A New Regulatory Region of the IL-2 Locus That Confers Position-Independent Transgene Expression

Mary A. Yui, Gabriela Hernández-Hoyos and Ellen V. Rothenberg

*J Immunol* 2001; 166:1730-1739; doi: 10.4049/jimmunol.166.3.1730

http://www.jimmunol.org/content/166/3/1730

**References**

This article cites 42 articles, 14 of which you can access for free at:

http://www.jimmunol.org/content/166/3/1730.full#ref-list-1

**Subscription**

Information about subscribing to *The Journal of Immunology* is online at:

http://jimmunol.org/subscription

**Permissions**

Submit copyright permission requests at:

http://www.aai.org/About/Publications/JI/copyright.html

**Email Alerts**

Receive free email-alerts when new articles cite this article. Sign up at:

http://jimmunol.org/alerts
A New Regulatory Region of the IL-2 Locus That Confers Position-Independent Transgene Expression

Mary A. Yui, Gabriela Hernández-Hoyos, and Ellen V. Rothenberg

Although the promoter/enhancer of the IL-2 gene mediates inducible reporter gene expression in vitro, it cannot drive consistent expression in transgenic mice. The location and existence of any regulatory elements that could open the IL-2 locus in vivo have remained unknown, preventing analysis of IL-2 regulation in developmental contexts. In this study, we report the identification of such a regulatory region, marked by novel DNase-hypersensitive sites upstream of the murine IL-2 promoter in unstimulated and stimulated T cells. Inclusion of most of these sites in an 8.4-kb IL-2 promoter green fluorescent protein transgene gives locus control region-like activity. Expression is efficient, tissue specific, and position independent. This transgene is expressed not only in peripheral T cells, but also in immature thymocytes and thymocytes undergoing positive selection, in agreement with endogenous IL-2 expression. In contrast, a 2-kb promoter green fluorescent protein transgene, lacking the new hypersensitive sites, is expressed in only a few founder lines, and expression is dysregulated in CD8+ cells. Thus, the 6.4 kb of additional upstream IL-2 sequence contains regulatory elements that provide integration site independence and differential regulation of transgene expression in CD8 vs CD4 cells. The Journal of Immunology, 2001, 166: 1730–1739.

Interleukin-2 is the primary cytokine produced by newly activated naive T cells, and it plays a key role in the regulation of immune responses. IL-2 expression is stringently controlled, requiring multiple signals for induction, and is activated only in a subset of mature cells that express other activation markers (1). IL-2 is also induced in certain stages of thymic development in which its significance is still unknown (2–5).

Although the regulation of IL-2 expression in response to stimulation has been studied in great detail in transient transfection assays (reviewed in Refs. 6 and 7), we know very little about the mechanisms regulating IL-2 expression in a developmental context. Available cell lines are wholly inadequate to address this question by in vitro transfection studies because they do not represent normal developmental states. However, the well-characterized promoter/enhancer of the IL-2 gene is notoriously poor at allowing expression of reporter genes in transgenic mice. Use of the known promoter/enhancer (600 bp) to drive transgene expression resulted in only 1 of 17 and 2 of 26 founder mice expressing the transgene properly (8, 9). Use of a more extended regulatory region (up to 2700 bp), including all sequences tested for function by transfection (10), resulted in expression in only two of five founders (11). In several of the transgenic founders that express the transgenes, there is evidence for insertional appropriation of host regulatory sequences: i.e., the transgene gives ectopic expression, causes lethality when homozgyous, or both (8, 9, 11) (our unpublished results). In this context of variable expression, it has been impossible until now to attempt mapping sequence elements that restrict IL-2 expression to particular cell types.

Such a pattern of poor expression is characteristic of transgene constructs that are known to contain incomplete positive regulatory regions (12–14). The regulatory regions of these transgenes are missing sequence elements that can cooperate with proximal promoters/enhancers by opening the local chromatin and maintaining a transcriptionally competent domain. Some such elements are designated locus control regions (LCRs), genetic regulatory elements that confer tissue-specific and physiological levels of transcription on linked genes irrespective of integration site (15). It is only recently that the potential roles of elements affecting chromatin structure have been addressed for regulation of any cytokine genes (reviewed in Ref. 16). Distal regulatory elements have been identified by DNase hypersensitivity for the IL-3/GM-CSF locus (17, 18), and differential chromatin remodeling has been described for IFN-γ, IL-4, and IL-13 in Th1 vs Th2 cells (19). In addition, histone acetylation has been shown to be important for IL-4 inducibility in mature T cells (20). The addition of a heterologous (CD2) LCR to the 2.7-kb IL-2 promoter/enhancer region was recently shown to yield a high incidence of transgene expression, in 6 of 7 founders (21), which suggests that an LCR-like element is indeed missing from the IL-2 promoter regions tested to date. Constructs with exogenous LCRs, however, cannot be assumed to recapitulate all aspects of normal developmental regulation that may be mediated by different protein-DNA interactions. Therefore, in this study, we sought molecular and functional evidence to locate any natural LCR-like elements that control the normal developmental regulation of the IL-2 gene.

Regions of DNA-protein interactions are often hypersensitive to DNase I digestion, and this characteristic has proven to be useful in detecting new and distant transcriptional regulatory sites. An early study using this approach found evidence for hypersensitive (HS) sites beyond the minimal enhancer of the human IL-2 gene, although they were not precisely mapped (22). We therefore used

---

4 Abbreviations used in this paper: LCR, locus control region; CBSS, Coffman’s balanced salt solution; GFP, green fluorescent protein; HS, hypersensitive site; HSA, heat-stable Ag (CD24).
this method to find a series of new HS sites several kilobases upstream of the known IL-2 promoter/enhancer. The inclusion of these newly identified distal HS sites in green fluorescent protein (GFP) reporter gene constructs, along with the known IL-2 promoter/enhancer region, resulted in dramatically improved and consistent expression of GFP in independent transgenic founders over a wide range of transgene copy numbers. Expression was inducible and cell type specific in all expressing lines. This is the first study to identify a distal region of the IL-2 locus containing cis-acting elements that are sufficient for developmentally regulated expression of a transgene independent of the site of chromosomal insertion.

Materials and Methods

Cell culture and stimulations

Splenocytes and purified T cells were cultured in DMEM supplemented with 10% FBS, 2 mM l-glutamine, 5 × 10^{-5} M 2-ME, 100 U/ml penicillin, and 100 μg/ml streptomycin (Life Technologies, Gaithersburg, MD) at 37°C in 6% CO2. For stimulations via the TCR, 96-well flat-bottom plates (Corning, Corning, NY) were coated with 10 μg/ml purified cCD3 and cCD28 (PharMingen, San Diego, CA) in PBS for 2 h at 37°C and washed with medium before the addition of 10^6 cells. For chemical stimulations, 10^6 cells were incubated with 175 nM A23187 and 10 ng/ml PMA, both from Sigma (St. Louis, MO), and prepared as stock solutions in DMSO.

DNase I-hypersensitive assays

For purified T cells, C57BL/6 (B6) splenocytes were stained with biotin-conjugated anti-B220 Ab (PharMingen), incubated with streptavidin-conjugated MACs Microbeads, and passed over a V5 MACS magnetic separation column (Miltenyi Biotec, Auburn, CA) to remove B cells. Purified T cells were then subjected to DNase I treatment either immediately or after 6 h of stimulation with PMA and A23187. Synergistic melanoma MCa1/2 cells were used for a non-T-cell line (kindly provided by Michael Nishimura, University of Chicago, Chicago, IL).

DNase I HS assays were conducted, as described previously (23), using aliquots of 100 × 10^6 cells permeabilized by mlecithin for DNase digestion and DNA extraction. Twenty micrograms of each DNA sample were digested with XhoI, electrophoresed in agarose gels, blotted, and hybridized as described previously. Three probes were used as follows (shown in Fig. 2): the distal upstream region, the 842-bp XhoI to EcoRI fragment, the promoter region, the 784-bp EcoRI to XhoI fragment including part of the second intron and third exon, and for the downstream region, the 727-bp XhoI to EcoRI fragment in the third intron.

Transgene construction

DNA from the IL-2 upstream sequence was cloned from a C57BL/6 α genomic library (Stratagene, La Jolla, CA) using a probe specific for the promoter region of the IL-2 gene. Upstream sequences, from the XhoI site at −8.4 kb to the PstI site at +4.5 bp, and from the HindIII site at −2 kb to the PstI site, were subcloned into an enhanced GFP vector pEGFP-N1 (Clontech Laboratories, Palo Alto, CA). In addition, the SV40 poly(A) site in the vector was deleted from NotI to AflII, and replaced with the 3′ splice and poly(A) site from human β-globin, as the BamHI site was cloned to EcoRI fragment from plasmid pE84 (kindly provided by Stephen Hedrick, University of California, San Diego, CA).

The 8.4-kb upstream IL-2 region used in this study was sequenced by subcloning progressive unidirectional deletions using the Erase-A-Base system (Promega, Madison, WI). Dideoxy sequencing on the subclones was then conducted using Applied Biosystems Prism Dye Terminator Cycle Sequencing Ready Reaction and Applied Biosystems Prism 9600 Sequence Detector System kit (PE Applied Biosystems, Foster City, CA). These sequence data have been submitted to the GenBank database under accession number AF290391.

Generation and characterization of transgenic mice

Transgenic mice were generated in the Caltech Transgenic Animal Facility (Pasadena, CA) by pronuclear injection of (B6 × DBA/2)F1 zygotes using standard methods, and transgene positive mice were identified by PCR analysis of tail DNA. PCR primers in the IL-2 proximal promoter (IL2-1F: 5′-CACTGCTGCATTTGTGTAACCCCTC-3′ and the GFP coding sequence (GFP-1R: 5′-GCTGAATTCTGCGGCCTTTAC-3′) were used, amplifying a 830-bp product in transgene-positive mice. PCR conditions were as follows: 94°C, 5 min, then 35 cycles of 93°C, 30 s; 62°C, 15 s; 72°C, 45 s, followed by a final 5 min at 72°C, using an MJ PTC-200 DNA Engine Thermal Cycler (MJ Research, Watertown, MA). Founder mice were prop- agated by serial backcrosses to C57BL/6 mice. All conventional transgenic mice used for this study were generated and maintained in the Caltech Transgenic and Knockout Core Facility under specific pathogen-free conditions. Transgene copy number was determined by Southern blot hybridization of an IL-2 promoter region probe to genomic DNA cleaved with XbaI and EcoRI. The 600-bp probe was generated by amplifying a PCR product using the tail-typing primers, which was then digested with PstI to remove the GFP sequence for purification. The size of the transgenic and the 2.4-kb endogenous (two-copy) IL-2 bands were determined by use of a PhosphorImager (Molecular Dynamics, Sunnyvale, CA) and ImageQuant software.

To obtain transgenic SCID mice, transgene-positive male mice from two founder lines (lines 12 and 4) were crossed with C57BL/6-scid/scid (B6-scid) mice from the Rothenberg laboratory breeding colony at Caltech (24). To prevent infection of the immunodeficient animals, breeders and their pups were kept on prophylactic antibiotic (Baytril) treatment and transferred to sterile microisolator cages. Heterozygous offspring were typed for the transgene by PCR analysis and then backcrossed to B6-scid mice. Homozygous scid/scid pups were selected by flow cytometry of peripheral blood samples to determine the absence of peripheral lymphocytes, and transgene-positive animals were identified by PCR analysis of tail DNA. The transgene-positive SCID lines were then maintained on sterile food, water, and bedding, in sterile microisolator cages in ventilated racks.

Flow cytometric staining and cell sorting

Single-cell suspensions of 10^6 RBC-depleted blood, splenocytes, or thymocytes were stained on ice in 96-well U-bottom plates. Abs were diluted in Coffman’s balanced salt solution (CBSS) buffer (1.25% HBSS without phenol red) supplemented with 0.25% BSA (CBSS/BSA) (Pentax Fraction V; Miles, Kankakee, IL) with or without 0.1% sodium azide, as previously described (24). Cells were precubated in Fc block (PharMingen), followed by staining with specific Abs and washing with CBSS/BSA. The following staining reagents were used: CD3-APC, TRC8-Chrome, CD44-PE, CD69-PE, NK1.1-PE, Thy-1-PE, CD8-APC, HSA-biotin, streptavidin-APC (PharMingen), CD4-PE (Becton Dickinson, Mountain View, CA), and Sca-1-TC (Caltag Laboratories, South San Francisco, CA). FACs analysis was performed using the cell sorter (FACs Sflow Dickinson). Cells were sorted using a Coulter Elite (Coulter, Hialeah, FL).

RNA purification and real-time quantitative RT-PCR

Total RNA was extracted using RNAzol (Leedo Medical, Houston, TX) in accordance with the manufacturer’s instructions. Stimulated cells were incubated for 6 h with plate-bound αCD3/αCD28. RNA from 2 × 10^6 cells was treated with RNase-free DNase to remove residual genomic DNA. First-strand cDNA synthesis reactions were then conducted using SuperScript reverse transcriptase II (Life Technologies) and 250 ng random hexamers (Pharmacia, Uppsala, Sweden) following standard protocols.

Real-time fluorescent PCR analysis (25) was conducted using the Applied Biosystems Prism 7700 Sequence Detection System (PE Applied Biosystems). Taqman probes and primers for murine GAPDH (VIC labeled) and IL-2 (FAM labeled) were purchased as sets from PE Applied Biosystems, and a GFP-specific oligo probe, 6FAM-GCTTTACTTGTA4TAMRA, was custom synthesized by the same company. Primers used to specifically amplify the transgenic message crossed the intron between GFP (5′-CATGATGTTCTGCTGAGTTC3′) and the 3′ β-globin sequences (5′-CAGCACACAGACAGCACGGTT3′). Thirty-microliter reactions were conducted in Taqman Universal PCR buffer (1× Taqman buffer A; 5 mM MgCl2; 200 μM each of dATP, dTTP, and dGTP; 400 μM dUTP; 8% glycerol; 200 μM primers; 100 μM fluorescent-labeled probe; 0.01 U/μl AmpliTaq Gold) using the following thermocycling conditions: 50°C, 2 min; 95°C, 10 min; then 40 cycles of 95°C, 15 s and 60°C, 1 min. The cycle thresholds (Ct) were determined by measurement of the amount of fluorescent dye released after PCR amplification at each cycle using cDNA samples and specific primers and probes for GFP, IL-2, and GAPDH in separate amplification tubes and selecting a threshold in the linear range of the amplification. Control GAPDH Ct8 were subtracted from GFP and IL-2 Cts for each sample (= ΔCt). Each value was then adjusted by subtracting the ΔCt for a reference sample within each experiment (the transgene-negative control for IL-2 and transgenic line 4 for GFP) (= ΔΔCt). Relative amounts of mRNA were calculated as 2^-ΔΔCt.
Because protein-DNA contacts in the known 2700 bp of the IL-2 gene in unstimulated and stimulated purified T cells were found upstream, but not downstream of splenocytes, using a GFP-specific primer incorporating an XhoI site for cloning (5'-AGACTTAGACCCATAGACGACGACG-ACT-3'), and a reverse primer specific for exon 3 of the β-globin 3'end (5'-TGATAGGCCGCTTGACCTG-3'). The product was then digested with XhoI and EcoRI to generate a 562-bp product that was cloned into pBlueScript (Stratagene).

Results

DNase I-HS sites are found upstream, but not downstream of the IL-2 gene in unstimulated and stimulated purified T cells.

Because protein-DNA contacts in the known 2700 bp of IL-2 5'-flanking DNA are insufficient to mediate full LCR-like activity, we sought evidence for other regions of the gene that might confer position-independent expression on transgenes. To identify distal sites of potential protein-DNA interactions in the IL-2 locus, we performed DNase I-hypersensitivity assays on unstimulated and stimulated purified mouse splenic T cells. Unstimulated T cells displayed a prominent HS site at ~4.5 kb upstream of the IL-2 start site, with a series of five weaker HS sites seen between ~3 and ~4.5 kb (Fig. 1A, left panel). The same HS sites were also found in unstimulated EL4 thymoma cells that express IL-2 only after stimulation (data not shown). These sites appeared to be cell type specific; the region from ~3 to ~4.5 kb was not hypersensitive in a syngeneic melanoma cell line (Fig. 1A, right panel). These six T cell-specific HS sites are of particular interest as they indicate that the region of the IL-2 gene upstream of ~2 kb has an accessible chromatin structure in resting T cells and is bound by putative transcription factors that may play a role in maintaining the locus in an open but transcriptionally silent state.

Six hours after stimulation with PMA and the calcium ionophore A23187, extensive chromatin remodeling was seen in the upstream region (Fig. 1A, middle panel). Four new HS sites appeared between ~2.5 and ~3 kb. At the same time, the six HS sites found in the unstimulated cells were still present, although the ~4.5-kb site became less prominent. Three new, more distal sites also appeared, one strong band at approximately ~8 kb, and two others at ~7 and ~10 kb. Again, a similar pattern of HS sites was also found in stimulated EL4 cells (data not shown).

In sharp contrast, no HS sites were observed in the region from the third intron to 5 kb 3' of the IL-2 gene in unstimulated or stimulated purified T cells (Fig. 1B) or in EL4 cells (data not shown). HS sites in the promoter-proximal region that have been reported in human and mouse T cell lines (4, 22, 27) were also observed in purified T cells but not in non-T cells, and no new sites were revealed (Fig. 1C). A summary of IL-2 locus HS sites found in resting and activated T cells is shown in Fig. 2. Overall, these results show that the distal upstream region of the IL-2 gene is maintained in an open chromatin structure in resting peripheral T cells, but not in non-T cells, and that this region undergoes rapid and extensive remodeling upon activation. The presence of at least 13 new distal HS sites upstream of the IL-2 gene suggests that this region has many previously uncharacterized DNA-protein interactions that may play in vivo roles in mediating IL-2 locus accessibility or providing tissue-specific enhancer and/or silencer functions.

Construction and in vitro testing of the IL2p-GFP transgenes utilizing 8.4 kb vs 2 kb of upstream IL-2 sequence

To characterize and test this upstream region for function in vitro and in vivo, 5'-flanking sequence from the IL-2 gene was cloned from a CS7BL/6 genomic DNA library and sequenced. Two fragments, 2 and 8.4 kb, were subcloned into a GFP-expressing vector, to give constructs designated 2kbIL2p-GFP and 8kbIL2p-GFP, respectively. Both constructs include the proximal 2 kb of 5'-flanking sequence (to the PstI site at +45) that was found to drive optimal inducible expression in transiently transfected cells (10). The 8kbIL2p-GFP construct also includes an additional 6.4 kb of flanking DNA containing all of the novel HS sites found in resting T cells, as well as most, but not all, of the activation-induced HS sites (Fig. 2). The SV40 poly(A) in the original plasmid was replaced with the human β-globin 3'-splice site and poly(A) tail, which reproducibly improved the inducible expression of GFP by 2–2.5× after transient transfections in EL4 cells (data not shown).

The 2- and 8.4-kb GFP constructs gave similar GFP expression after transient transfections in EL4 cells (data not shown). The 8kbIL2p-GFP construct also includes an additional 6.4 kb of flanking DNA extending from within the third intron to 5 kb 3' of IL-2 coding sequence in stimulated or unstimulated T cells. C, HS sites in the region from ~2 kb in the promoter to +2 kb in the second intron. Two sites in the proximal promoter region have been reported previously in EL4 cells (23). An unlabeled arrow indicates a weak site of hypersensitivity in the second intron of the stimulated cells.

FIGURE 1. DNase I-HS sites in the IL-2 locus in unstimulated and stimulated purified splenic T cells. A, HS sites located from ~2.5 kb to >14 kb upstream of the IL-2 gene in unstimulated and stimulated purified T cells. No HS sites were detected in a melanoma line, MCA205 (non-T cells). HS sites are indicated by arrows. Locations, given in kb, are relative to the IL-2 gene start site. B, No HS sites were detected in the 6-kb XhoI fragment extending from within the third intron to 5 kb 3' of IL-2 coding sequence in stimulated or unstimulated T cells. C, HS sites in the region from ~2 kb in the promoter to +2 kb in the second intron. Two sites in the proximal promoter region have been reported previously in EL4 cells (23). An unlabeled arrow indicates a weak site of hypersensitivity in the second intron of the stimulated cells.
The 8kbIL2p-GFP-transgenic mouse lines express the transgene more consistently than 2kbIL2p-GFP lines

The two constructs, 2kbIL2p-GFP and 8kbIL2p-GFP, were injected into mouse embryo pronuclei, and transgenic founders were produced. Eleven independent 8kbIL2p-GFP and twelve independent 2kbIL2p-GFP founders were obtained. Founders were bred to C57BL/6 mice and more detailed analyses performed on cells from transgene-positive founders and/or their progeny. Transgene copy numbers ranged from 1 to >40 for both constructs, as determined by Southern blot analysis.

Transgene-positive mice were screened for activation-induced GFP expression among CD3+ cells. GFP expression was readily detected in CD3+-positive cells from both kinds of founders, as shown in Fig. 3A. Although spontaneous GFP expression was observed in some CD3+ cells (see below), the percentage and mean fluorescence values among those GFP-positive cells were given above each plot. Percentages of cells in each quadrant are indicated.

Transgenic line number and transgene copy numbers (in parentheses) are given above each plot. The percentages of GFP-positive cells among CD3+ cells were observed in 8kbIL2p-GFP transgenic lines: comparison with 2kbIL2p-GFP transgenic lines. A, Flow cytometric analysis of GFP expression in CD3 cells, from two 8kbIL2p-GFP and two 2kbIL2p-GFP lines, without stimulation (Unst) and 16 h after stimulation with PMA and A23187 (Stim). B, GFP expression levels in stimulated and unstimulated peripheral T cells from 11 independent 8kbIL2p-GFP and 12 independent 2kbIL2p-GFP transgenic lines. Splenocytes and/or peripheral blood cells (+) were stained with anti-CD3 Abs and analyzed by flow cytometry either without stimulation (open bars) or after 16 h of stimulation with PMA and A23187 (filled bars). Both the percentages of GFP-positive cells among CD3+ gated cells and the mean fluorescence values among those GFP+CD3+ cells are given for each transgenic line (+, missing data).

transgensics, with 10 of 11 (91%) lines expressing GFP in stimulated CD3+ cells (>3% positive under the conditions used in Fig. 3A). The only 8kbIL2p-GFP line that failed to express GFP carries only one copy of the transgene. By contrast, only 4 of 12 (33%) of the 2kbIL2p-GFP lines expressed GFP at that level; the remainder expressed the transgene very poorly or failed to express it at all over a wide range of copy numbers (Fig. 3B, left panels). The characteristics of the expression patterns seen in different 8kbIL2p-GFP lines were qualitatively very consistent between lines and among the progeny of most individual lines (see below, and M. A. Yui and E. V. Rothenberg, in preparation). The expression in these lines seemed unlikely to be a result of insertional mutagenesis, since there was no evidence of morbidity in transgene-positive animals. These results show that the additional 6.4 kb of IL-2 gene sequence from 1, A–C XbaI study. The 2kbIL2p-GFP and 8kbIL2p-GFP transgene constructs used in this study. The XbaI sites and the probes (gray boxes) used for HS assays (Fig. 1, A–C) are as shown. Features of an analysis of the upstream DNA sequence from ~2 kb to ~8.4 kb (Genbank AF290391) are also indicated: a putative matrix attachment region identified by use of MARFinder (43) (M), a region of high homology with an intronic sequence in the Bruton’s thymidine kinase gene (btk), and a region of high homology with an upstream sequence in the human IL-2 gene (A). To construct the transgenes used in this study, the 2-kb HinII-PstI and the 8.4-kb XhoI-PstI fragments were subcloned into a GFP reporter vector with a 1.7-kb BamHI-PstI human β-globin 3’-splice site and poly(A) (transgene constructs 2kbIL2p-GFP and 8kbIL2p-GFP, respectively). Upstream sequences included in these constructs are shown relative to the IL-2 gene and HS sites (X, XbaI; H, HinII; P, PstI).
upstream sequence in the 8kbIL2p-GFP transgene sharply increases the probability that an individual cell will open and express the locus, independent of integration site.

**GFP induction in purified GFP− αβ T cells**

Transgenic TCRαβ+ splenocytes that are initially GFP negative synthesize GFP de novo in response both to TCR ligation and to PMA/A23187 stimulation, as shown for two independent 8kbIL2p-GFP transgenic lines in Fig. 4. A subset of TCRαβ+ GFP− cells was seen to express GFP by 16 h after stimulation, and the percentage of GFP+ cells increased further by 44 h. GFP was readily induced using either plate-bound anti-CD3/anti-CD28 or PMA/A23187 stimulation, although a higher percentage of GFP+ cells was always observed in PMA/A2387-stimulated cells, with the mean fluorescence reaching peak levels at earlier time points. These results show that, as with endogenous IL-2, a subset of T cells that are initially GFP negative respond rapidly to specific activating signals by producing high levels of GFP.

**Quantitative analysis of transgene expression: similar efficiencies of expression of transgenes and endogenous IL-2 genes**

To assess whether the 8kbIL2p-GFP transgene includes regulatory elements allowing transgene expression at a level comparable with that of the endogenous IL-2 genes, GFP and IL-2 mRNA were quantitated using real time fluorescent RT-PCR (25) in samples from TCR/CD28-stimulated cells. RNA was prepared from splenocytes from nontransgenic and transgenic mice, with or without 6-h stimulation with anti-CD3/anti-CD28, and cDNA from these samples was analyzed in parallel with cDNA plasmid standards. Fig. 5A shows that with this assay, the threshold cycle number (Ct) (the PCR cycle at which the product of each sample crossed a detection threshold) was linear with the log of sample dilution over five orders of magnitude, for three independent sets of plasmid standard dilutions. These standards were used to calculate absolute levels of IL-2 and GFP mRNA in the cDNA from transgenic and control mice. As shown in Fig. 5B, the lower copy number transgenic lines (three to four copies) could express GFP mRNA at similar levels to the two-copy endogenous IL-2 genes, with ratios of about 1. Higher GFP:IL-2 ratios were seen at higher transgene copy numbers. This finding suggests that the 8.4 kb of upstream sequence in the 8kbIL2p-GFP transgene encompasses most, if not all, of the positive regulatory elements required for inducible expression, in addition to conferring positional independence.

In the 2kbIL2p-GFP transgenics, only a few lines (with 10–17 copies) show significant expression. Among these few 2-kb lines that do express well, the levels of GFP mRNA induced at 6 h spanned a 10-fold range that overlapped the expression levels seen in 8kbIL2p-GFP transgenic lines (Fig. 5C). Thus, the elements mapping between −2 and −8.4 kb affect the likelihood that transgenes will be expressed more than they affect the level of expression once induced.

**Effect of copy number on expression of 8kbIL2p-GFP transgenes in response to stimulation**

With the exception of line 175 (23 copies), the ratio of GFP:IL-2 expression showed only a modest increase over a 10-fold increase in copy number (Fig. 5B). Such a nonlinear response to copy number could be a result of competition for a limiting pool of trans-acting positive regulatory factors. However, GFP expression did not occur at the expense of endogenous IL-2 expression, even at high transgene copy numbers (Fig. 5C, left panel). In agreement with this result, we have found that at both low and high copy numbers, IL-2 and GFP protein can both be detected in individual stimulated T cells (data not shown).

GFP mRNA expression in the 8kbIL2p-GFP lines (Fig. 5C, right panel) generally paralleled protein expression (Fig. 3B), reflecting a composite of the percentage of T cells that becomes GFP+ and the mean fluorescence intensities of the GFP+ cells. The two lines with the lowest transgene copy numbers also displayed the lowest mean fluorescence intensities among GFP+ T cells (Fig. 3B, right panel). However, the percentage of cells expressing GFP did not show a consistent increase as a function of copy number (Fig. 3B, left panel). This suggests that the number of transgene copies (per integrated array) has only limited effect on the likelihood that these genes will be activated transcriptionally in response to a particular episode of stimulation.

**Spontaneous GFP+ cells have an effector/memory cell phenotype**

Although expression of the GFP transgenes is highly responsive to activation in the population as a whole, there is a reproducible background of transgene expression, at both the RNA and protein levels, before stimulation (Figs. 3B and 5C). This background expression is seen in all the expressing transgenic lines, regardless of whether they are derived with the 2-kb or the 8.4-kb promoter construct. Fig. 5C shows that background GFP RNA expression can be seen even in populations of unstimulated cells that express...
FIGURE 5. Quantitative analysis of GFP and IL-2 mRNA induction in 8kbIL2p-GFP and 2kbIL2p-GFP transgenic lines. A, Threshold cycle data (the PCR cycle number at which the specific amplification product crossed a threshold \( C_{T} \)) from real-time fluorescent RT-PCR analysis for IL-2 (left panel) and GFP (right panel) for mRNA quantitation in unstimulated and anti-CD3/anti-CD28-stimulated splenocyte samples using plasmid standard curves. Shown are C\( _{T} \) values for 10-fold serial dilutions of three independent plasmid preparations. Standard curves were adjusted for the difference in plasmid size, with the 3.65-kb pGmIL2 starting at 0.1 ng and the 3.46-kb pGFP-Bg at 0.95 ng/reaction (\( 10^{4} \) molecules). Also shown are the no-template controls (NTC), and cDNA from unstimulated (Unst) and stimulated (Stim) splenocytes (\( 10^{4} \) cell equivalents/reaction). Samples are from 8kbIL2p-GFP lines 4, 8, 17, 175, 177, 2kbIL2p-GFP line 214, and a transgene-negative (Tg\(^{-}\)) mouse. The maximum C\( _{T} \) value is 40 cycles. Standard curves for both IL-2 and GFP were linear over the entire dilution range used. Duplicate values are shown for each sample (dark circles and hollow triangles). B, Graph showing the ratios of GFP to IL-2 mRNA levels relative to transgene copy numbers for stimulated splenocytes from 8kbIL2p-GFP lines 4, 8, 17, 175, 177, 2kbIL2p-GFP line 214, and a transgene-negative (Tg\(^{-}\)) mouse. The maximum C\( _{T} \) value is 40 cycles. Standard curves for both IL-2 and GFP were linear over the entire dilution range used. Duplicate values are shown for each sample (dark circles and hollow triangles). C, Relative amounts of IL-2 (left panel) and GFP (right panel) mRNA determined by real-time fluorescent RT-PCR analysis, for RNA extracted from unstimulated (gray bars) and \( \alpha \)CD3/CD28 stimulated (filled bars) splenocytes. Transgene line and copy number (in parentheses) are given for each sample. C\( _{T} \) values were adjusted for control GAPDH levels and percentages of CD3\(^{+}\) cells in the input cell population. The relative values for IL-2 expression are presented on a scale with the average level in activated nontransgenic cells set as 1. The relative values for IL-2 are consistently in the more primitive thymocytes, in which expression cannot be determined by conventional TCR-dependent signals. However, the immature populations in which IL-2 mRNA is found are rare, altogether representing <0.1% of normal thymocytes (4). Therefore, to enrich these subsets ~100-fold, we crossed the 8kbIL2-GFP transgenes from two separate founder lines onto the B6-SCID homozygous background. GFP-expressing thymocytes from these mice could then be identified by four-color flow cytometry, using Sca-1 (Ly6-A/E) and HSA (CD24) expression for primary subdivision of cell types (reviewed in Ref. 28), as shown in Fig. 7. Both lines of SCID 8kbIL2p-GFP transgenics showed clear GFP expression that was restricted to the Sca-1\(^{-}\) HSA\(^{low}\) and Sca-1\(^{-}\) HSA\(^{high}\) populations. These two populations express IL-2 mRNA in vivo and are thought to represent pluripotent precursors and NK-like cells, respectively (28). GFP\(^{+}\) cells constituted a fraction of each of these subsets: 25–30% of the Sca-1\(^{-}\) HSA\(^{low}\) subset and no detectable IL-2 RNA (see data for lines 4, 8, and 214, one experiment each, and for lines 186, 177, and 17). The spontaneous transgene expression, however, is associated with the same cell types that express IL-2 in response to induction. As shown in Fig. 3A (and data not shown), spontaneously GFP\(^{+}\) cells are overwhelmingly concentrated in the CD3\(^{+}\) population, in both CD4\(^{+}\) and CD8\(^{+}\) subsets (see below). To determine whether this spontaneous expression of GFP is dependent on prior immunological activation, freshly isolated transgenic splenocytes were analyzed for surface phenotype markers that distinguish naive from previously activated cells. All of the GFP\(^{+}\) cells in the CD4\(^{+}\) population were found in the minority subset that is CD44\(^{hi}\) (Fig. 6), and predominantly CD62L\(^{lo}\) (data not shown), in every 8kb- and 2kbIL2p-GFP line tested. This phenotype is a hallmark of previously activated effector/memory cells. The majority of the GFP\(^{+}\) cells were found to be small, resting cells, and most were also CD69\(^{lo}\), indicating that only a minority of GFP\(^{+}\) cells were recently activated. The cells expressing GFP spontaneously in vivo can still respond to TCR stimulation in vitro by turning on even higher levels of GFP expression (data not shown), which is consistent with a memory cell phenotype. These results indicate that the background GFP expression, in peripheral CD4 cells at least, is primarily associated with effector/memory type T cells, presumably elicited by environmental stimulation. The 8kbIL2p-GFP transgene is also expressed very consistently in TCR\(^{+}\) and NKT cells, both of which are known to be activatable by endogenous Ags (data not shown and M.A.Yui and E.V.Rothenberg, in preparation).

Expression of the 8kbIL2p-GFP transgene in a developmental context: expression in primitive thymocyte subsets in SCID-transgenic mice

IL-2 is spontaneously induced both in TCR\(^{hi}\) thymocytes undergoing positive selection to CD4 and CD8 lineages (2–5) and in much more primitive cells before TCR gene rearrangement (4). All of our IL2-GFP-transgenic lines that express in activated peripheral T cells also show spontaneous expression in TCR\(^{hi}\) thymocytes (data not shown; and see below). It was of particular interest to determine whether the 8kbIL2p-GFP transgene is also expressed in the more primitive thymocytes, in which expression cannot be driven by conventional TCR-dependent signals. However, the immature populations in which IL-2 mRNA is found are rare, altogether representing <0.1% of normal thymocytes (4). Therefore, to enrich these subsets ~100-fold, we crossed the 8kbIL2-GFP transgenes from two separate founder lines onto the B6-SCID homzygous background. GFP-expressing thymocytes from these mice could then be identified by four-color flow cytometry, using Sca-1 (Ly6-A/E) and HSA (CD24) expression for primary subdivision of cell types (reviewed in Ref. 28), as shown in Fig. 7. Both lines of SCID 8kbIL2p-GFP transgenics showed clear GFP expression that was restricted to the Sca-1\(^{-}\) HSA\(^{low}\) and Sca-1\(^{-}\) HSA\(^{high}\) populations. These two populations express IL-2 mRNA in vivo and are thought to represent pluripotent precursors and NK-like cells, respectively (28). GFP\(^{+}\) cells constituted a fraction of each of these subsets: 25–30% of the Sca-1\(^{-}\) HSA\(^{low}\) subset and
FIGURE 6. Spontaneously GFP⁺ CD4⁺ T cells are CD44 bright, previously activated cells. Freshly isolated splenocytes from two 8kbIL2p-GFP, two 2kbIL2p-GFP, and transgene-negative lines were analyzed by flow-cytometric analysis for expression of GFP, CD4, and CD44, an effector/memory cell marker for cells within the CD4⁺ population. The correlation of GFP and CD44 fluorescence is shown in cells gated for CD4 expresion. Percentages of cells in each quadrant are given; GFP⁺ CD4⁺ cells were disproporionately CD44 bright in all transgenic lines tested (>85% CD44 bright; cf. <25% CD44 bright for the rest of the CD4⁺ population). Tranergenic line number and transgene copy numbers (in parentheses) are given above each plot.

35–40% of the Sca-1⁻ HSA low subset, in the example shown in Fig. 7. In each of these subsets, GFP⁺ cells were typical of the predominant cell type: the precursor-like Sca-1⁺ subset was Thy-1⁺ and NK1.1⁻ (Fig. 7A), while the NK-like Sca-1⁻ subset was Thy-1 heterogeneous and NK1.1⁺ (Fig. 7B). Most, but not all, of the GFP⁺ cells were also CD69⁺, in agreement with IL-2 mRNA expression data (4). In sharp contrast, the population including the majority of transgenic SCID thymocytes, the HSA⁺ subset, did not express the transgene (Fig. 7C). The HSA⁺ subset consists of cells undergoing specification and commitment to the T cell lineage, corresponding to the CD25⁺ CD44⁻⁻/⁻ pro-T cell populations in the conventional fractionation scheme (24, 28). Despite the more advanced developmental state of these cells, they do not express IL-2 mRNA in vivo (4). Results in SCID progeny of one of the 8kbIL2p-GFP transgenic line, line 12, were similar except that the percentage of cells expressing GFP in the Sca-1⁺ population was about 3-fold lower (data not shown). The expression of the transgene thus recapitulates faithfully the developmental expression pattern of the endogenous gene, even in a stage of lymphocyte development in which the inducing signal must be TCR independent.

Qualitative distortions in expression pattern of the 2kbIL2p-GFP transgene: preferential expression in CD8⁺ T cells

In those lines in which it is expressed, the 2kbIL2p-GFP transgene correctly directs expression to T cells and thymocytes, similar to the 8kbIL2p-GFP transgene, but the detailed pattern of transgene expression in T cell subsets differs. The most conspicuous difference is that the shorter transgene frequently drives high level expression in CD8⁺, but not CD4⁺ T cells. Whereas transgenic lines utilizing the 8.4-kb promoter construct showed similar low levels of background expression in CD4⁺ and CD8⁺ T cells, a distinctly biased pattern of endogenous expression of GFP was observed in the 2-kb transgenic lines, favoring the CD8⁺ cells (Fig. 8). This contrasts with the pattern of expression of endogenous IL-2, which is expressed at least as efficiently in CD4⁺ cells (29–31). Of the four 2kbIL2p-GFP-transgenic lines that express GFP, three express the transgene in percentages of splenic CD8⁺ cells 5–28 times higher than in CD4⁺ cells (Fig. 8C). This effect is most pronounced in 2kbIL2p-GFP line 227. By contrast, 8kbIL2p-GFP-transgenic lines express GFP in CD4⁺ and CD8⁺ cells at similar frequencies. Line 227 expresses GFP in 10 times more CD8⁺ cells than an 8kbIL2p-GFP line matched for spontaneous expression in CD4⁺ cells (Fig. 8A).

This preferential expression in CD8 cells is not an artifact of response to peripheral Ag exposure because it appears before the cells leave the thymus, in single-positive thymocytes emerging from positive selection (Fig. 8, B and D). In agreement with the transient expression of endogenous IL-2 at this stage (2–5), small percentages of CD4⁺ and CD8⁺ single-positive thymocytes express GFP, in all of the expression-positive transgenic lines with either the 8- or 2-kb construct (Fig. 8, B and D; and data not shown). However, the frequency of GFP expression in CD8⁺ cells relative to CD4⁺ cells is elevated in three of four 2-kb lines (Fig. 8D), and the percentage of GFP⁺ cells among CD8⁺ thymocytes is greatly expanded in line 227 (Fig. 8B).

These results suggest that IL-2 may be regulated in part through distinct mechanisms in CD4 vs CD8 cells. The additional 6.4 kb of upstream sequence appears to be needed to exert a negative regulatory effect to limit expression in CD8⁺, but not CD4⁺, cells.
Discussion

The well-characterized 600-bp IL-2 promoter/enhancer includes cis-acting elements required for inducible and T cell-specific expression in transient transfection studies. Nevertheless, constructs utilizing this promoter region consistently express transgenes very poorly, demonstrating that this region is missing key regulatory elements needed for proper gene expression in a developmental context. This is the first study directed to localize and functionally characterize these missing regulatory elements. We have now identified a 6.4-kb region upstream from the IL-2 promoter/enhancer, encompassing multiple HS sites in resting and activated T cells, which confers integration site-independent and developmentally consistent transgene expression. The ability to obtain reproducible, position-independent activity of a transgene utilizing IL-2 regulatory sequences makes it possible to map regulatory elements and to test the significance of IL-2 expression in distinct cell types and developmental states.

Inclusion of the novel upstream HS sites in a transgene confers strong, position-independent activity

The 8.4-kb flanking sequence transgene was designed to include a set of candidate regulatory sites that we have located by DNase-hypersensitivity analysis. The presence of at least 13 tissue-specific HS sites upstream of any previously identified regulatory sequences provided the first evidence that multiple, uncharacterized regulatory elements in this region may contribute to normal IL-2 gene regulation in vivo, although they do not appear to be important in transient transfection in vitro. Our transgenic studies confirm this prediction. A construct utilizing 2 kb of IL-2 upstream sequence results in transgene expression in only 4 of 12 independently derived lines, consistent with previous reports (11). In sharp contrast, when the new 8.4-kb transgene construct is used, GFP is expressed in 10 of 11 independent lines. Thus, the addition of 6.4 kb of upstream sequence, including all of the 5'-flanking HS sites found in resting T cells and most of the sites observed in stimulated T cells, greatly increases the likelihood of transgene expression, suggesting a relative independence from local integration site positional effects. By contrast, expression from the 2-kb promoter construct appears to depend acutely upon factors related to the site of integration.

Position-independent expression and resistance of transgenes to heterochromatization of high copy number arrays are both properties correlated with the presence of an LCR (13, 14, 32), or of an enhancer mediating LCR-like functions (33, 34).

Normally, sites of regulatory importance are phylogenetically conserved, e.g., the murine and human versions of the proximal 600-bp IL-2 promoter/enhancer, which are ~85% identical. This comparison cannot be made for most of the new upstream region, however, because the human sequence is interrupted at about 3300 by a >6-kb LINE element (35), and 5' sequence beyond this is currently unavailable. Only one significantly conserved region is seen between -2000 and this breakpoint, corresponding to -2206 to -2262 in the murine gene (27, 35). Transfection studies...
in a transformed cell line have not revealed any net enhancer activity in the new 6.4-kb upstream region (data not shown), but chromatin-opening functions need not be detectable in transient expression assays. In view of the activities reported for certain nuclear matrix attachment regions (12), it is of interest that the 8.4-kb upstream region of the murine IL-2 gene includes a potential matrix attachment region (see Fig. 2). Alternatively, the new sequences could exert their main effects through enhancer activity specific for particular, early developmental stages in the thymus when IL-2 is first expressed (cf Ref. 13).

**GFP transgene expression in response to stimulation in vivo**

Expression of the 8kbIL2p-GFP transgenes (and 2kbIL2p-GFP transgenes, in lines that allow expression at all) is cell type specific and inducible by TCR/CD28 ligation. All 8kbIL2p-GFP and 2kbILp-GFP transgenic lines that are capable of inducibly expressing the transgene also express GFP spontaneously in specific subsets of thymic and peripheral cells. In the thymus, these are the same cell populations that are known to express IL-2 in response to developmental signals (2–5), including two subsets of TCR-negative thymocytes, as well as mature TCRαβ and TCRγδ T cells (M. A. Yui and E. V. Rothenberg, manuscript in preparation). Thus, both transgenic constructs appear to be expressed in the compartments in which IL-2 is normally expressed.

IL-2 regulation is complex, and it will require further work to determine all the respects in which the transgenes may or may not reproduce the regulation of the endogenous genes. However, based on quantitative PCR analysis of mRNA from 8-kb transgenic splenocytes after stimulation, GFP mRNA appears to be expressed from three to four transgene copies at a comparable level to IL-2 splenocytes after stimulation, GFP mRNA is also likely to be more stable than IL-2 mRNA, as it is designed to lack the 3′ untranslated region destabilization motif found in the IL-2 gene (40, 41). This would exempt transgene expression from any negative regulatory mechanism in memory cells that acts at the level of IL-2 mRNA stability. In any case, the 8kbIL2p-GFP construct reveals a possible regulatory difference between naive and memory T cells, and offers a tool to dissect its mechanism.

**Cell specificity in the 8kbIL2p-GFP and 2kbIL2p-GFP transgenic lines: differential regulation in CD4+ and CD8+ cells**

In the few 2kbIL2p-GFP transgenic lines that do express the transgene, the inducibility of GFP and fidelity of transgene expression are largely similar to those of the 8kbIL2p-GFP transgenics. However, this work presents evidence of one regulatory defect shown by the 2kbIL2p-GFP lines even when they permit expression. Most 2kbIL2p-GFP lines have relatively high background levels of GFP in CD8+ cells in thymus as well as spleen, while expression in CD4+ cells in both tissues remains indistinguishable from that in the 8.4-kb lines. Although both CD4+ and CD8+ cells can express IL-2 normally (29–31), there is no evidence for any natural situation in which CD8+ cells express more IL-2 than CD4+ cells.

This suggests that there is a regulatory element or elements with distinct activities in CD8+ and CD4+ cells, which map(s) between 2 and 8.4 kb upstream of the IL-2 promoter. Alternatively, the CD8+ cell bias could be an artifact of the integration sites permissive for 2kbIL2p-GFP expression. In either case, these results offer the first evidence that a component of IL-2 regulation in CD8+ cells can be mediated by cis elements that are physically separable from those that regulate expression in CD4+ cells.

**A tool for analysis of IL-2 developmental regulation**

IL-2 is expressed in various developmental contexts that may ultimately make it most interesting as a marker for early hematopoietic lineage decisions (4, 5, 42). The signals driving the earliest intrathymic IL-2 expression are unknown, but the transcription factors mediating this programmed induction could be identified, once the critical cis-regulatory sequences are characterized. The 8kbIL2p-GFP transgene clearly includes these sequences and can be used as a starting point for this search. In addition, the ability to isolate primitive IL-2-expressing cells, without perturbation, based on strong, unambiguous GFP fluorescence, will make it possible to determine precisely their range of developmental potentials.

In conclusion, the results shown in this study locate element(s) conferring efficient, reproducible, position-independent expression in vivo that is separate from most of the elements in the IL-2 gene that mediate inducibility and T cell specificity. The identification of novel regulatory sequences between −8.4 and −2 kb makes it
possible to begin to dissect the molecular basis of in vivo developmental regulation of IL-2 for the first time. In addition, these new transgenic lines, which provide highly sensitive detection of individual live GFP/IL-2-expressing cells, should be a valuable tool to define the developmental significance of IL-2 induction and inducibility in lymphocyte subsets and lymphoid precursors.

Acknowledgments

We thank Stephen Hedrick for kindly providing the plasmid containing the human β-globin 3′ splicing and poly(A) addition sites and Michael Nishimura for the melanoma cell line MCA102. We gratefully acknowledge help from Robert Chen for screening the genomic library for the IL-2 promoter clone and to Xiao Sun for sequencing assistance. Shelley Diamond and Pat Koen of the Flow Cytometry Core Facility provided invaluable help with flow cytometry and cell sorting. We also thank Shirley Pease, Xin Yu, Bruce Kennedy, and Alba Granados of the Caltech Transgenic and Knockout Core Facility for the generation and maintenance of the transgenic mice. All primers were made at the Caltech Biopolymer Synthesis Facility, and sequencing was performed at the Caltech DNA Sequencing Core Facility.

References