Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 8831–8835, September 1995
Developmental Biology

Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in prolactin signal transduction in mouse mammary tissue

(Janus kinase/signal transducers and activators of transcription/mammary gland factor/milk protein)

XIUWEN LIU*, GERTRAUD W. ROBINSON*, FABRICE GOUILLEUX†, BERND GRONER†, AND LOTHAR HENNIGHAUSEN*

*Laboratory of Biochemistry and Metabolism, National Institute of Diabetes, Digestive, and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892; and †Institute for Experimental Cancer Research, Tumor Biology Center, Breisacher Strasse 117, D-79106 Freiburg, Germany

Communicated by Gilbert Ashwell, National Institutes of Health, Bethesda, MD, May 19, 1995 (received for review March 20, 1995)

ABSTRACT **Prolactin (PRL) induces transcriptional activation of milk protein genes, such as the whey acidic protein (WAP), β-casein, and β-lactoglobulin genes, through a signaling cascade encompassing the Janus kinase Jak2 and the mammary gland factor (MGF; also called Stat5), which belongs to the family of proteins of signal transducers and activators of transcription (STAT). We isolated and sequenced from mouse mammary tissue Stat5 mRNA and a previously unreported member, which we named Stat5b (Stat5 is renamed to Stat5a). On the protein level Stat5a and Stat5b show a 96% sequence similarity. The 5′ and 3′ untranslated regions of the two mRNAs are not conserved. Stat5a comprises 793 amino acids and is encoded by a mRNA of 4.2 kb. The Stat5b mRNA has a size of 5.6 kb and encodes a protein of 786 amino acids. Both Stat5a and Stat5b recognized the GAS site (γ-interferon-activating sequence; TTCNNNGAA) *in vitro* and mediated PRL-induced transcription in COS cells transfected with a PRL receptor. Stat5b also induced basal transcription in the absence of PRL. Similar levels of Stat5a and Stat5b mRNAs were found in most tissues of virgin and lactating mice, but a differential accumulation of the Stat5 mRNAs was found in muscle and mammary tissue. The two RNAs are present in mammary tissue of immature virgin mice, and their levels increase up to day 16 of pregnancy, followed by a decline during lactation. The increase of Stat5 expression during pregnancy coincides with the activation of the WAP gene.**

After binding of cytokines to the extracellular ligand-binding domains of their receptors, receptor-associated cytoplasmic tyrosine kinases belonging to the Janus kinase (JAK) family are activated (see refs. 1–3 for recent reviews). The JAKs cross phosphorylate each other and their associated membrane receptors. The latent cytoplasmic STAT (signal transducer and activator of transcription) proteins are recruited to the phosphotyrosine and serve as substrates for the JAKs. Subsequent to phosphorylation, the STATs dimerize and translocate to the nucleus where they bind to specific DNA sequences [GAS sites (γ-interferon-activating sequences)] and induce transcription. The JAK–STAT pathways are activated by many cytokines and other growth factors, including prolactin (PRL) (1–3).

PRL plays a central role in the development, differentiation, and function of the mammary gland (4). The binding of PRL to its cell surface receptor regulates the transcriptional activation of milk protein genes (5–8). Molecular mechanisms of PRL-mediated signal transduction have been studied in mammary epithelial cells, which can be induced to synthesize β-casein, in cultured cells under specialized conditions (5) and in transgenic mice (6–12). Through use of transgenic mice carrying whey acidic protein (WAP) genes, PRL response elements have been localized in the promoter/upstream region (6, 7, 9–12). Mammary gland factor (MGF; also called Stat5) recognizes the PRL response element TTCNNNGAA (GAS sequence) in the promoter region of the β-casein (5, 13), WAP (10), and β-lactoglobulin (8) genes. Mutations in this sequence in sheep β-lactoglobulin (8) and rat WAP transgenes (10) led to reduced expression of these genes in transgenic mice. This suggests that Stat5 mediates in part the transcriptional activation of milk protein genes, which is induced during pregnancy and lactation by lactogenic hormones. Here we describe the characterization of Stat5 and a closely related member (Stat5b) from mouse mammary tissue.‡

MATERIALS AND METHODS

Construction and Screening of cDNA Libraries. Total RNA was extracted from mammary glands of C57/B6 mice on day 1 of lactation. Poly(A)$^+$ RNA was isolated by using the Oligotex mRNA kit from Qiagen (Chatsworth, CA). A cDNA library was made with the ZAP cDNA synthesis kit from Stratagene. The sheep MGF/Stat5 cDNA clone (13) was used as a probe to screen for the mouse homologue. The Prime-It II kit from Stratagene was used to label the probe by random-priming with [α-^{32}P]dCTP. Hybridization was performed in 0.43 M sodium phosphate, pH 7.2/6.5% SDS/1% bovine serum albumin/0.02 M EDTA at 60°C. Positive plaques were excised (Stratagene), and plasmids were isolated by using Wizard Miniprep columns (Promega).

Sequence Analysis. Sequence analysis was done by using the Genetics Computer Group package (14), including DNA-sequence entering (FRAGMENT ASSEMBLY programs), editing (SEQED), and mapping (MAP) as well as data-base searching (FASTA and FETCH) and multiple sequence analysis (PILEUP, PLOTSIMILARITY, DISTANCES).

Plasmid Construction and DNA Transfections. The plasmids encoding the sheep MGF/Stat5 (pXM-MGF), the long form of the murine PRL receptor, and the β-casein promoter with the luciferase reporter gene (pZZ1) have been described (13, 15). pXM-Stat5a was constructed by cloning a 2.7-kb fragment containing 40 bp of 5′ and 250 bp of 3′ untranslated sequences plus 2.4 kb of coding sequence into the pXM expression vector (16). pXM-Stat5b containing mouse Stat5b cDNA was constructed similarly and contains 2.4 kb of coding sequence plus 25 bp of 5′ and about 2.0 kb of 3′ untranslated sequences. Transfections were performed by the calcium phos-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: JAK, Janus kinase; STAT, signal transducer and activator of transcription; PRL, prolactin; WAP, whey acidic protein; MGF, mammary gland factor; GAS, γ-interferon activating sequence; IL-4, interleuken 4.

‡The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U21103 for mouse Stat5a and U21110 for mouse Stat5b; X78428 for sheep Stat5 was updated Dec. 14, 1994).

phate precipitation method (13, 15). Five micrograms of the PRL receptor construct plus 5 μg of one of the plasmids pXM-MGF, pXM-Stat5a, pXM-Stat5b, or pXM vector control were used in each transfection. In addition, 0.5 μg of the plasmid pCH110 containing the β-galactosidase gene under the control of the simian virus 40 promoter (15) was included as an internal control for transfection efficiency. When the reporter gene assays were carried out, 2 μg of the plasmid pZZ1 was used, and the amount of the PRL receptor and of the STAT expression constructs was reduced to 3 μg.

Culture of COS cells, luciferase and β-galactosidase assays, and electrophoretic mobility-shift assay (EMSA) were performed as described (13, 15). The probe in the EMSA was derived from the β-casein gene promoter (5′-AGATTTCTAGGAATTCAAATC-3′) as described (13, 15). The antiserum against the sheep Stat5 (glutathione *S*-transferase-Stat5 fusion protein containing the N-terminal part of sheep Stat5 from amino acid residues 6 to 132) used in the supershift was generated in chickens.

RNA Extraction, Probe Preparation, and RNA Blot-Hybridization Analysis. RNA was isolated either by acidic phenol extraction (17) or the RNeasy columns (Qiagen). The gene-specific probe for Stat5a was derived from a 1.0-kb PCR fragment immediately downstream of the TGA codon. Similarly, a 1.5-kb PCR fragment from the 3′ untranslated region of Stat5b was used as a Stat5b-specific probe. The oligonucleotide probe for the WAP messages has been described (9). Total RNA was separated in 1.0% formaldehyde-agarose gels (18).

RESULTS

Isolation and Characterization of Mouse Stat5 cDNAs. A phage λ library was prepared from mRNA of mouse mammary tissue and screened with a sheep Stat5 cDNA probe. Fifteen independent clones were obtained and sequenced, of which 13 represented the mouse homologue of sheep Stat5 and 2 encoded a Stat5-like protein (now named Stat5b). The longest mouse Stat5 (renamed Stat5a) cDNA consisted of 2.4 kb of protein coding sequence, plus 0.2 and 1.2 kb of 5′ and 3′ untranslated sequence, respectively. The deduced mouse Stat5a sequence encompasses 793 amino acids (Fig. 1). Although the sheep Stat5 and the mouse Stat5a sequences are highly conserved, several differences are noteworthy. The mouse and sheep sequences differ significantly in four regions (amino acid addition or deletion, underlined and labeled A through D in Fig. 1). An extra amino acid in the sheep protein is located at position 97 within a stretch of 17 amino acids that is not conserved between the two species (Region A in Fig. 1). We resequenced this region in a sheep Stat5 cDNA clone (13) and detected the insertion of three T residues over a stretch of 51 nucleotides in the original sequence (data presented to reviewers). These erroneous frameshifts resulted in an insertion of a single amino acid and an incorrect sequence spanning 17 amino acids. The corrected sheep sequence in this region corresponds to the mouse Stat5a sequence. We also sequenced a second region in the sheep Stat5 clone corresponding to amino acids 259–270 (Region B in Fig. 1) and found the protein sequence should also read the same as the corresponding mouse Stat5 proteins. It is not clear at this point whether the additional differences in region C and D are genuine or the result of erroneous frameshifts in the sheep sequence.

Stat5b Is an Additional Member of the STAT Family. The Stat5b cDNA clone has a size of 5.6 kb containing 0.5 kb and 2.7 kb of 5′ and 3′ untranslated sequence, respectively, and 2.4-kb of protein coding sequence. The 5′ and 3′ untranslated regions are not conserved between Stat5a and Stat5b. Overall,

```
     *            *   *               *       *  *       ****                                         *
5a MAGWIQAQQLQGDALRQMQV LYGQHFPIEVRHYLAQWIES QPWDAIDLDNPQDRGQATQL LEGLVQELQKKAEHQVGEDG FLLKIKLGHYATQLQNT.YDR 100
5b MAmWIQAQQLQGDALhQMQa LYGQHFPIEVRHYLsQWIES QaWDsIDLDNPQenikATQL LEGLVQELQKKAEHQVGEDG FLLKIKLGHYATQLQsT.YDR 100
sp MAGWIQAQQLQGDALRQMQV LYGQHFPIEVRHYLAQWIES QPWDAIDLDNPQDRaQvTQL LEGLVQELQKKAEHQVGEDG FLLKIKLGHYvhvssrTrtta 101
                                                                                                  Region A

                             *     * *             *
5a CPMELVRCIRHILYNEQRLV REANNCSSPAGVLVDAMSQK HLQINQRFEELRLITQDTEN ELKKLQQTQEYFIIQYQESL RIQAQFAQLGQLNPQERMSR  200
5b CPMELVRCIRHILYNEQRLV REANNgSSPAGsLaDAMSQK HLQINQtFEELRLITQDTEN ELKKLQQTQEYFIIQYQESL RIQAQFAQLGQLNPQERMSR  200
sp aPwswlRCIRHILYNEQRLV REAtNgnSsAGiLVDAMSQK HLQINQtFEELRLVTQDTEN ELKKLQQTQEYFIIQYQESL RIQAQFAQLaQLNPQERLSR  201

5a ETALQQKQVSLETWLQREAQ TLQQYRVELAEKHQKTLQLL RKQQTIILDDELIQWKRRQQ LAGNGGPPEGSLDVLQSWCE KLAEIIWQNRQQIRRAEHLC  300
5b ETALQQKQVSLETWLQREAQ TLQQYRVELAEKHQKTLQLL RKQQTIILDDELIQWKRRQQ LAGNGGPPEGSLDVLQSWCE KLAEIIWQNRQQIRRAEHLC  300
sp ETALQQKQVSLEaWLQREAQ TLQQYRVELAEKHQKTLQLL RKQQTIILDDELIQWKRRhd wrGmeaPPr.SLDVLQSWCE KLAEIIWQNRQQIRRAEHLC  300
                                                                 Region B
                                                                                                 **
5a QQLPIPGPVEEMLAEVNATI TDIISALVTSTFIIEKQPPQ VLKTQTKFAATVRLLVGGKL NVHMNPPQVKATIISEQQAK SLLKNENTRNECSGEILNNC  400
5b QQLPIPGPVEEMLAEVNATI TDIISALVTSTFIIEKQPPQ VLKTQTKFAATVRLLVGGKL NVHMNPPQVKATIISEQQAK SLLKNENTRNdySGEILNNC  400
sp QQLPIPGPVEEMLAEVNATI TDIISALVTSTFIIEKQPPQ VLKTQTKFAATVRLLVGGKL NVHMNPPQVKATIISEQQAK SLLKNENTRNECSGEILNNC  400

                              *     *         *  *      *                        *
5a CVMEYHQATGTLSAHFRNMS LKRIKRADRRGAESVTEEKF TVLFESQFSVGSNELVFQVK TLSLPVVVIVHGSQDHNATA TVLWDNAFAEPGRVPFAVPD  500
5b CVMEYHQATGTLSAHFRNMS LKRIKRsDRRGAgSVTEEKF TiLFdSQFSVGgNELVFQVK TLSLPVVVIVHGSQDnNATA TVLWDNAFAEPGRVPFAVPD  500
sp CVMEYHQrTGTLSAHFRNMS LKRIKRADRRGAESVTEEKF TVLFESQFSVGSNELVFQVK TLSLPVVVIVHGSQDHNATA TVLWDNAFAEPGRVPFAVPD  500

                                                                       *                   *
5a KVLWPQLCEALNMKFKAEVQ SNRGLTKENLVFLAQKLFNI SSNHLEDYNSMSVSWSQFNR ENLPGWNYTFWQWFDGVMEV LKKHHKPHWNDGAILGFVNK  600
5b KVLWPQLCEALNMKFKAEVQ SNRGLTKENLVFLAQKLFNI SSNHLEDYNSMSVSWSQFNR ENLPGrNYTFWQWFDGVMEV LKKHlKPHWNDGAILGFVNK  600
sp KVLWPQLCEALNMKFKAEVQ SNRGLTKENLlFLAQKLFNn SSsHLEDYNgMSVSWSQFNR ENLPGWNYTFWQWFDGVMEV LKKHHKPHWNDGAILGFVNK  600

                                       ** **    *                                   **       ******   *
5a QQAHDLLINKPDGTFLLRFS DSEIGGITIAWKFDSPDRNL WNLKPFTTRDFSIRSLADRL GDLNYLIYVFPDRPKDEVFA KYYTPV.....LAKAVDGYV  695
5b QQAHDLLINKPDGTFLLRFS DSEIGGITIAWKFDSqeRmf WNLmPFTTRDFSIRSLADRL GDLNYLIYVFPDRPKDEVys KYYTPVpcepatAKAaDGYV  700
sp QQAHDLLINKPDGTFLLRFS DSEIGGITIAWKFDSPDRNL WNLKPFTTREgSIRSLADRL GDLNYLIYVFPDRPKDEVFs KYYTPV.....LAKAVDGYV  695

              *         **               *           *   *  *   * **     *     *     * *******
5a KPQIKQVVPEFVNASTD.AG ASATYMDQAPSPVVCPQPHY NMYPPNPDPVLDQDGEFDLD ESMDVARHVEELLRRPMDSL DARLSPP.AGLFTSARSSLS* 793
5b KPQIKQVVPEFaNASTD.AG sgATYMDQAPSPVVCPQaHY NMYPPNPDsVLDtDGdFDLe dtMDVARrVEELLgRPMDSq wiphaqs*              786
sp KPQIKQVVPEFVsASaDsAG sSATYMDQAPSPaVCPQPHY NMYPqNPDPVLDQDGEFDLD EtMDVARHVEELLRRPnqqs qp.LSPPpAGLFTpARgSLS* 794
              Region C                                                              Region D
```

FIG. 1. Alignment of the deduced protein sequences of mouse Stat5a (line 5a), mouse Stat5b (line 5b), and sheep Stat5 (line sp). Amino acids different from Stat5a are shown in lowercase letters and the differences between Stat5a and Stat5b are indicated by the asterisks. The sheep Stat5 amino acid sequence is derived from the corrected nucleotide sequence. Regions where amino acid addition or deletion is involved between the sheep Stat5 and the mouse Stat5a are underlined and labeled A through D.

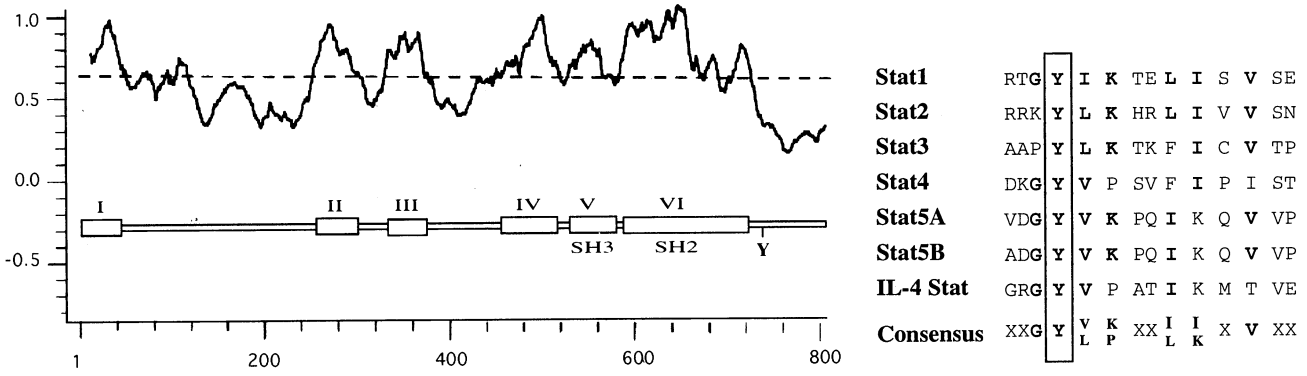
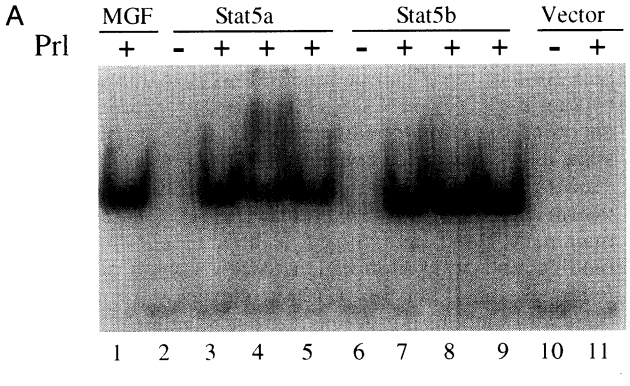


FIG. 2. Sequence comparison of STAT proteins. (*Left*) Similarity plot of STAT proteins. Members of the STAT family (Stat1 to Stat6) were aligned by using the PILEUP program of the Genetics Computer Group package and edited with LINEUP. PLOTSIMILARITY was used to analyze the conservation of amino acids at each position in the aligned sequences. The program generates a score of 1.5 for perfect matches and values <1.5 (even negative values) for nonmatches. The broken line represents the average similarity across the entire alignment. The numbers refer to that in the overall alignment after introducing gaps in each individual STAT protein. The diagram below shows some of the known structural features of the STAT proteins. Stat2 and IL-4 Stat (Stat6) are human while the rest are mouse sequences. (*Right*) Alignment of the amino acids around the putative phosphotyrosine. GenBank accession numbers are as follows: Stat1, U06924; Stat2, M97934; Stat3, U06922; Stat4, U06923; Stat5a, U21103; Stat5b, U21110; and Stat6, U16030.

the two proteins show a sequence similarity of 96% (Fig. 1). Of the 45 amino acid changes between amino acid residues 1 and 780, 20 represent conservative changes. The main difference between Stat5a and Stat5b lies within the C terminus. The C-terminal 8 amino acids in Stat5b are completely diverged from those in Stat5a. Stat5b is also 12 amino acids shorter than Stat5a at the C terminus.

Sequence Conservation Within the STAT Family. A comparison of the different STAT proteins revealed that the interleukin 4 (IL-4) Stat (ref. 19; now named Stat6) is the closest relative to Stat5, showing 43% and 42% similarity to Stat5a and Stat5b, respectively (data not shown). In the PILEUP alignment of all of the STAT proteins, there are 11 blocks where Stat5a, Stat5b, and IL-4 Stat (Stat6) have similar amino acid insertions or deletions compared to the other STAT members (data not shown). In the similarity plot (Fig. 2 *Left*), six conserved regions have similarity scores well above the average. Two of those correspond to the previously identified Src homology (SH) domains SH3 (labeled V in Fig. 2 *Left*) and SH2 (labeled VI in Fig. 2 *Left*). No similar sequences were found in the data base for the other four conserved domains (labeled I through IV in Fig. 2 *Left*), implying that these regions may represent new functional motifs specific for the STAT proteins. The alignment of all of the STAT proteins at the region of the putative tyrosine phosphorylation sites revealed a consensus sequence of Gly-Tyr-(Val or Leu)-(Lys or Pro)-Xaa-Xaa-(Ile or Leu)-(Ile or Lys)-Xaa-Val (Fig. 2 *Right*), which is different from the traditional tyrosine phosphorylation sites.

Stat5a and Stat5b Bind to the GAS Site and Activate PRL-Induced Transcription. COS cells expressing the PRL receptor and either Stat5a or Stat5b were induced with PRL. Nuclear extracts were prepared and band-shift experiments were performed with a oligonucleotide containing the GAS (Fig. 3*A*). Binding of Stat5a and Stat5b to the GAS site was observed after inducing the cells with PRL (Fig. 3*A*, lanes 3 and 7) but not in the absence of the hormone (Fig. 3*A*, lanes 2 and 6). The mouse Stat5a complex and the sheep MGF/Stat5 complex comigrated, but the Stat5b complex had a faster mobility (Fig. 3*A*, lanes 1, 3, and 7). The Stat5a complex was



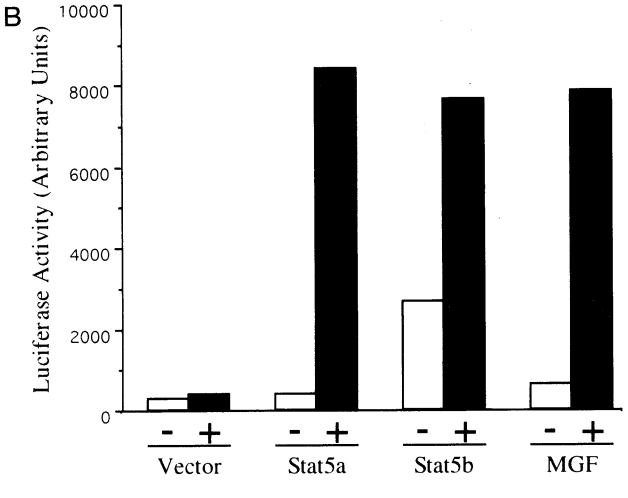


FIG. 3. Binding of Stat5a and Stat5b to the GAS site (*A*) and PRL-dependent activation of the β-casein promoter in COS cells by Stat5a and Stat5b (*B*). (*A*) COS cells were transfected with ovine MGF/Stat5 (lane 1) or the mouse Stat5a (lanes 2–5) or Stat5b (lanes 6–9) cDNAs or the vector only (lanes 10 and 11) as indicated plus the murine PRL receptor construct. Transfected cells were treated with PRL for 30 min (+) or not treated (−). Nuclear extracts were prepared and analyzed in band-shift assays with a labeled DNA probe containing the Stat5 binding site in the bovine β-casein promoter. Supershift was performed with an antiserum against the sheep MGF (lanes 4 and 8) or with the preimmune serum as a control. (*B*) COS cells were transfected with the β-casein-luciferase construct, PRL receptor cDNA, and one of the STAT constructs. Each transfection also included the simian virus 40 promoter-driven β-galactosidase gene as an internal control to monitor the transfection efficiency. Luciferase activity (arbitrary units) was normalized against the β-galactosidase activity values.

supershifted with an antibody generated against sheep MGF, but the Stat5b complex was not recognized by this antibody (Fig. 3*A*, lanes 4 and 8). No supershift complexes were observed with preimmune serum (Fig. 3*A*, lanes 5 and 9).

We determined whether Stat5a and Stat5b were able to induce transcription from the β-casein promoter containing a GAS site. A β-casein-luciferase reporter gene and a PRL receptor expression vector were transfected into COS cells together with either ä Stat5a or Stat5b expression vector, and luciferase activity was measured in the absence and presence of PRL (Fig. 3*B*). In the absence of PRL, Stat5a was unable to activate transcription, and an ≈20-fold induction was seen in the presence of PRL (Fig. 3*B*). Transactivation by Stat5b was observed already in the absence of PRL and was induced an additional 3-fold upon PRL stimulation (Fig. 3*B*).

Expression of Stat5a and Stat5b in Tissues and During Mammary Development. We analyzed the steady-state levels of Stat5a and Stat5b RNA in tissues taken from 6-week-old virgins (Fig. 4*A*) and from lactating females (Fig. 4*B*). Stat5a and Stat5b RNAs had sizes of 4.2 and 5.6 kb, respectively. In the virgin both RNA species were found at similar levels in most tissues with the exception of brain and muscle, which contained predominantly Stat5b (Fig. 4 *A* and *C*). With some exceptions, the steady-state levels of Stat5a and Stat5b on day 1 of lactation were similar to those found in the virgin (Fig. 4*B*). Most strikingly, the Stat5b message was highly abundant in muscle tissue of virgin and lactating females and in males (Fig. 4 *B* and *C*), whereas Stat5a was abundant in mammary tissue during lactation (Fig. 4 *B* and *C*). High levels of Stat5a RNA were also found in other secretory organs, such as salivary gland and seminal vesicle (Fig. 4*C*).

Stat5 is believed to play an important role in the activation of milk protein genes (5, 8, 10, 13). We therefore evaluated the

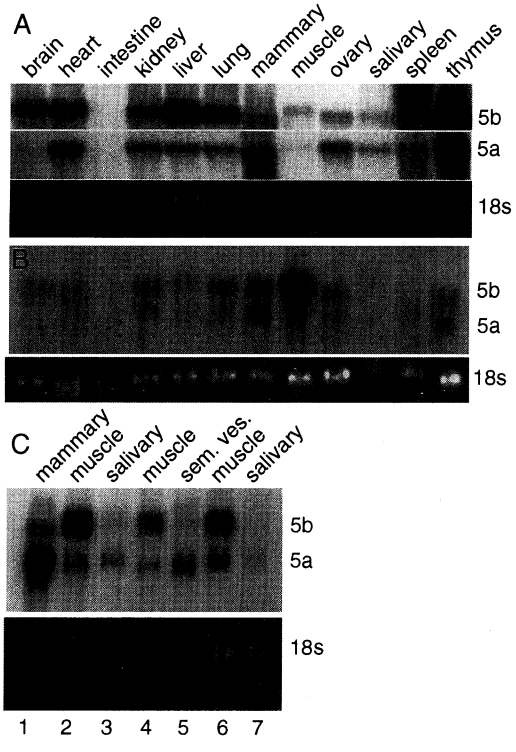


FIG. 4. RNA blot analysis of Stat5 messages from mouse tissues. Twenty micrograms of total RNA was separated on formaldehyde-agarose gels and probed with gene-specific probes for Stat5a and Stat5b. Tissues were from 6-week-old virgin female mice (*A*) or from mice on lactation day 1 (*B*). (*C*) RNA from a 1-day lactating mouse (lane 1), a mouse pregnant for 17 days (lanes 2 and 3), a virgin mouse (lanes 6 and 7), and a male mouse (lanes 4 and 5) was probed for Stat5a and Stat5b. 18s, 18S RNA.

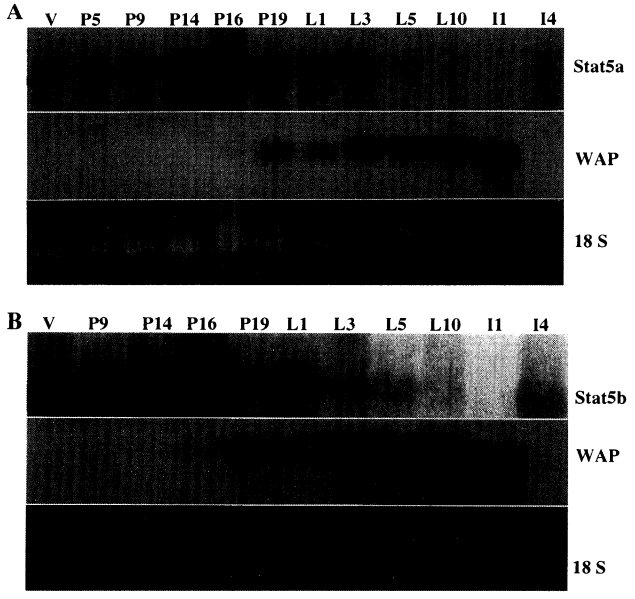


FIG. 5. Northern blot analysis of Stat5 messages from mouse mammary glands of different developmental stages. Twenty micrograms of total RNA was loaded in each lane, and the blots were hybridized with gene-specific probes for Stat5a (*A*) and Stat5b (*B*) as well as for the WAP messages. Ethidium bromide-stained 18S rRNA is shown for quantitation. V, P, L, and I represent virgin, pregnant, lactating, and involution, respectively; and the numbers indicate the time (in days) at each stage when the mammary glands were taken.

regulation of Stat5 messages in mammary tissue throughout its development. The steady-state levels of Stat5a and Stat5b mRNAs increased during pregnancy, with the highest concentration on day 16 (Fig. 5), coincident with the onset of expression of the WAP (Fig. 5; ref. 7). However, both Stat5 messages gradually declined during lactation, a time period when the milk protein production remains high (Fig. 5; ref. 7).

DISCUSSION

Specificity and Redundancy in the Stat5 Pathway. Several STAT proteins appear to participate in multiple cytokine pathways (1–3) and, when activated, bind to similar or identical GAS sites (20). Stat5 has been linked to the PRL-signaling pathway (13), but little is known about its specificity. The activation of several cytokine pathways results in the phosphorylation of Stat5 and its concomitant ability to bind to GAS sites, but it does not necessarily parallel transcriptional activation. For example, Stat5 binding but not transcriptional activation of the β-casein gene promoter is induced by erythropoietin (EPO) in COS cells carrying expression vectors encoding the corresponding receptor (21). Stat5a and Stat5b show a 96% sequence conservation, suggesting that they may execute related or identical functions. However, their tissue distribution does not completely overlap (Stat5a is high in mammary gland, and Stat5b is high in muscle), and their binding and transcriptional activities exhibit some distinct features, suggesting that these molecules do not display redundant functions but define distinct signaling characteristics. Although the Stat5 proteins are probably necessary for the activation of milk protein genes during late pregnancy and early lactation, the presence of Stat5 mRNA in nonmammary tissues suggests an additional and less restrictive role.

Stat5 Proteins and the Activation of Milk Protein Genes. Mutations in the GAS sites of the promoters of rat WAP (10) and sheep β-lactoglobulin (8) genes resulted in an ≈80–90% reduction in expression of these transgenes during lactation. However, the tissue specificity was not affected, suggesting that

Stat5 proteins participate in the PRL-induced expression of milk protein genes but are not essential for their mammary specificity. Although Stat5 proteins can be found in many tissues of virgin and lactating mice, the expression of milk protein genes, and in particular the WAP gene, is confined to mammary tissue (22). This suggests that the GAS binding sites within the promoter regions are not recognized by STAT proteins in the context of native chromatin in nonmammary cells or that specific elements repress the WAP gene in nonmammary tissues.

Activation of the WAP gene in organ explant cultures from either virgin or pregnant mice requires the synergistic presence of PRL and glucocorticoids (6, 7), suggesting that neither signaling pathway by itself is sufficient to induce transcription. In the mouse and rat WAP genes, the target sequences for Stat5 are located at −570 and −725, respectively, and they are adjacent to binding sites for the glucocorticoid receptor and nuclear factor 1 (NF1) (9, 10). Since mutations of either the Stat5 or the NF1 site have dramatic effects on the level of WAP transgene expression at early lactation (10), it is likely that transcription is controlled by a synergistic interplay between transcription factors (9).

Stat5a- and Stat5b-specific mRNAs accumulate during pregnancy with similar kinetics, and their steady-state levels were highest on day 16 of pregnancy, followed by a decline during late pregnancy and lactation. By day 10 of lactation, the time of maximum milk protein gene expression, very little Stat5 mRNA was detected. However, the decline of the corresponding proteins is less pronounced, and they can be detected during midlactation (K. Heermeier, X.L., and L.H., unpublished data). The Stat expression pattern is in agreement with several transcriptional features of endogenous and transgenic WAP genes. The WAP gene is active in the virgin mouse during estrus (6, 23) and transcription can be induced with PRL (6), suggesting that the PRL signaling pathway can be activated prior to pregnancy. The peak of Stat5 RNA on day 16 of pregnancy coincides with the activation of the WAP gene (6, 7). The decline of Stat5 RNA during lactation, the period at which the WAP gene is expressed at its highest level, further strengthens and provides a mechanistic basis for our hypothesis that WAP gene induction and maintenance requires different signals (7). It is possible that the Stat5-mediated induction of WAP gene transcription during pregnancy is the result of an unmasking of additional transcription factor binding sites in the promoter. Once accessible, the role of Stat5 proteins in the maintenance of milk protein gene expression would be less pronounced. This is supported by our studies in transgenic mice. While activation of the endogenous WAP gene requires the presence of PRL, expression of WAP transgenes can be obtained in some lines in the absence of PRL (7, 24). It can be hypothesized that the initial stimulation during late pregnancy results in chromatin changes (11) that create an opportunity for other transcription factors to interact with WAP gene promoter sequences.

Note. After this work was submitted for review, two reports on murine Stat5 sequences were published (25, 26).

1. Darnell, J. E., Jr., Kerr, I. M. & Stark, G. R. (1994) *Science* **264,** 1415–1421.
2. Ihle, J. N. & Kerr, I. M. (1995) *Trends Genet.* **11,** 69–74.
3. Taniguchi, T. (1995) *Science* **268,** 251–255.
4. Topper, Y. J. & Freeman, C. S. (1980) *Physiol. Rev.* **60,** 1049–1056.
5. Groner, B., Altiok, S. & Meier, V. (1994) *Mol. Cell. Endocrinol.* **100,** 109–114.
6. Pittius, C. W., Sankaran, S., Topper, Y. & Hennighausen, L. (1988) *Mol. Endocrinol.* **2,** 1027–1032.
7. Burdon, T., Sankaran, L., Wall, R. J., Spencer, M. & Hennighausen, L. (1991) *J. Biol. Chem.* **266,** 6909–6914.
8. Burdon, T. G., Maitland, K. A., Clark, A. J., Wallace, R. & Watson, C. J. (1994) *Mol. Endocrinol.* **8,** 1528–1536.
9. McKnight, R. A., Spencer, M., Dittmer, J., Brady, J. N., Wall, R. J. & Hennighausen, L. (1995) *Mol. Endocrinol.* **9,** 717–724.
10. Li, S. & Rosen, J. M. (1995) *Mol. Cell. Biol.* **15,** 2063–2070.
11. Li, S. & Rosen, J. M. (1994) *Mol. Endocrinol.* **8,** 1328–1335.
12. Li, S. & Rosen, J. M. (1994) *J. Biol. Chem.* **269,** 14235–14243.
13. Wakao, H., Gouilleux, F. & Groner, B. (1994) *EMBO J.* **13,** 2182–2191.
14. Genetics Computer Group (1994) *Program Manual for the Wisconsin Package, Version 8* (Genetics Computer Group, Madison, WI).
15. Gouilleux, F., Wakao, H., Mundt, M. & Groner, B. (1994) *EMBO J.* **13,** 4361–4369.
16. Yang, Y.-C., Ciarletta, A. B., Temple, P. A., Chung, M. P., Kovacic, S., Witek-Giannotti, J. S., Leary, A. C., Kriz, R., Donahue, R. E., Wong, G. G. & Clark, S. C. (1986) *Cell* **47,** 3–10.
17. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162,** 156–159.
18. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1990) *Current Protocols in Molecular Biology* (Greene/Wiley Interscience, New York).
19. Hou, J., Schindler, U., Henzel, W. J., Ho, T. C., Brasseur, M. & McKnight, S. L. (1994) *Science* **265,** 1701–1706.
20. Standke, G. J., Meier, V. S. & Groner, B. (1994) *Mol. Endocrinol.* **8,** 469–477.
21. Gouilleux, F., Pallard, C., Dusanter, H., Wakao, H., Haldosen, L. A., Norstedt, G., Levy, D. & Groner, B. (1995) *EMBO J.* **14,** 2005–2013.
22. Pittius, C. W., Hennighausen, L., Lee, E., Westphal, H., Nichols, E., Vitale, J. & Gordon, K. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5874–5878.
23. Kordon, E. C., McKnight, R. A., Jhappan, C., Hennighausen, L., Merlino, G. & Smith, G. H. (1995) *Dev. Biol.* **168,** 47–61.
24. Shamay, A., Pursel, V. G., Wall, R. J. & Hennighausen, L. (1992) *Mol. Endocrinol.* **6,** 191–197.
25. Mui, A.-L.-F., Wakao, H., O'Farrell, A.-M., Harada, N. & Miyajima, A. (1995) *EMBO J.* **14,** 1166–1175.
26. Azam, M., Erdjument-Bromage, H., Kreider, B. L., Xia, M., Quelle, F., Basu, R., Saris, C., Tempst, P., Ihle, J. N. & Schindler, C. (1995) *EMBO J.* **14,** 1402–1411.