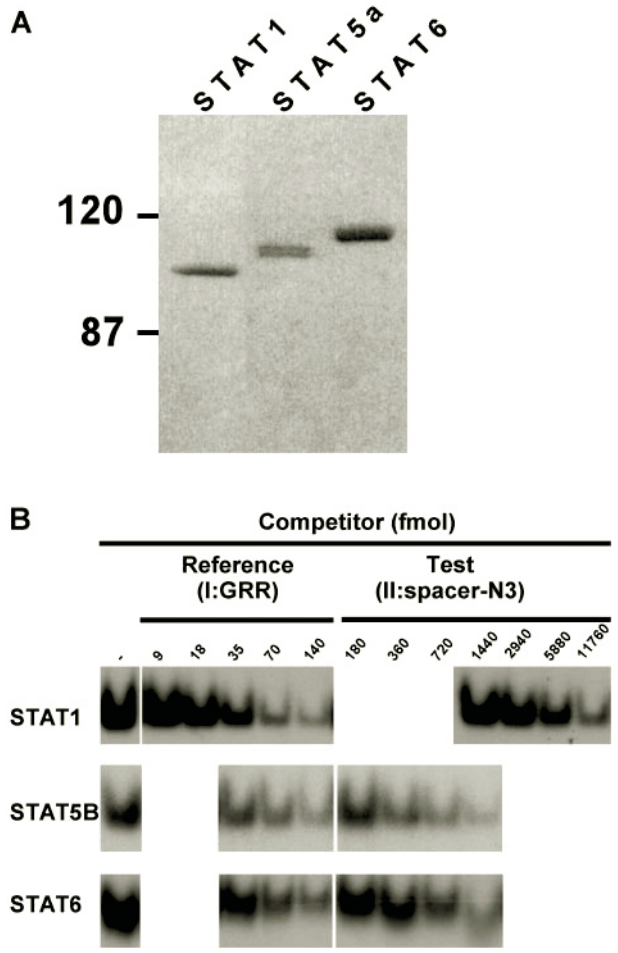

FIG. 3. *A,* characterization of recombinant STAT. Histidine-tagged proteins were produced in insect cells as described previously. Cells were coinfected with two baculovirus samples, one encoding JAK2 tyrosine kinase and the other a STAT protein. Tagged proteins were purified by Ni^{2+} affinity and analyzed by SDS-polyacrylamide gel electrophoresis. 500 ng of each protein were loaded on the gel and stained with Coomassie Brilliant Blue. *Numbers* on the side refer to molecular weight standards. We do not know why STAT5A migrates as a doublet; the *upper band* does not represent the phosphorylated portion. *B,* experimental procedure for the determination of relative affinities of STAT proteins. Different amounts of carefully quantified reference or test competitor were added to bandshift reactions containing a constant amount of STAT protein and labeled reference probe (see "Experimental Procedures"). Reactions loaded onto the same gel were compared by assessment of the resulting autoradiographs. This allows one to determine with an accuracy of less than 2-fold the amount of test competitor required to reduce the signal due to the shifted probe to the same level as a given quantity of reference competitor.

Effect of Changing Spacer Length between Palindromic Half-sites—To investigate the role of the spacer of the half-palindromes in the binding of different STATs to low affinity binding sites, we added a base to the center of the N_3 spaced palin-

TABLE II

Comparison of the affinities of different DNA sequences for STAT1, STAT5A, STAT5B and STAT6

Lowercase letters in the XI, SIE element indicate single-stranded overhangs in the annealed oligo used for competition. The following abbreviations are used: GRR, IFN-γ response region of the *Fc-γRI* gene (19, 41); mIL-2rE site I/II, distal/proximal GAS element of mouse *IL-2Rα* gene (42); hIL-2rE site I/II, distal/proximal GAS element of human *IL-2Rα* gene (43, 44); APRE, acute phase response element of rat α_2-*macroglobulin* gene (45); ϵ, GLS of human *Ig heavy chain germ line* gene (46); IL-4R, STAT6-binding site of human *IL-4R* gene (30); SIE, sis-inducible element of human c-*fos* gene (47); N3-selected, sequence selected in random selection experiment with STAT5A; N4 artificial 1–3, artificial N4 sequence 1–3; N4-selected, sequence selected in random selection experiment with STAT6; N3 artificial 1, artificial N3 sequence 1; half-site 1–5, artificial half-sites. The sequences designed to test effect of specific base changes and spacer length are as follows: spacer N3, reference sequence (N_3 spaced palindrome); spacer cttc, A → C change at position −5; spacer ttcc, G → C change at position −1; spacer N4, conversion to N_4 site by insertion of a base (see text); spacer N3/N4, A → T change at position 4 (see text); spacer half, mutation in one side of the palindrome to create a half-site (see text); no STAT1–2, sequences without STAT-binding element nor half-sites; suffix var-1, -2. . . denominates sequence variations of the original sites.

Element [b]	Sequence	Fold reduction in relative affinity [a]			
		STAT1	STAT5A	STAT5B	STAT6
class I	*TATTTCCCAGAAARS*				
GRR	GTATTTCCCAGAAAAGGAAC	1	1	1	1
GRR-var1	GTATTTCCCAGAAAAGCTAC	1	1	1	1
GRR-var2	GTTATTTCCCAGAAAGGCCAGACATGGAGTCGCCGAGAG	3	3	ND	6
GRR-var3	GTATTTCCCAGAAAGGGAAC	2	3	ND	5
GRR-var4	GTTATTTCCCAGAAAGCCCAGACATTTCTGATAAGAGAG	5	6	ND	6
class II	*AAMTTCSCSGAAKWT*				
spacer-N3	GACAAATTCGCCGAATTTGG	170	5	10	10
spacer-cttc	GACAACTTCGCCGAATTTGG	140	8	17	20
spacer-ttcc	GACAAATTCCCCGAATTTGG	5	3	5	10
N3-selected	GATCTGTTAACTTCCCGGAAGATG	3	5	5	40
class III	*TYCTTCTRRGAAKTM*				
hIL-2rE site I	CAGTTTCTTCTAGGAAGTACCA	60	10	10	7
mIL-2rE site I	CAGTTTCTTCTGAGAAGTACCA	90	10	ND	5
APRE	GATCCTTCTGGGAATTCCTA	5	2	2	2
class IV	*CATTTCTGRTAAKAG*				
hIL-2rE site II	CAGACATTTCTGATAATAGAG	60	70	70	70
mIL-2rE site II	CAGACATTTCTGATAAGAGAG	50	100	180	100
mIL-2rE site II-var1	CAGACATTTCTGGTAAGAGAGTTG	20	100	100	100
class V	*GACTTCCCAAGAACAG*				
ε	GATCCTCGACTTCCCAAGAACAGCA	>200	20	80	2
ε-var1	GATCCCGCTGTTGCTCAATCGACTTCCCAAGAACAGCA	>200	20	40	2
ε-var2	GATCCCGCTGCGTCGAGGCCGACTTCCCAAGAACAGCA	>200	25	100	5
class VI	*WRRTTCGCGCGAAYYW*				
N4-artificial1	TGCATGATATGATGTGCAGTTGGTTCGCGCGAACCACCAGACA	>180	>200	>100	10
spacer-N4	GACAAATTCGCGCGAATTTG	>200	60	130	3
class VII	*WKYTTCYBRAGAASYW*				
mIL-2rE site I-var1	CAGTTTCTTCCTGAGAAGTACCAGAC	>200	80	>100	3
mIL-2rE site I-var2	GCAGTTTCTTCTCGAGAAGTACCA	>200	>200	>150	10
N4-artificial2	GATCCATGGGAGTTTCTGAAGAACCTTCAGCTGTGCACAG	>200	80	>200	5
class VIII	*TWMTTCAYMYGAAWTA*				
IL-4R	TGATATGATGTGCAGTTTCTTCATCTGAAATACCAGACA	>200	200	>100	15
N4-selected	CATAATTCACACGAATTACAGATC	>200	30	40	3
class IX	*CRKTTCBYGMTAASMG*				
mIL-2rE site II-var2	GAAGTACCAGACATTTCCTGATAAGAGAG	>200	110	>100	7
mIL-2rE site II-var3	CAGACATTTCTCGATAAGAGAGTTG	>200	>200	>200	20
N4-artificial3	ACCAGACGGTTCGCGCTAACCGAGTTGAGCAACTTCCTGATATG	>200	>200	>200	25
class X	*CGTTCACTGTAGTC*				
N3-artificial1	CGTTCACTGTAGTCTGGCTGC	>200	>200	ND	>200
class XI	*RWMTTCBSGBSWWKK*				
SIE	tccGTATTCCCGTCAATGCA	10	>200	>200	>200
spacer-N3/N4	GACAAATTCGCGCGATTTTG	>200	130	>200	25
N4-artificial4	CGCTGCTGCCCCCCAGCATCTTCTGGGGTAGGAAGTGGTTG	>200	>200	ND	>200
class XII	*RHTTCMYSDRSNB*				
spacer-half	GACAAATTCCCCAGCTTTGG	>200	>200	>200	>200
half-site1	ATAGAGTTGAGCAACTTCCTGA	>200	>200	>200	>200
half-site2	CAAAGAGGGATTTCACCTACAT	120	>200	ND	>200
half-site3	CAAAGACTCATTTCACCTACAT	>200	>200	ND	>200
half-site4	GAAAGGGACCAACCACTTCCTGGGGGCAGCAGCGCCACGT	>200	>200	ND	>200
half-site5	AAGGAGAGGGAGATTCCCCTGCCGTTG	>200	>200	>200	45
class XIII	*RRYYWKGSRYKGRCWDKRYKSK*				
noSTAT1	GATCATGCATTGACATCATTGT	>170	>160	ND	>160
noSTAT2	AGCTTGGGGCGGGCTGGGCGCGGGCG	>200	>200	ND	>200

[a] The relative affinity of a STAT protein for different oligos was determined by comparing their capacity to compete for STAT binding in a band shift assay (see Fig. 3*B*). For normalization the results were compared to the competition with a reference, the IFN-γ response element of the Fc-γRI gene (GRR). The numbers indicate the ratio between the concentration of the competitor and of the reference oligo that reduce band shift intensity to the same degree.

[b] Sequences are classified into 13 classes according to sequence similarity, number of nucleotides between half-sites, substitutions within half-sites, and bases between half-sites or outside the palindrome.

drome of the distal STAT consensus site in the IL-2-responsive enhancer of the murine *IL-2Rα* gene (III, mIL-2rE site I; . . . GTTTC TTCTGAGAA GTACCA; half-sites of core palindrome are underlined) (see Table II for complete results).

Insertion of a C at position −1 (VII: mIL-2rE site I-var1: . . . GTTTC TTC*C*TGAGAA GTACC . . . ; inserted bases are in italics) resulted in an 8-fold reduction of STAT5A binding, whereas insertion of a C at position −0 (VII: mIL-2rE site I-var2; . . . GTTTC TTCTC*C*GAGAA GTACC . . .) reduced binding more than 20 times. This difference is presumably due to the fact that insertion of a C at position −1 gives rise to a half-site TTC*C*, characteristic for high affinity binding sites. Surprisingly, neither insertion significantly improved STAT6 affinity. In III, mIL-2rE site I the N_3 site overlaps with an N_4

site (CAGT TTCTTCTGAG AAGTACCA) which has a mismatch in one of the palindromic cores (dot below letter) but lacks the unfavorable C and G residues at position −5 and +5, respectively. It is possible that STAT6 binds to this element that is destroyed by the insertions in the mIL-2rE site I variants.

It is an inherent characteristic of N_4 sites that they contain N_3 spaced palindromes with more or less mismatches. Thus the VII mIL-2rE site I-var1 sequence contains an N_3 site with one mismatch (CAGTTTC TTCCTGAGAA GTAC . . .) in addition to the N_4 site. To test whether such "hidden" sites affect STAT binding, we designed a pair of oligonucleotide sequences with low predicted affinity that contain an N_3 (II: spacer N3, GACAAA TTCGCCGAA TTTGG) or an N_4 site (VI: spacer N4, GACAAA TTCGCGCGAA TTTG) but no hidden sites with a single mismatch.

Both STAT5A and STAT5B bound about 12 times better to the spacer N3 than to the spacer N4 sequence. STAT6 favored the spacer N4 site, but binding was only 3 times stronger than to the spacer N3 site. The affinity of STAT1 for the spacer N3 sequence was already very low, and we could not detect significant binding to the N_4 sites.

These data suggest that STAT5 can bind, albeit weakly, to N_4 spaced sites with two conserved TTC half-palindromes. However, as any such site also contains an N_3 spaced site with two mismatches in one of the half-sites, we considered the possibility that STAT5 does not bind to the N_4 half-site but rather to the N_3-spaced half-site with two mismatches. To test this, we introduced a mismatch into one of the conserved N_4-spaced half-sites of sequence spacer N4 that left the N_3-spaced half-sites with two mismatches unchanged (XI: spacer N3/N4; GACAAA TTCGCGCGAT TTTG). This resulted in a significant reduction of binding to both STAT5A and STAT5B. Inevitably, the change also affects the flanking nucleotide of the mismatched N_3 half-site. But the binding site selection experiments had revealed that sequences with A or T at this position (+5) were selected with equal frequencies. Thus, we consider it unlikely that STAT5 recognizes sequence spacer N4 as an N_3 site with two mismatches in one half-site, and we conclude that STAT5 can bind N_4 sites.

Of the 16 N_4 spaced sequences tested (see Table II), STAT1 bound only one with detectable affinity. This sequence, the SIE element of the c-*fos* promoter, contains one perfect half-site. In addition it contains a half-site with one mismatch (XI: SIE, tccGTA TTCCCGTCAA TGCA) at a distance of 4 nucleotides, and a half-site with two mismatches (tccGTA TTCCCGTCA ATGCA) three nucleotides from the conserved half-site. In a previous analysis (31) it was found that an A → C mutation that induces a second mismatch in the N_4 spaced palindrome (tccGTA TTCCCGTCAC TGCA) does not change the affinity of the sequence for STAT1. On the other hand, a mutation of C → A that creates an N_3 site with one instead of two mismatches (tccGTA TTCCCGTAA ATGCA) increases the affinity. These data suggest that STAT1 cannot bind to N_4 spaced half-sites and recognizes the SIE element as an N_3 site, even though it contains two mismatches.

Comparison with Natural Binding Sites in STAT-responsive Genes—We compared these spacer preferences with the configuration of the natural sites described for STAT1, -5, and -6. Genes reported to be activated by STAT3 and -4 were included for completeness. The current literature was screened, and enhancer elements claimed to be targets of STAT1, -3–6 listed in Table III. The conditions for inclusion in the table are outlined there. Table III also indicates whether data derived from gene targeting, with STAT-deficient cell lines, dominant negative STAT mutants, antisense RNA, or electroporation of specific antibodies support a role for a STAT protein. In the following the genes containing the enhancers compiled in this fashion are referred to as STAT1- and STAT3–6-activated genes, respectively. Note that in this analysis, we relied on the published interpretations of the experiments. The sequences in Table III were thus aligned as shown in the original publications. Reanalysis of these sites by the same HMM-based protocol as used for the *in vitro* selected oligonucleotides was not possible because for many sites we were lacking the experimental details delimiting the sequence range shown to interact with the STAT protein.

All genes found to be activated STAT4 and STAT5 contained a palindrome spaced by N_3 (see Fig. 4 for a graphic representation of the base frequency matrix derived from the STAT-activated genes, the numerical values are presented in Table IV). STAT1-activated genes also contained almost exclusively N_3 sites, with the exception of the 3′-interferon response region element of the *HLA-E* gene, which is a half-site. Most of the STAT6 targets were sites with N_4 spacers, but some STAT6-responsive enhancers did contain N_3 sites. It was reported that one of these sites conferred only a weak response to STAT6 (32). Among the STAT3 targets, one canonical site with an N_2 spacer could be identified (CRP-APRE element of C-reactive protein), and there were 2 sites with a N_4 spacer. We will address the question if STAT1 can bind to sites with a N_4 spacer below.

STATs Do Not Bind to Half-sites in Vitro—In relation to these experiments, the question arose whether STATs can bind with detectable affinity to isolated half-sites. Since the binding site selection experiments had indicated a strong preference for C at position −1 for all proteins tested, and STAT1 had, in addition, selected C or G at position 0, we reasoned that TTCCC would be an optimal half-site. We designed an oligonucleotide with this motif, but without a second half-site (XII: spacer-half, GACAAA TTCCCCAGCTTTGG). This oligonucleotide did not compete for binding to any of the four STAT proteins analyzed. Thus, it appears that these STATs cannot bind with significant affinity to isolated half-sites.

STAT1 Is More Permissive for Mismatches in Position ±2 of the Palindrome Than STAT5 and STAT6—15 of the 26 elements of "natural" STAT1-activated genes contained mismatches in the palindromic sequence. All besides one sequence (3′-interferon response region element of *HLA-E* gene) showed only a substitution of a single base per palindromic half-site. In 10 cases the substitution was in one half-site, and in the 4 other cases both palindromic half-sites were mutated (cis-acting regions of genes *CD86* (both elements) and *GBP-1*, distal site of *MIG*). The most frequent substitution was TTC → TTA (see Table III). In contrast to STAT1, the 11 cis-acting elements of STAT6-activated genes all contained perfect palindromes besides one (mCD23a STAT6 site). For STAT5 the situation was more complex; all 22 STAT5-activated genes contained one cis-acting element with a perfect palindrome, but in three of the six genes regulated by more than one binding site for STAT5, one of the elements contains a TAA half-site (instead of GAA). These differences suggest that STAT1 is more permissive for mismatches *in vivo*. This is consistent with our *in vitro* measurements of relative affinity, as illustrated by the following. Sequence IV, mIL-2rE site II (. . . ACAT TTCTGATAA GAGAG) contains a TAA half-site and binds at least as well to STAT1 as the III, mIL-2rE site I (. . . TTTC TTCTGAGAA GTACCA) with a perfect palindrome, whereas the latter sequence binds 10 times better to STAT5A. Similar results were obtained with two more sequences with a TAA half-site (III, hIL-2rE site I (. . . TTCTAGGAA . . .) *versus* IV, hIL-2rE site II (. . . TTCTGATAA . . .); III, APRE (. . . TTCTGGGAA . . .)

TABLE III

Cis-acting elements of STAT-responsive genes

Elements were included in this table if the experimental evidence demonstrates the following: (i) mutations in the potential binding site in a homologous context reduce the induction of an appropriate reporter (the homologous sequence must have a minimal length of 100 base pairs; if the mutation is done by deletion the deletion must be smaller than 50 base pairs; monomers without mutations are not considered); (ii) the binding protein is identified by immunological methods or by purification of bound protein; (iii) the results from gene targeting experiments do not speak against the involvement of a particular STAT protein. Putative interactions between a site and a STAT protein that lead to inhibition of induction of the gene in *cis* are not considered. In view of the quantity of STAT literature, we apologize to authors whose work could not be included in the table due to space constraints. IVFP indicates references that describe *in vivo* footprinting. ND indicates not done. * indicates references to experiments with mice.

Gene/enhancer	Species [a]	Response to (in studies indicated)	Putative cis-acting element			Experimental evidence		
			Name	Sequence [b]	Position [c]	Identification of STAT [d]	Functional assay [e]	Effect of STAT deficiency [f]
STAT1:								
α1-antichymotrypsin	Hs	OSM	ACT-A	CGTAIIACCAGAAATTA	-129	(48)	(48)	ND
α1-antichymotrypsin	Hs	OSM	ACT-B	TCCAGICCGAGAACAGA	-99	(48)	(48)	ND
α2-macroglobulin	Rn	IFN-γ, IL-6	APRE	ATCCIICTGGGAATTCC	-170	(45,49,50)	(45)	ND
bcl-x	Hs	LIF	GAS	GCATTICGGAGAAGACG	-46	(51)	(51)	(52,53)
caprine arthritis-encephalitis virus LTR	viral	IFN-γ	GAS	CAAAIICCTGTAAATCA	-175	(54)	(54)	ND
CD40	Hs	IFN-γ	mGAS	ACTCICCTTGAAACGC	-535	(55)	(55)	(55)
CD86	Hs	IFN-γ	GAS3	CATTIIGGTCTAAACTA	-1176	(56)	(56)	ND
CD86	Hs	IFN-γ	GAS4	TAACIIGCTTTAAAGCT	-1126	(56)	(56)	ND
CIITA	Hs	IFN-γ	GAS	CCACIICTGATAAAGCA	-146	(57,58)	(57,58)	(59)*(60,61)
Fc-γRI	Hs	IFN-γ, IL-10, LIF	GRR	GTATIICCCAGAAAAGG	-34	(19,41)	(19,29)	ND
GBP-1	Hs	IFN-γ	GAS	CATAIIACTCTAAATCC	-119	(62)	(24)	(59)*
gp130	Hs	IL-6/OSM	SBE	CGCGIIACGGGAATCGC	-232	(63)	(63)	ND
HLA-E	Hs	IFN-γ	5'IRR	TGATIGCTGGGAAACTC	-191	(64)	(64)	ND
HLA-E	Hs	IFN-γ	3'IRR	CAGTIICCCGTTCCTCT	-171	(64)	(64)	ND
ICAM-1	Hs	IFN-γ, IL-6	pIγRE	AGGTIICGGGAAAGCA	-79	(65-69)	(65-67,69)	(70)
ICAM-1	Hs	IFN-γ	GAS-2787	CCTTIICTTGAAAAAC	-2792	(71)	(71)	ND
INDO	Hs	IFN-γ	PEII	CATTIICCTGTAAAATG	-414	(72), (73)	(72), (73)	ND
iNOS	Mm	IFN-γ	GAS	CCTTIICCCCTAACACT	-946	(74)	(74)	(59)
IRF-1	Rn	IFN-γ, Prl, GH	GAS	TGATIICCCCGAAATGA	-126	(75-77)	(78)	(59)*(79)*
IRF-1	Hs	IFN-γ, IFN-α, Prl	GAS	TGATIICCCCGAATGAC	-128	(80)	(80-82) [IVFP]	(59)*(79)*
MCP-1	Hs	IFN-γ	GAS	CTACIICCTGGAAATCC	-216	(83)	(83)	ND
MIG	Mm	IFN-γ	prox. GAS	CCGTIIATGTGAAATGG	-184	(84)	(85)	(86)
MIG	Mm	IFN-γ	dist. GAS	TCCCIIACTATAAACTC	-202	(84)	(85)	(86)
SOCS-3	Mm	LIF	STAT1/STAT3 element	GCAGIICCAGGAATCGG	-76	(87)	(87)	ND
SPI-3	Rn	IFN-γ	SPI STAT1/STAT3	CATGIICCCAGAAATCA	-136	(88,89)	(88)	ND
TAP-1	Hs	IFN-γ	GAS	CGCTIICCCCTAAATGG	-193	(90,91)	(90)	ND

[a] Hs, human; Mm, mouse; Rn, rat (*Rattus norvegicus*); Ma, Syrian hamster (*Mesocricetus auratus*); Oc, rabbit (*Oryctolagus cuniculus*).

[b] Orientation is 5′ to 3′ towards transcription start. Palindromic core half-sites are shaded and nucleotides underlined with an unbroken line when they correspond to the optimal sequence and a dashed underline if the nucleotide is a mismatch to the canonical palindromic sequence.

[c] Position of the first nucleotide of sequence shown in this table.

[d] By *in vitro* (electrophoretic mobility shift assay + supershift or purification of bound protein) or by *in vivo* footprinting.

[e] By mutagenesis of binding site in homologous context.

[f] By gene targeting, STAT-deficient cell lines, dominant negative protein, antisense RNA, or electroporation of specific antibodies.

versus IV, mIL-2rE site II-var1 (. . . TTCTGGTAA . . .), in the last case we noted a detectable increase in affinity for STAT1, but the increase in affinity for STAT5 was more than 10 times stronger). These observations suggest that STAT1, but neither STAT5 nor STAT6, interacts with TAA half-sites almost as well as with the canonical GAA half-site. The fact that this toler-

TABLE III—*continued*

STAT3:							
α 1-antichymotrypsin	Hs	OSM	ACT-A	CGTATTACCCAGAAATTA	-129	(48)	(48)	ND
α 1-antichymotrypsin	Hs	OSM	ACT-B	TCCAGTCCGAGAACAGA	-99	(48)	(48)	ND
α 2-macroglobulin	Rn	IL-6, IFN-γ	APRE	ATCCTTCTGGGAATTCT	-170	(49,50,92-94)	(45)	ND
angiotensinogen	Rn	CT-1, AngII	St-domain	AGGGTTCCTGGAAGGGA	-178	(95,96)	(95,96)	ND
c-fos	Hs	PDGF, EGF	SIE	GCAGTTCCCGTCAATCC	-352	(31,97)	(31,98)	(99)
c-myc	Hs	IL-6	STAT3 binding site	AGGCTTGGCGGGAAAAAG	+91	(100)	(100)	(100)
C/EBP δ	Mm	growth arrest, IL-6	δAPRE	GTCGTTCCCAGCAGCAC	-115	(101,102)	(101)	(101)
C/EBP δ	Rn	IL-6	δ IV site	GTCGTTCCCAGCAGCAC	-115	(103)	(103)	ND
C-reactive protein	Hs	IL-6	CRP-APRE	CCTCTTCCCCGAAGCTCT	-116	(104)	(104)	ND
CYP19 gene = aromatase	Hs	IL-11	GAS	GTGTTTCCTGTGAAAGTT	-286	(105)	(105)	ND
gamma-fibrinogen	Rn	IL-6	site II	CAAAATTCTGGGAATCCC	-158	(106)	(106)	ND
gamma-fibrinogen	Rn	IL-6	site III	CCCAGACTGGGAATTCA	-49	(106)	(106)	ND
gp130	Hs	IL-6, OSM	SBE	CGCGTTACGGGAATCGC	-232	(63)	(63)	ND
junB	Mm	IL-6	JRE-IL6	GCGCTTCCTGACAGTGA	-151	(107,108)	(107,109)	(100)
junB	Mm	IL-6	APRF element	CAGATTCCGGGAATCCC	+2135	(110)	(110)	ND
LBP	Hs	IL-6	STAT3 RE	ACTGCACTGGGAATCTA	+91	(111)	(111)	ND
Ly6-A/E	Mm	IL-6, IL-9	δ GAS	CATATTCCTGTAAGTGA	-1232	(112)	(112)	ND
p21WAF	Hs	OSM	P21-SIE1	CTCCTTCCCGGAAGCAT	-637	(113)	(113)	(113)
perforin	Hs	IL-2	PSTAT	CAGATTCCGAGAAGACA	-1093	(114)	(114)	ND
SOCS-3	Mm	LIF	STAT1/STAT3 element	GCAGTTCCAGGAATCGG	-76	(87)	(87)	(87)
SPI-3	Rn	IL-6	APRF/STAT3 site	CATGTTCCCAGAAATCA	-136	(88,89)	(88)	ND
STAT3	Mm	IL-6	SBE	GCACTTGCCAGGAACTCA	-326	(115)	(115)	(115)
TIMP-1	Rn	OSM, IL-6	STAT RE	ATGCGTCCAGGAAGCCT	-53	(116)	(116)	ND
VIP	Hs	LIF	CyRE	AGATTTCCCTGGAATTAA	-1318	(117)	(117)	ND
STAT4:							
c-myc	Hs	IL-2	GAS-like RE	GCATTTCCAATAATAAA	-1405	(118)	(118)	ND
IFN-gamma	Hs	IL-12	STAT RE	TGCCTTCCAAAGAATCCC	-370	(119) IVFP	(119)	ND
IRF-1	Hs	IL-12	SBE	TGATTTCCCCCGAAATGA	-266	(120)	(120)	ND
STAT5:							
3β-hydroxysteroid dehydrogenase	Hs	Prl	STAT5 RE	ATTATTCTGAGAAAAGG	-122	(121)	(121)	ND
α s1 casein	Oc	Prl	F4	TACTTTCTTAGAAAATT	-3331	(122)	(122)	ND
ALS	Mm	GH	ALS-GAS1	GGTGTTCCTAGAAGAGG	-637	(123)	(123)	ND
Bcl-XL	Hs	IL-3	STAT binding element	GCATTTCGGGAGAAGACG	-335	(124,125)	(124)	(53,124,125)
Bcl-XL	Mm	Epo, IL-2, GM-CSF, IL-3	STAT binding element	GCATTTCGGGAGAAAAGG	-345	(126)	(126)	(125,127,128)
β casein	Rn	Prl	prox. MGF site	GGACTTCTTGGAATTAA	-101	(129)	(130)	ND
CIS	Hs	IL-2	CIS1	GCGGTTCTAGGAAGACG	-257	(131)	(131)	(132)*
CIS	Hs	IL-2	CIS2	CTGCTTCCGGGAAGGGC	-240	(131)	(131)	(132)*
CIS	Hs	IL-2	CIS3	CGTTTTCCTGGAAAGTT	-369	(131)	(131)	(132)*
CIS	Hs	IL-2	CIS4	AAAGTTCTTGGAAATCT	-358	(131)	(131)	(132)*

TABLE III—*continued*

cyclin D1	Hs	IL-3, constit. active STAT5	D1-SIE1	GGCGTTCTTGGAAATGC	-644	(133)	(133)	(133)
CYP2C12	Rn	GH	GHRE1	AAATTTCCTAGAAGTGA	-4200	(134)	(134)	ND
CYP2C12	Rn	GH	GHRE2	TAAATTCCTAGAACTCA	-4175	(134)	(134)	ND
CYP 3A 10/6 β hydroxylase	Ma	GH	CYP GHRE	CAAGTTCCTGGAAGCGT	-95	(135)	(135)	(136)
HNF-6	Rn	GH	H6 STAT	AGGATTCTAAGAAGAGG	-120	(137)	(137)	(137)
insulin 1	Rn	GH	ins-GLE	AACTTTCTGGGAAATGA	-334	(138)	(138)	ND
IL-2R α	Mm	IL-2	mGASI	TTTCTTCTGAGAAGTAC	-1369	(6,12)^{IVFP}	(42)	(139)
IL-2R α	Mm	IL-2	mGASII	ACATTTCTGATAAGAGA	-1349	(6,12)^{IVFP}	(42)	(139)
IL-2R α	Hs	IL-2	GASd	TTTCTTCTAGGAAGTAC	-4285	(43,44)^{IVFP}	(43,44)	(139)*
IL-2R α	Hs	IL-2	GASp	ACATTTCTGATAAGAGA	-4265	(43,44)^{IVFP}	(44)	(139)*
immunoglobulin J chain	Mm	IL-2	GAS	GCAATTCTATGAAAAGC	-549	(140)^{IVFP}	(140)	ND
lymphotoxin α	Mm	IL-2	GASL1	CCCTTTCCCAGAACGCA	-193	(141)	(141)	(141)
OSM	Mm	Epo, IL-3	prox. GAS	CATGTTCCCAGAAGGCC	-158	ND	(142)	(132,143)
OSM	Hs	GM-CSF	M-STAT	CGAATTCGAAGAAAACA	-195	(144)	(144)	ND
perforin	Hs	IL-2	pSTAT	CAGATTCCGAGAAGACA	-1093	(114)	(114)	(145)*
prolactin receptor (exon 1A promoter)	Rn	GH, Prl	GLE	TGAGTTCTAGGAATAAA	-160	(146)	(146)	ND
spi 2.1	Rn	GH	SPI-GLE1	CATGTTCTGAGAAATCA	-133	(147)	(148,149)	ND
spi 2.1	Rn	GH	SPI-GLE2	ACGCTTCTACTAATCCA	-148	(149)	(149)	ND
TCRγ (Jγ1 region)	Mm	IL-7	STAT binding motif 1	CCTCTTCTCAGAAATAT	-229	(150)	(150)	ND
TCRγ (Jγ1 region)	Mm	IL-7	STAT binding motif 2	TGAGTTCCTGGAAATTA	-124	(150)	(150)	ND
TCRγ (Jγ1 region)	Mm	IL-7	STAT binding motif 3	AAAATTCCTAGAATACT	-107	(150)	(150)	ND
STAT6:								
12/15 lipoxygenase	Hs	IL-4	ST6 TLO	GACTTTCCTGAGAAACCG	-957	(151)	(151)	ND
angiotensinogen	Rn	AngII	St-domain	AGGGTTCCTGGAAGGGA	-178	(96)	(96)	ND
c-γ 1 sterile transcript	Mm	IL-4	IL-4RE	CACATTCACATGAAGTAA	-127	(152)	(153)	ND
c-γ 3 sterile transcript	Hs	IL-4	GAS	TGATTTCCTAGGAAGAC A	-89	(154,155)	(154), (155)	ND
eotaxin	Hs	IL-4	STAT6 element	AGGCTTCCCCTGGAATCTC	-78	(156)	(156)	ND
Fc-εRII/CD23 (CD23a promoter)	Mm	IL-4	m CD23a STAT6 site	TGCTTACCCTGAGAAATAA	-117	(157)	(158)	(159)
Fc-εRII/CD23 (CD23b promoter)	Hs	IL-4	IL-4RE	GAATTTCTAAGAAAGGG	-230	(160)	(161)	(159)*
Ig heavy chain germ line gene	Hs	IL-4	GLS	CGACTTCCCCAAGAACAGA	-154	(162)	(46, 163)	(164-166)*(167)
lymphotoxin α	Hs	IL-4	STAT site	CCCTTTCCCAGAACTCA	-201	(32)	(32)	ND
P-selectin	Hs	IL-4	site 1	TAACTTCATGGGAAGGGC	-153	(168)	(168)	ND
sIL-1Ra	Hs	IL-4	SBE1	CTTCTTCCCAGGAACTCA	-251	(169)	(169)	ND

ance is not observed in the STAT1-binding site model derived from the *in vitro* selected sequences (Fig. 2) may be a reflection of the high stringency of the selection conditions.

The SIE element in the c-*fos* promoter (XI: SIE′, tccGTA TTCCCGTCA ATGCA) is the only sequence with two mismatches in one half-site showing clear affinity for STAT1 among all the sequences tested. This sequence is not bound by STAT5A, STAT5B, or STAT6, but it does bind to STAT1. We have argued above that STAT1 recognizes this oligonucleotide as an N_3 site with two mismatches rather than as an N_4 site with one mismatch. As an N_3 site the SIE element contains a T at position +2. Thus, the ability of STAT1 to recognize this sequence may be related to its capacity to bind TAA half-sites. Examining 15 other N_3 sites with two mismatches in one palindrome (all elements from classes V to IX and XI, spacer N3/N4 and N4 artificial 4), we observed no significant binding to STAT1.

One of these (VIII, IL-4R) does contain a T at position +2. This suggests that STAT1 binding to TTCN_3TAA sites with two mismatches may depend on other residues. Indeed, the G at position +7 in the SIE element may contribute to its

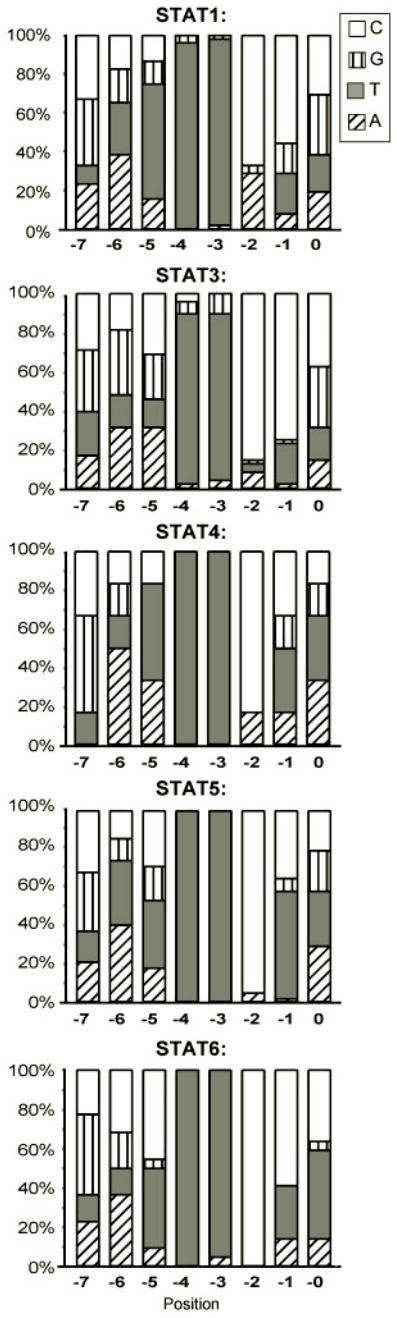

FIG. 4. **Base frequencies by position derived from manual alignment of the half-sites of the STAT-responsive elements described in the literature.** The number of half-sites analyzed was 52 for STAT1, 48 for STAT3, 6 for STAT4, 62 for STAT5, and 22 for STAT6 (for numerical values see Table IV, for full sequences see Table III).

TABLE IV
Base frequency matrices of STAT half-sites of cis-acting elements of STAT-responsive genes (see Fig. 4 and Table III)

	POSITION:							
STAT1	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
A	23	38	15	*0*	*2*	*29*	8	19
T	10	27	60	*96*	*96*	*0*	21	19
G	35	17	12	*4*	*2*	*4*	15	31
C	33	17	13	*0*	*0*	*67*	56	31
STAT3	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
A	17	31	31	*2*	*4*	*8*	2	15
T	23	17	15	*88*	*85*	*4*	21	17
G	31	33	23	*6*	*10*	*2*	2	31
C	29	19	31	*4*	*0*	*85*	75	38
STAT4	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
A	0	50	33	*0*	*0*	*17*	17	33
T	17	17	50	*100*	*100*	*0*	33	33
G	50	17	0	*0*	*0*	*0*	17	17
C	33	17	17	*0*	*0*	*83*	33	17
STAT5A	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**0**
A	21	40	18	*0*	*0*	*5*	2	29
T	16	34	35	*100*	*100*	*0*	56	29
G	31	11	18	*0*	*0*	*0*	6	21
C	32	15	29	*0*	*0*	*95*	35	21
STAT6	**-7**	**-6**	**-5**	**-4**	**-3**	**-2**	**-1**	**-0**
A	23	36	9	*0*	*5*	*0*	14	14
T	14	14	41	*100*	*95*	*0*	27	45
G	41	18	5	*0*	*0*	*0*	0	5
C	23	32	45	*0*	*0*	*100*	59	36

[a] The palindromic nucleotides are shaded.

recognition by STAT1, since STAT1, but not STAT5A or STAT6, showed a preference for C at position −7 in the site selection experiment (Fig. 2). Note that the TTCN$_3$TAA site with two mismatches which fails to bind to STAT1 has a T at position +7. A role for this G is consistent with the crystal structure of the STAT1-DNA complex (21) which shows an interaction of Glu-421 with the exocyclic amino group of guanine at position +7.

Defined Bases Inside and Outside of the Palindrome Contribute to STAT Binding Specificity—After having analyzed effects of the spacer and the sequence of the palindromic half-sites on STAT1, -5, and -6 binding, we wanted to examine the role of the bases between the half-sites and flanking the palindrome. In the site selection experiment STAT1 selected very frequently a C in position −1 (Fig. 2). Comparing two N$_3$ sequences with conserved half-sites (II: spacer N3, GACAAA TTCGCCGAA TTTGG, and II: spacer ttcc, GACAAA TTC*C*CCGAA TTTGG) (substituted bases are underlined), we found that substituting a C for a G at position −1 indeed resulted in a very strong increase in the affinity for STAT1 but barely had any effect on the affinities of STAT5 or STAT6.

On the other hand, a T to C substitution at the same position next to a half-site with one mismatch (reverse IV: mIL-2rE site II′, CTCTC TTA*T*CAGAA ATGT . . . *versus* reverse IV: mIL-2rE site II-var1, . . . CTCTC TTA*C*CAGAA ATGTC . . .) hardly had any effect on the binding of any of the STATs tested. A similar result was obtained by Soldaini *et al.* (28). Whereas these experiments suggest that C at position −1 does improve the affinity of a TTC half-site for STAT1 but not for STAT5 or STAT6, it should be noted that among natural STAT targets TTCC is as frequent in the STAT5 as among the STAT1-responsive elements.

In the site selection experiments, there was a bias against C at position −5 in the STAT5- and STAT6-specific high affinity binding site models, whereas STAT1 showed no such bias. We did not observe any impact of base changes at this position in lower affinity sites. Substitution of A by C in two weak binding sites (II: spacer N3, GACAA*A* TTCGCCGAA TTTGG *versus* II: spacer cttc, GACAA*C* TTCGCCGAA TTTGG, and reverse IV: hIL-2rE site II, CTCT*A* TTATCAGAA ATGTC . . . *versus* reverse IV: mIL-2rE site II, CTCT*C* TTATCAGAA ATGTC . . .) resulted only in a slight decrease of affinity for STAT5 and STAT6 and had no significant effect on STAT1 binding. About 30% of STAT enhancer elements in STAT5- and in STAT6-activated genes contain a C in position −5.

The Specificities of STAT5A and STAT5B Are Indistinguishable—We investigated if weak affinity binding sites can distinguish between STAT5A and STAT5B (33–35). The amino acid

sequences of these two proteins are 96% identical, and until now, to our knowledge, no gene has been shown to be transactivated selectively by only one of the two STAT5 proteins. Nevertheless, analysis of mice lacking STAT5A (36, 37), STAT5B (38), or both proteins (39) have revealed that each has essential functions. Comparing the relative affinity of these two proteins for 29 sequences (Table II), we found no difference that was significantly greater than two. Thus, it seems likely that the nonredundancy of STAT5A and STAT5B during development is a reflection of nonoverlapping cell type-specific expression of the two proteins or of their capacity to interact with different proteins required for target gene activation rather than a result of differences in DNA binding specificity. Our results are in agreement with the findings of Soldaini *et al.* (28). Others have reported that STAT5B binds much more strongly to the APRE element of the rat α_2-macroglobulin promoter than STAT5A (40). They showed that this is due to a single amino acid difference at position 433 between STAT5A (glutamine) and STAT5B (glycine). The significance of their finding is unclear since all of the human, mouse, or rat STAT5A and STAT5B data base entries predict a glutamine at this position.

DISCUSSION

Earlier work from many groups had shown that the sequence specificity of different STAT proteins is very similar but not identical. The experiments described here define more precisely and quantify the specificity differences between STAT1, STAT5, and STAT6. As demonstrated previously (8), STAT6 differs from the other STATs in its preference for sites in which the two palindromic halves are separated by 4 rather than 3 nucleotides. Our results confirm earlier findings of others (15, 30) that this preference is not absolute and demonstrate that STAT6 can bind N_3 sites. In the site selection experiments STAT6 discriminated against N_3 sites no more than against single base changes in the core consensus region. The ratio between N_4 sites and N_3 sites selected by STAT6 (93:7) is the same as the ratio between sequences containing T at position −4 and those with other nucleotides at this position. With regard to STAT5 we show here that this protein can bind N_4 sites but that its affinity for N_3 sites is much higher. STAT1 appears incapable of recognizing N_4 sites. On the other hand, the site selection experiments suggest that STAT1 but not STAT5 is capable of binding to N_2 sites. Previous experiments indicate that STAT3 may also bind to N_2 sites (15). STAT1 also shows a stronger preference than STAT5 and STAT6 for C at position −1. Indeed among the 40 sequences tested in our experiments, the relative affinity of any sequence without a C in position −1 was at least 50 times lower than that of the reference probe. In this and in other more subtle aspects, STAT5 and STAT6 are more similar to each other than to STAT1 (see below), reflecting, presumably, the closer similarity of the STAT5 and STAT6 DNA binding domains.

In general the differences suggested by the matrices derived from the sites that were selected for high affinity binding (Fig. 2) are consistent with the affinity differences of the various STAT proteins for weaker binding sites. This indicates that the STAT-specific differences between the high affinity matrices are not due to variations in the site selection conditions.

There are two characteristics in which STAT1 differs from STAT5 and STAT6, apart from its inability to bind N_4 sites. One is the preference for C at position −7. The crystal structure of the STAT1-DNA complex suggests that the interaction between Glu-421 and the C at this position is a structural correlate of this preference (see Supplemental Material). It has not been found in STAT3 crystals, and it will be interesting to observe whether is also absent in STAT5 and STAT6 crystals. The other STAT1-specific feature is the capacity of this protein to bind palindromes with one TTA half-site that are much less well recognized by STAT5 or STAT6. This difference might be explained by the fact that the asparagine in STAT1 which makes contacts with positions 1–3 of the binding site is replaced by a histidine in STAT5 and STAT6 (see Supplemental Material). A combination of these STAT1 characteristics may be responsible for the apparently greater tolerance of this protein for mismatches in the canonical palindrome.

The experiments reported here, together with previous studies from other laboratories, reveal clear differences in the fine specificity of STAT proteins. The finding that these differences are reflected among the natural target sites of different STATs indicates that they do contribute to selective gene activation by these proteins. They may account, for example, for the observation that in the same T cell line the *IL-2Rα* gene is induced by IL-2 whereas the *IL-4R* gene responds to IL-4 (7). On the other hand, the example of the different phenotypes of STAT5A- and of STAT5B-deficient mice reveals that other mechanisms, such as cell type-specific expression of receptors and signaling proteins, protein-protein interactions, and regulation of the chromatin conformation of target genes contribute to STAT target selection. To determine the relative importance of these mechanisms will require quite complex experiments.

Acknowledgments—The computational analysis of the *in vitro* selected sequences was performed with resources maintained at the Swiss Institute of Bioinformatics. We thank C. Rusterholz and P. Corthésy-Henrioud for help with the gel-shift experiments and advice; S. Cherpillod for help with the preparation of the manuscript; and F. Guilleux and O. Hagenbüchle for DNA probes. We also thank W. Meyer, F. Meylan, V. Imbert, C. Saner, O. Zilian, and F. Radtke for very useful suggestions and support.

REFERENCES

1. Fu, X.-Y., Kessler, D. S., Veals, S. A., Levy, D. E., and Darnell, J. E. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 8555–8559
2. Kessler, D. S., Veals, S. A., Fu, X. Y., and Levy, D. E. (1990) *Genes Dev.* **4,** 1753–1765
3. Schindler, C., and Darnell, J. E., Jr. (1995) *Annu. Rev. Biochem.* **64,** 621–651
4. Darnell, J. E., Jr. (1997) *Science* **277,** 1630–1635
5. Gaffen, S. L., Lai, S. Y., Longmore, G. D., Liu, K. D., and Goldsmith, M. A. (1999) *Blood* **94,** 74–86
6. Rusterholz, C., Corthésy Henrioud, P., and Nabholz, M. (1999) *Mol. Cell. Biol.* **19,** 2681–2689
7. Jankovic, D. L., Gibert, M., Baran, D., Ohara, J., Paul, W. E., and Theze, J. (1989) *J. Immunol.* **142,** 3113–3120
8. Schindler, U., Wu, P., Rothe, M., Brasseur, M., and McKnight, S. L. (1995) *Immunity* **2,** 689–697
9. Horvath, C. M., Wen, Z., and Darnell, J. E., Jr. (1995) *Genes Dev.* **9,** 984–994
10. Xu, X., Sun, Y.-L., and Hoey, T. (1996) *Science* **273,** 794–797
11. Vinkemeier, U., Cohen, S. L., Moarefi, I., Chait, B. T., Kuriyan, J., and Darnell, J. J. E. (1996) *EMBO J.* **15,** 5616–5626
12. Meyer, W. K., Reichenbach, P., Schindler, U., Soldaini, E., and Nabholz, M. (1997) *J. Biol. Chem.* **272,** 31821–31828
13. John, S., Vinkemeier, U., Soldaini, E., Darnell, J. E., Jr., and Leonard, W. J. (1999) *Mol. Cell. Biol.* **19,** 1910–1918
14. Lamb, P., Seidel, H. M., Haslam, J., Milocco, L., Kessler, L. V., Stein, R. B., and Rosen, J. (1995) *Nucleic Acids Res.* **23,** 3283–3289
15. Seidel, H. M., Milocco, L. H., Lamb, P., Darnell, J. E., Stein, R. B., and Rosen, J. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 3041–3045
16. Dreier, B., Fritz, S., Volkert, G., Brachvogel, B., Lottspeich, F., and Fey, G. H. (1997) *J. Biol. Chem.* **272,** 22118–22124
17. Pollock, R., and Treisman, R. (1990) *Nucleic Acids Res.* **18,** 6197–6204
18. Schindler, U., Terzaghi, W., Beckmann, H., Kadesch, T., and Cashmore, A. R. (1992) *EMBO J.* **11,** 1275–1289
19. Pearse, R. N., Feinman, R., Shuai, K., Darnell, J. J. E., and Ravetch, J. V. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 4314–4318
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E., and Kuriyan, J. (1998) *Cell* **93,** 827–839
22. Cardon, L. R., and Stormo, G. D. (1992) *J. Mol. Biol.* **223,** 159–170
23. Durbin, R., Eddy, S., Krogh, A., and Mitchison, G. (1998) *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, UK
24. Lew, D. J., Decker, T., Strehlow, I., and Darnell, J. E. (1991) *Mol. Cell. Biol.* **11,** 182–191
25. Decker, T., Lew, D. J., Mirkovitch, J., and Darnell, J. E. J. (1991) *EMBO J.*

10, 927–932
26. Decker, T., Kovarik, P., and Meinke, A. (1997) *J. Interferon Cytokine Res.* **17,** 121–134
27. Becker, S., Groner, B., and Muller, C. W. (1998) *Nature* **394,** 145–151
28. Soldaini, E., John, S., Moro, S., Bollenbacher, J., Schindler, U., and Leonard, W. J. (2000) *Mol. Cell. Biol.* **20,** 389–401
29. Pearse, R. N., Feinman, R., and Ravetch, J. V. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 11305–11309
30. Kotanides, H., and Reich, N. C. (1993) *Science* **262,** 1265–1267
31. Wagner, B. J., Hayes, T. E., Hoban, C. J., and Cochran, B. H. (1990) *EMBO J.* **9,** 4477–4484
32. Worm, M. M., Tsytsykova, A., and Geha, R. S. (1998) *Eur. J. Immunol.* **28,** 901–906
33. Mui, A. L.-F., Wakao, H., O'Farrell, A.-M., Harada, N., and Miyajima, A. (1995) *EMBO J.* **14,** 1166–1175
34. Liu, X., Robinson, G. W., Gouilleux, F., Groner, B., and Hennighausen, L. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 8831–8835
35. Lin, J.-X., Mietz, J., Modi, W. S., John, S., and Leonard, W. J. (1996) *J. Biol. Chem.* **271,** 10738–10744
36. Liu, X., Robinson, G. W., Wagner, K.-W., Garrett, L., Wynshaw-Boris, A., and Hennighausen, L. (1997) *Genes Dev.* **11,** 179–186
37. Feldman, G. M., Rosenthal, L. A., Liu, X., Hayes, M. P., Wynshaw-Boris, A., Leonard, W. J., Hennighausen, L., and Finbloom, D. S. (1997) *Blood* **90,** 1768–1776
38. Udy, G. B., Towers, R. P., Snell, R. G., Wilkins, R. J., Park, S.-H., Ram, P. A., Waxman, D. J., and Davey, H. W. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 7239–7244
39. Teglund, S., McKay, C., Schuetz, E., van Deursen, J. M., Stravopodis, D., Wang, D., Brown, M., Bodner, S., Grosveld, G., and Ihle, J. N. (1998) *Cell* **93,** 841–850
40. Boucheron, C., Dumon, S., Santos, S. C., Moriggl, R., Hennighausen, L., Gisselbrecht, S., and Gouilleux, F. (1998) *J. Biol. Chem.* **273,** 33936–33941
41. Larner, A. C., David, M., Feldman, G. M., Igarashi, K., Hackett, R. H., Webb, D. S. A., Sweitzer, S. M., Petricoin, E. F., III, and Finbloom, D. S. (1993) *Science* **261,** 1730–1733
42. Sperisen, P., Wang, S. M., Soldaini, E., Pla, M., Bucher, P., Rusterholz, C., Corthésy, P., Reichenbach, P., and Nabholz, M. (1995) *J. Biol. Chem.* **270,** 10743–10753
43. Lécine, P., Algarté, M., Rameil, P., Beadling, C., Bucher, P., Nabholz, M., and Imbert, J. (1996) *Mol. Cell. Biol.* **16,** 6829–6840
44. John, S., Robbins, C. M., and Leonard, W. J. (1996) *EMBO J.* **15,** 5627–5635
45. Yuan, J., Wegenka, U. M., Lutticken, C., Buschmann, J., Decker, T., Schindler, C., Heinrich, P. C., and Horn, F. (1994) *Mol. Cell. Biol.* **14,** 1657–1668
46. Rothman, P., Li, S. C., Gorham, B., Glimcher, L., Alt, F., and Boothby, M. (1991) *Mol. Cell. Biol.* **11,** 5551–5561
47. Hayes, T. E., Kitchen, A. M., and Cochran, B. H. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84,** 1272–1276
48. Kordula, T., Rydel, R. E., Brigham, E. F., Horn, F., Heinrich, P. C., and Travis, J. (1998) *J. Biol. Chem.* **273,** 4112–4118
49. Ripperger, J., Fritz, S., Richter, K., Dreier, B., Schneider, K., Lochner, K., Marschalek, R., Hocke, G., Lottspeich, F., and Fey, G. H. (1995) *Ann. N. Y. Acad. Sci.* **762,** 252–260
50. Fujitani, Y., Nakajima, K., Kojima, H., Nakae, K., Takeda, T., and Hirano, T. (1994) *Biochem. Biophys. Res. Commun.* **202,** 1181–1187
51. Fujio, Y., Kunisada, K., Hirota, H., Yamauchi-Takihara, K., and Kishimoto, T. (1997) *J. Clin. Invest.* **99,** 2898–2905
52. Stephanou, A., Brar, B. K., Knight, R. A., and Latchman, D. S. (2000) *Cell Death Differ.* **7,** 329–330
53. Horita, M., Andreu, E. J., Benito, A., Arbona, C., Sanz, C., Benet, I., Prosper, F., and Fernandez-Luna, J. L. (2000) *J. Exp. Med.* **191,** 977–984
54. Sepp, T., and Tong-Starksen, S. E. (1997) *J. Virol.* **71,** 771–777
55. Nguyen, V. T., and Benveniste, E. N. (2000) *J. Biol. Chem.* **275,** 23674–23684
56. Li, J., Colovai, A. I., Cortesini, R., and Suciu-Foca, N. (2000) *Hum. Immunol.* **61,** 486–498
57. Muhlethaler-Mottet, A., Di Berardino, W., Otten, L. A., and Mach, B. (1998) *Immunity* **8,** 157–166
58. Dong, Y., Rohn, W. M., and Benveniste, E. N. (1999) *J. Immunol.* **162,** 4731–4739
59. Meraz, M. A., White, J. M., Sheehan, K. C. F., Bach, E. A., Rodig, S. J., Dighe, A. S., Kaplan, D. H., Riley, J. K., Greenlund, A. C., Campbell, D., Carver-Moore, K., DuBois, R. N., Clark, R., Aguet, M., and Schreiber, R. D. (1996) *Cell* **84,** 431–450
60. Lee, Y. J., and Benveniste, E. N. (1996) *J. Immunol.* **157,** 1559–1568
61. Piskurich, J. F., Wang, Y., Linhoff, M. W., White, L. C., and Ting, J. P. (1998) *J. Immunol.* **160,** 233–240
62. Shuai, K., Schindler, C., Prezioso, V. R., and Darnell, J. J. E. (1992) *Science* **258,** 1808–1812
63. O'Brien, C. A., and Manolagas, S. C. (1997) *J. Biol. Chem.* **272,** 15003–15010
64. Gustafson, K. S., and Ginder, G. D. (1996) *J. Biol. Chem.* **271,** 20035–20046
65. Caldenhoven, E., Coffer, P., Yuan, J. P., Vandestolpe, A., Horn, F., Kruijer, W., and Vandersaag, P. T. (1994) *J. Biol. Chem.* **269,** 21146–21154
66. Jahnke, A., and Johnson, J. P. (1994) *FEBS Lett.* **354,** 220–226
67. Look, D. C., Pelletier, M. R., Tidwell, R. M., Roswit, W. T., and Holtzman, M. J. (1995) *J. Biol. Chem.* **270,** 30264–30267
68. Naik, S. M., Shibagaki, N., Li, L. J., Quinlan, K. L., Paxton, L. L., and Caughman, S. W. (1997) *J. Biol. Chem.* **272,** 1283–1290
69. Hou, J., Baichwal, V., and Cao, Z. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 11641–11645
70. Walter, M. J., Look, D. C., Tidwell, R. M., Roswit, W. T., and Holtzman, M. J. (1997) *J. Biol. Chem.* **272,** 28582–28589
71. Tessitore, A., Pastore, L., Rispoli, A., Cilenti, L., Toniato, E., Flati, V., Farina, A. R., Frati, L., Gulino, A., and Martinotti, S. (1998) *Eur. J. Biochem.* **258,** 968–975
72. Chon, S. Y., Hassanain, H. H., Pine, R., and Gupta, S. L. (1995) *J. Interferon Cytokine Res.* **15,** 517–526
73. Chon, S. Y., Hassanain, H. H., and Gupta, S. L. (1996) *J. Biol. Chem.* **271,** 17247–17252
74. Gao, J., Morrison, D. C., Parmely, T. J., Russell, S. W., and Murphy, W. J. (1997) *J. Biol. Chem.* **272,** 1226–1230
75. David, M., Petricoin, E. F., III, Igarashi, K., Feldman, G. M., Finbloom, D. S., and Larner, A. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 7174–7178
76. Wang, Y. F., and Yu-Lee, L. Y. (1996) *Mol. Cell. Endocrinol.* **121,** 19–28
77. Le Stunff, C., and Rotwein, P. (1998) *Endocrinology* **139,** 859–866
78. Wang, Y., O'Neal, K. D., and Yu-Lee, L. (1997) *Mol. Endocrinol.* **11,** 1353–1364
79. Durbin, J. E., Hackenmiller, R., Simon, M. C., and Levy, D. E. (1996) *Cell* **84,** 443–450
80. Pine, R., Canova, A., and Schindler, C. (1994) *EMBO J.* **13,** 158–167
81. Sims, S. H., Cha, Y., Romine, M. F., Gao, P. Q., Gottlieb, K., and Deisseroth, A. B. (1993) *Mol. Cell. Biol.* **13,** 690–702
82. Rein, T., Muller, M., and Zorbas, H. (1994) *Nucleic Acids Res.* **22,** 3033–3037
83. Valente, A. J., Xie, J. F., Abramova, M. A., Wenzel, U. O., Abboud, H. E., and Graves, D. T. (1998) *J. Immunol.* **161,** 3719–3728
84. Guyer, N. B., Severns, C. W., Wong, P., Feghali, C. A., and Wright, T. M. (1995) *J. Immunol.* **155,** 3472–3480
85. Wong, P., Severns, C. W., Guyer, N. B., and Wright, T. M. (1994) *Mol. Cell. Biol.* **14,** 914–922
86. Ohmori, Y., and Hamilton, T. A. (1998) *J. Biol. Chem.* **273,** 29202–29209
87. Auernhammer, C. J., Bousquet, C., and Melmed, S. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 6964–6969
88. Kordula, T., and Travis, J. (1995) *Biochem. J.* **309,** 63–67
89. Kordula, T., and Travis, J. (1996) *Biochem. J.* **313,** 1019–1027
90. Min, W., Pober, J. S., and Johnson, D. R. (1996) *J. Immunol.* **156,** 3174–3183
91. Chatterjee-Kishore, M., Kishore, R., Hicklin, D. J., Marincola, F. M., and Ferrone, S. (1998) *J. Biol. Chem.* **273,** 16177–16183
92. Wegenka, U. M., Buschmann, J., Lutticken, C., Heinrich, P. C., and Horn, F. (1993) *Mol. Cell. Biol.* **13,** 276–288
93. Akira, S., Nishio, Y., Inoue, M., Wang, X. J., Wei, S., Matsusaka, T., Yoshida, K., Sudo, T., Naruto, M., and Kishimoto, T. (1994) *Cell* **77,** 63–71
94. Ripperger, J. A., Fritz, S., Richter, K., Hocke, G. M., Lottspeich, F., and Fey, G. H. (1995) *J. Biol. Chem.* **270,** 29998–30006
95. Fukuzawa, J., Booz, G. W., Hunt, R. A., Shimizu, N., Karoor, V., Baker, K. M., and Dostal, D. E. (2000) *Hypertension* **35,** 1191–1196
96. Mascareno, E., Dhar, M., and Siddiqui, M. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 5590–5594
97. Ruff-Jamison, S., Zhong, Z., Wen, Z., Chen, K., Darnell, J. E., Jr., and Cohen, S. (1994) *J. Biol. Chem.* **269,** 21933–21935
98. Robertson, L. M., Kerppola, T. K., Vendrell, M., Luk, D., Smeyne, R. J., Bocchiaro, C., Morgan, J. I., and Curran, T. (1995) *Neuron* **14,** 241–252
99. Kunisada, K., Tone, E., Fujio, Y., Matsui, H., Yamauchi-Takihara, K., and Kishimoto, T. (1998) *Circulation* **98,** 346–352
100. Kiuchi, N., Nakajima, K., Ichiba, M., Fukada, T., Narimatsu, M., Mizuno, K., Hibi, M., and Hirano, T. (1999) *J. Exp. Med.* **189,** 63–73
101. Hutt, J. A., O'Rourke, J. P., and DeWille, J. W. (2000) *J. Biol. Chem.* **275,** 29123–29131
102. Cantwell, C. A., Sterneck, E., and Johnson, P. F. (1998) *Mol. Cell. Biol.* **18,** 2108–2117
103. Yamada, T., Tobita, K., Osada, S., Nishihara, T., and Imagawa, M. (1997) *J. Biochem.* (*Tokyo*) **121,** 731–738
104. Zhang, D., Sun, M., Samols, D., and Kushner, I. (1996) *J. Biol. Chem.* **271,** 9503–9509
105. Zhao, Y., Nichols, J. E., Bulun, S. E., Mendelson, C. R., and Simpson, E. R. (1995) *J. Biol. Chem.* **270,** 16449–16457
106. Zhang, Z., Fuentes, N. L., and Fuller, G. M. (1995) *J. Biol. Chem.* **270,** 24287–24291
107. Coffer, P., Lutticken, C., van Puijenbroek, A., Klop-de Jonge, M., Horn, F., and Kruijer, W. (1995) *Oncogene* **10,** 985–994
108. Kojima, H., Nakajima, K., and Hirano, T. (1996) *Oncogene* **12,** 547–554
109. Nakajima, K., Kusafuka, T., Takeda, T., Fujitani, Y., Nakae, K., and Hirano, T. (1993) *Mol. Cell. Biol.* **13,** 3027–3041
110. Brown, R. T., Ades, I. Z., and Nordan, R. P. (1995) *J. Biol. Chem.* **270,** 31129–31135
111. Schumann, R. R., Kirschning, C. J., Unbehaun, A., Aberle, H. P., Knope, H. P., Lamping, N., Ulevitch, R. J., and Herrmann, F. (1996) *Mol. Cell. Biol.* **16,** 3490–3503
112. Demoulin, J. B., Maisin, D., and Renauld, J. C. (1999) *Eur. Cytokine Netw.* **10,** 49–56
113. Bellido, T., O'Brien, C. A., Roberson, P. K., and Manolagas, S. C. (1998) *J. Biol. Chem.* **273,** 21137–21144
114. Yu, C. R., Ortaldo, J. R., Curiel, R. E., Young, H. A., Anderson, S. K., and Gosselin, P. (1999) *J. Immunol.* **162,** 2785–2790
115. Ichiba, M., Nakajima, K., Yamanaka, Y., Kiuchi, N., and Hirano, T. (1998) *J. Biol. Chem.* **273,** 6132–6138
116. Bugno, M., Graeve, L., Gatsios, P., Koj, A., Heinrich, P. C., Travis, J., and Kordula, T. (1995) *Nucleic Acids Res.* **23,** 5041–5047
117. Symes, A. J., Corpus, L., and Fink, J. S. (1995) *J. Neurochem.* **65,** 1926–1933
118. Grigorieva, I., Grigoriev, V. G., Rowney, M. K., and Hoover, R. G. (2000) *J. Biol. Chem.* **275,** 7343–7350
119. Barbulescu, K., Becker, C., Schlaak, J. F., Schmitt, E., Meyer zum Buschenfelde, K. H., and Neurath, M. F. (1998) *J. Immunol.* **160,** 3642–3647
120. Galon, J., Sudarshan, C., Ito, S., Finbloom, D., and O'Shea, J. J. (1999) *J. Immunol.* **162,** 7256–7262
121. Feltus, F. A., Groner, B., and Melner, M. H. (1999) *Mol. Endocrinol.* **13,** 1084–1093

122. Jolivet, G., L'Hotte, C., Pierre, S., Tourkine, N., and Houdebine, L. M. (1996) *FEBS Lett.* **389,** 257–262
123. Ooi, G. T., Hurst, K. R., Poy, M. N., Rechler, M. M., and Boisclair, Y. R. (1998) *Mol. Endocrinol.* **12,** 675–687
124. Dumon, S., Santos, S. C., Debierre-Grockiego, F., Gouilleux-Gruart, V., Cocault, L., Boucheron, C., Mollat, P., Gisselbrecht, S., and Gouilleux, F. (1999) *Oncogene* **18,** 4191–4199
125. Socolovsky, M., Fallon, A. E., Wang, S., Brugnara, C., and Lodish, H. F. (1999) *Cell* **98,** 181–191
126. Silva, M., Benito, A., Sanz, C., Prosper, F., Ekhterae, D., Nunez, G., and Fernandez-Luna, J. L. (1999) *J. Biol. Chem.* **274,** 22165–22169
127. Kieslinger, M., Woldman, I., Moriggl, R., Hofmann, J., Marine, J. C., Ihle, J. N., Beug, H., and Decker, T. (2000) *Genes Dev.* **14,** 232–244
128. Lord, J. D., McIntosh, B. C., Greenberg, P. D., and Nelson, B. H. (2000) *J. Immunol.* **164,** 2533–2541
129. Wakao, H., Schmitt-Ney, M., and Groner, B. (1992) *J. Biol. Chem.* **267,** 16365–16370
130. Schmitt-Ney, M., Doppler, W., Ball, R. K., and Groner, B. (1991) *Mol. Cell. Biol.* **11,** 3745–3755
131. Verdier, F., Rabionet, R., Gouilleux, F., Beisenherz-Huss, C., Varlet, P., Muller, O., Mayeux, P., Lacombe, C., Gisselbrecht, S., and Chretien, S. (1998) *Mol. Cell. Biol.* **18,** 5852–5860
132. Mui, A. L. F., Wakao, H., Kinoshita, T., Kitamura, T., and Miyajima, A. (1996) *EMBO J.* **15,** 2425–2433
133. Matsumura, I., Kitamura, T., Wakao, H., Tanaka, H., Hashimoto, K., Albanese, C., Downward, J., Pestell, R. G., and Kanakura, Y. (1999) *EMBO J.* **18,** 1367–1377
134. Sasaki, Y., Takahashi, Y., Nakayama, K., and Kamataki, T. (1999) *J. Biol. Chem.* **274,** 37117–37124
135. Subramanian, A., Teixeira, J., Wang, J., and Gil, G. (1995) *Mol. Cell. Biol.* **15,** 4672–4682
136. Subramanian, A., Wang, J., and Gil, G. (1998) *Nucleic Acids Res.* **26,** 2173–2178
137. Lahuna, O., Rastegar, M., Maiter, D., Thissen, J. P., Lemaigre, F. P., and Rousseau, G. G. (2000) *Mol. Endocrinol.* **14,** 285–294
138. Galsgaard, E. D., Gouilleux, F., Groner, B., Serup, P., Nielsen, J. H., and Billestrup, N. (1996) *Mol. Endocrinol.* **10,** 652–660
139. Nakajima, H., Liu, X. W., Wynshaw-Boris, A., Rosenthal, L. A., Imada, K., Finbloom, D. S., Hennighausen, L., and Leonard, W. J. (1997) *Immunity* **7,** 691–701
140. Kang, C.-J., Sheridan, C., and Koshland, M. E. (1998) *Immunity* **8,** 285–295
141. Lu, L., Zhu, J., Zheng, Z., Yan, M., Xu, W., Sun, L., Theze, J., and Liu, X. (1998) *Eur. J. Immunol.* **28,** 805–810
142. Yoshimura, A., Ichihara, M., Kinjyo, I., Moriyama, M., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., Hara, T., and Miyajima, A. (1996) *EMBO J.* **15,** 1055–1063
143. Wang, D., Stravopodis, D., Teglund, S., Kitazawa, J., and Ihle, J. N. (1996) *Mol. Cell. Biol,* **16,** 6141–6148
144. Ma, Y., Streiff, R. J., Liu, J., Spence, M. J., and Vestal, R. E. (1999) *Nucleic Acids Res.* **27,** 4649–4657
145. Zhang, J., Scordi, I., Smyth, M. J., and Lichtenheld, M. G. (1999) *J. Exp. Med.* **190,** 1297–1308
146. Galsgaard, E. D., Nielsen, J. H., and Moldrup, A. (1999) *J. Biol. Chem.* **274,** 18686–18692
147. Wood, T. J., Sliva, D., Lobie, P. E., Pircher, T. J., Gouilleux, F., Wakao, H., Gustafsson, J. A., Groner, B., Norstedt, G., and Haldosen, L. A. (1995) *J. Biol. Chem.* **270,** 9448–9453
148. Le Cam, A., Pantescu, V., Paquereau, L., Legraverend, C., Fauconnier, G., and Asins, G. (1994) *J. Biol. Chem.* **269,** 21532–21539
149. Bergad, P. L., Shih, H.-M., Towles, C. T., Schwarzenberg, S. J., and Berry, S. A. (1995) *J. Biol. Chem.* **270,** 24903–24910
150. Ye, S. K., Maki, K., Kitamura, T., Sunaga, S., Akashi, K., Domen, J., Weissman, I. L., Honjo, T., and Ikuta, K. (1999) *Immunity* **11,** 213–223
151. Conrad, D. J., and Lu, M. (2000) *Am. J. Respir. Cell Mol. Biol.* **22,** 226–234
152. Berton, M. T., and Linehan, L. A. (1995) *J. Immunol.* **154,** 4513–4525
153. Xu, M. Z., and Stavnezer, J. (1992) *EMBO J.* **11,** 145–155
154. Schaffer, A., Cerutti, A., Shah, S., Zan, H., and Casali, P. (1999) *J. Immunol.* **162,** 5327–5336
155. Pan, Q., Petit-Frere, C., Stavnezer, J., and Hammarstrom, L. (2000) *Eur. J. Immunol.* **30,** 1019–1029
156. Matsukura, S., Stellato, C., Plitt, J. R., Bickel, C., Miura, K., Georas, S. N., Casolaro, V., and Schleimer, R. P. (1999) *J. Immunol.* **163,** 6876–6883
157. Tinnell, S. B., Jacobs-Helber, S. M., Sterneck, E., Sawyer, S. T., and Conrad, D. H. (1998) *Int. Immunol.* **10,** 1529–1538
158. Richards, M. L., and Katz, D. H. (1997) *J. Immunol.* **158,** 263–272
159. Takeda, K., Tanaka, T., Shi, W., Matsumoto, M., Minami, M., Kashiwamura, S., Nakanishi, K., Yoshida, N., Kishimoto, T., and Akira, S. (1996) *Nature* **380,** 627–630
160. Park, H. J., So, E. Y., and Lee, C. E. (1998) *Mol. Immunol.* **35,** 239–247
161. Kohler, I., and Rieber, E. P. (1993) *Eur. J. Immunol.* **23,** 3066–3071
162. Fenghao, X., Saxon, A., Nguyen, A., Ke, Z., Diaz-Sanchez, D., and Nel, A. (1995) *J. Clin. Invest.* **96,** 907–914
163. Albrecht, B., Peiritsch, S., and Woisetschlager, M. (1994) *Int. Immunol.* **6,** 1143–1151
164. Mikita, T., Campbell, D., Wu, P., Williamson, K., and Schindler, U. (1996) *Mol. Cell. Biol.* **16,** 5811–5820
165. Hill, S., Herlaar, E., Le Cardinal, A., van Heeke, G., and Nicklin, P. (1999) *Am. J. Respir. Cell Mol. Biol.* **21,** 728–737
166. Linehan, L. A., Warren, W. D., Thompson, P. A., Grusby, M. J., and Berton, M. T. (1998) *J. Immunol.* **161,** 302–310
167. Kaplan, M. H., Schindler, U., Smiley, S. T., and Grusby, M. J. (1996) *Immunity* **4,** 313–319
168. Khew-Goodall, Y., Wadham, C., Stein, B. N., Gamble, J. R., and Vadas, M. A. (1999) *Arterioscler. Thromb. Vasc. Biol.* **19,** 1421–1429
169. Ohmori, Y., Smith, M. F., Jr., and Hamilton, T. A. (1996) *J. Immunol.* **157,** 2058–2065