B Lineage-Specific Interactions of an Immunoglobulin Enhancer with Cellular Factors in vivo

Anne Ephrussi, George M. Church, Susumu Tonegawa, Walter Gilbert


Stable URL:
http://links.jstor.org/sici?sici=0036-8075%2819850111%293%3A227%3A4683%3C134%3ABLIOAI%3E2.0.CO%3B2-I
leaving a trace, but by measuring all of the other particles in the event, the presence of neutrinos could be deduced by imbalances of energy and momentum. Next, Rubbia used multiple detection techniques along the paths of all particles so that electrons, photons, and muons could be identified with high reliability. These processes are also harbingers of new physics. Such considerations were essential because only about one W or Z particle was expected for every hundred million proton-antiproton interactions.

As in the case of the SPS-AA complex, UA-1 was constructed in an amazingly short time, and it began to take data in 1981. A second major detector known as UA-2 was also constructed during this period. It was built by a collaboration of approximately 60 physicists from six European institutions, under the leadership of Pierre Darrilul and Luigi DiLella, both of CERN. The UA-2 is somewhat simpler than the UA-1 apparatus, but it is used to deal with many of the same physics questions.

Both the UA-1 and UA-2 experiments resulted in a number of important observations during the early running. Most notable was the detection of very clean jets of hadrons that follow from the quark-gluon substructure of protons and antiprotons. By the end of 1982, enough data had accumulated for signs of the long sought W’s to be seen. The signature was quite clear; each group of investigators had a few events containing a very energetic electron produced at large angles with respect to the colliding beam direction, but with no other visible particles to balance the electron’s transverse momentum. In early 1983, both groups published their findings, which were fully consistent with predicted properties of the W boson. UA-1 had six events (6), whereas UA-2 presented four events (7). Soon thereafter, both groups obtained equally convincing evidence for the Z0. In this case, the production rates are expected to be lower than those for the W, but the experimental signatures are easier to interpret. By the end of 1983, Rubbia’s group had collected some 50 W events and five Z0 events, and the experimenters were using the data to learn more about the nature of the production process. Given the previous successes of the unified electro-weak theory, these results were immediately accepted as conclusive evidence for the existence of the W and Z bosons. This exploit closed the 50-year-old quest for a clarification of the weak force, but the technique that evolved opens a new phase of research in elementary particle physics. Henceforth, all new colliding beam accelerators and detectors will look back on the model pioneered by the 1984 Nobel laureates in physics.

**Leon M. Lederman**
**Roy F. Schwitter**

Leon M. Lederman is director of Fermi National Accelerator Laboratory, Batavia, Illinois 60510. Roy F. Schwitter is co-manager of the Collider Detector Facility at the Fermi National Accelerator Laboratory and professor of physics at Harvard University.

**References and Notes**


---

**RESEARCH ARTICLE**

**B Lineage–Specific Interactions of an Immunoglobulin Enhancer with Cellular Factors in Vivo**

Anne Ephrussi, George M. Church, Susumu Tonegawa, Walter Gilbert

Enhancers are DNA sequences that increase the level of transcription from some promoters when placed within several kilobases of them (1, 2). Mouse immunoglobulin heavy chain genes contain a tissue-specific enhancer, which is located in the intron separating the V region J segment and the constant region (Fig. 1) (3–5). The immunoglobulin heavy chain genes are extensively transcribed in myeloma cells. Cloned copies of these genes are also efficiently transcribed when transfected into cells of the B lineage but not in non-B cells such as L cells, HeLa cells, or lung cells (3, 4).

Analysis of deletions showed that removal of a portion of the intron of the immunoglobulin γ2c gene greatly reduces the level of transcription of the gene upon its transfection into myeloma cells (3). Further experiments revealed that restitution of a 990 base-pair fragment of DNA from the intron to these deletions restores high level transcription in myeloma cells and that this restoration is independent of both the position and the orientation of the fragment (4). A smaller fragment, 307 base pairs in length (Fig. 1, base pairs 376 to 683) is capable of enhancing transcription in myeloma cells from heterologous promoters and genes such as β-globin and SV40 T antigen (4).

How enhancers work is not known. Although viral transcription enhancers function in most cells regardless of species or tissue of origin, they are most active in cells that are the virus’s natural hosts (6–8). The immunoglobulin heavy chain enhancer has a somewhat more restricted tissue specificity and functions most efficiently in cells of the B lineage (3, 4). It seems likely that enhancer activity is related to and dependent on specific cellular factors. In the case of an enhancer showing tissue specificity one might expect to find proteins binding to the enhancer only in cells in which the enhancer is active.

Proteins that interact with specific sequences can prevent or increase the methylation by dimethyl sulfate (DMS).
of specific guanine (G) residues at the N-7 position in the major groove (9). We have used DMS as a chemical probe to observe the G residues in and around the heavy chain enhancer in the DNA of living cells of diverse origins. The DMS penetrates cells and methylates the accessible G residues. Genomic sequencing (10) can display the state of methylation of the G residues in extracted DNA. We compared the patterns of G reactivity to DMS in several different types of cells, searching for alterations in the intensity of specific G bands that might indicate that sequence-specific proteins were bound. In cells of the B lineage but not in cells of non-B lineage, the DNA of the immunoglobulin heavy chain enhancer exhibited protections and enhancements in its G pattern. The affected G residues were mainly in four clusters, in and around sequences homologous to 5’CAGGTGGC 3’ (C, cytosine; A, adenine; G, guanine; T, thymine). The alterations in the pattern of G reactivity are consistent with the tissue-specific binding of molecules to the mouse immunoglobulin heavy chain enhancer.

While guanine residues in single-stranded DNA are equally reactive to DMS, in double-stranded DNA they have distinctive reactivities. [An example is a method known as the systematic underreaction of the middle G in the sequence 5’T GGA 3′ in double-stranded DNA (12).] Moreover, in living cells, the double-stranded DNA may have proteins bound to it which could affect the reactivity of the G residues. The N-7 position of guanine that is methylated by DMS lies in the major groove; proteins that come close to this N-7 atom can prevent or increase methylation by DMS at that position (9, 13). To determine whether sequence-specific binding was occurring, we looked for cases in which a G was less reactive than it was in naked DNA (protected), equally reactive (no effect), or more reactive than in naked DNA (enhanced). If a protein is binding, we can expect to find several such protections and enhancements over a short stretch of bases.

Control experiments in which cells were not treated with DMS showed that the patterns observed were not due to the existence of cell-specific endonucleases or depurination; controls without piperin...
enzyme cleaved plasmid XF² (X18 poly-linker plasmid containing the 683-bp Xba I-Eco RI fragment) with 50 µg of sonicated salmon DNA, in 80 percent formamide, 10 mM tris-HCl, pH 8.3, 1 mM EDTA at 90°C for 30 seconds, adding DMS to a concentration of 0.5 percent, incubating at 20°C for 1.5 minutes, and precipitating with ethanol. The genomic sequencing protocol (10) was followed, with simplifications. The first two washings after hybridization and the sealing procedure for the last washing were omitted. The damp membranes were placed between taut layers of Saran Wrap for autoradiography. Exposure times were 2 days, with intensifier screens. For exposures without screens, the exposure time was 2 weeks. All DNA's were cut with Eco RI, cleaved with piperidine, and separated by size on a denaturing gel. In Fig. 3 (right), B lineage-specific protections and enhancements on the upper strand: 421 to 347. DNA's were cut with Dde I, cleaved with piperidine, separated by size on a denaturing gel, transferred to nylon, and probed with the Dde 3' upper probe. Naked DNA methylated with DMS in vitro (a) and a non-B lineage cell which was treated with DMS in vivo (b) have very similar patterns of reactivity of G residues with DMS. Their pattern is distinct from the pattern of DNA from a B-lineage cell which was also treated with DMS in vivo (c). Compared with (a) and (b), (c) has protected G residues at 421, 411, 406, 404, 399, 391, 387, 384, 381, 359, 356, 355, and 352; and G residues with enhanced reactivities at 419, 412, and 407. Since these residues are all in G clusters, it was necessary to observe these enhancements on shorter exposures, in order to resolve which band was indeed enhanced. The protection in (c) at G378 was not confirmed in other experiments with J558L nor was G378 protected in other cells of the B lineage. Similarly, G395 (b) was not protected in other experiments with MEL and with other non-B cells. G411 and G406 were slightly enhanced in (b) relative to naked DNA (occasionally in chromatin the first G in 5′GGA 3′ and 5′GGC 3′ sequences is enhanced compared with naked DNA (39)). The sequence of this region is shown in Fig. 5.
dine rule out the possibility that the patterns are due to tissue-specific DMS-activated nucleases.

We compared two categories of cells. The first consisted of immortalized cells of the B lineage that express their immunoglobulin genes, namely, two myelomas (J558L and MPC11), a B lymphoma (A20-2J), and an Abelson virus-transformed early B cell (RAW8). The second was composed of non-B cells: a Friend virus mouse erythroleukaemia line, (MEL) and two T lymphomas (S49 and BW5147). All of the B-lineage cells produce a form of immunoglobulin RNA (14–16). Little or no immunoglobulin RNA has been detected in the non-B cells (16, 17).

Changes in G reactivity are B lineage-specific. We examined nearly 1000 bases of DNA for alterations in reactivity of G residues with DMS and found 28 protections and four enhancements. These alterations are specific to cells of the B lineage and all are localized almost precisely to the region defined by transfection experiments as the enhancer (3, 4).

We first examined DNA from cells treated with DMS in vivo using the Eco 3' upper probe (Fig. 1). The sequence is most distinct beginning around the Hinf I site and extending 5' from it (Fig. 2). Cells of the B lineage show a G pattern of reactivity distinct from the pattern of non-B cells and of naked DNA. In B cells (Fig. 2, lanes b and g), G's at positions 545, 536, 507, 494, 421, and 411 are protected; G412 and G419 have enhanced reactivities to DMS. In some experiments G460 and G458 were protected. There is a strong enhancement at the 406-407 band, which is followed by a stretch of protected G residues extending to 347. There are no differences in relative G reactivity between DNA's from any cells and naked DNA 5' to 347.

The stretch from 407 to 347 was examined by means of Deel digested DNA and the Deel 3' upper probe (Fig. 1). This region is protected exclusively in cells of the B lineage. In those cells, the protected region begins with the strong enhancement at 407 (Fig. 3). The G at position 406 is protected as are residues 404, 399, 387, 384, 381, 359, 356, 355, and 352. In some experiments G391 was protected. G388 is a weakly reactive G (the middle G of a 5'GGC 3' sequence) and may, in some experiments, be protected. From G347 to G24, the pattern in B cells is indistinguishable from the pattern in non-B lineage cells or in naked DNA.

The patterns of B-lineage and non-B cells (with one possible exception) are identical at all positions 3' to 545 (data not shown). DNA digested with Pvu II and probed with the Pvu 5' upper probe reveals the G's between 480 and 750; a Hinf I digest probed with the Hinf 5' upper probe reveals those between 670 and 992 (Fig. 1). Over this entire 500-bp region all the differentiated cells examined had identical patterns of G reactivity to DMS in vivo.

The G at 708 is protected in all cells, with one exception. In one experiment, the embryonal carcinoma line PSA-1 (J8) was used and showed a pattern indistinguishable from that of naked DNA at all positions including a visible band at 708. The 708 G residue is not an intrinsically weak G residue: it appears as a strong band in all the naked DNA's. In RAW8, A20-2J, J558L, and MPC11 it is completely obliterated, as it is in BW5147 and MEL cells.

The same strategy was used to examine the lower strand in and around the mouse heavy chain enhancer (data not shown). Figure 4 shows results obtained with the Eco 5' lower probe (Fig. 1); the G residues at 534, 533, 531, and 523 are protected, and that at 530 is enhanced in cells of the B lineage. A second cluster of Gs is protected in these cells, at 413, 408, 401, 399, 358, and 357. No other changes appear in the stretch of DNA between 275 and 575.

Outside the enhancer region, we examined some additional 130 bp with the Eco 3' lower probe on Eco RI digested DNA's. No alterations in the G pattern appeared between 818 and 951 (data not shown). In all cases the whole cell patterns were identical and were indistinguishable from naked DNA controls.

Altered G reactivity occurs in conserved sequences within the enhancer. On both strands, more than 30 G residues are protected or enhanced in cells of the B lineage (Fig. 5) within a region of 970 bp on the top strand and 430 bp on the lower. These tissue-specific effects are all localized within a 193-bp stretch of DNA in the middle of the 992-bp region. Most of the protections and enhancements occur in four clusters located at sequences that share homology to the sequence 5'CAGTTGGC 3'. Three of the four clusters match this sequence at seven out of eight positions; the other matches at six out of eight positions (Fig. 6). Three of the clusters have the 5'CAGTTGGC 3' homologous sequence on the top strand; in the fourth cluster, it is found on the bottom strand (Fig. 5).

Within these sequences on the top strand, the G's at positions 3, 4, and 6 are protected; the G at position 7 is either protected, enhanced, or normal. On the lower strand G1 and G8 are protected. We suspect that these sequences may represent recognition sites for enhancer binding proteins. Given the similarity of these sequences, the binding proteins might be identical.

The reason residue G7, in the octamer, is protected in one case, normal in the second, and enhanced in the third and fourth, may be that the positioning of binding molecules is slightly different in each case because of variations in the DNA sequence. An amino acid residue in one position could protect a G from reacting with DMS while a slight shift in its position could expose the G and allow it to react and perhaps even be enhanced because of the proximity of the protein. This happens in the binding of phage lambda repressor to the operator site O41 (19), where the nearly symmetrical site O41 is divided into two half-sites, each of which is recognized by one subunit of the repressor dimer. Since the
center of symmetry passes through the central base pair, the repressor monomers do not bind with perfect symmetry; a G that is protected in one half-site, is in the other enhanced.

In a companion study, where we examined the heavy chain enhancer in vitro in nuclei (20), we compared B cell to L and Friend cell nuclei. We demonstrated tissue-specific contacts, in the enhancer region, to two sequences inverted with respect to one another and homologous to GCAAGGTGCC (sites 3 and 4). In the study presented here, we show (i) that the same contacts can be seen in living cells, and (ii) that in these in vivo experiments a further set of protections appears, mostly upstream of sites 3 and 4. The additional contacts to sites 1 and 2 could reflect the binding of the same factor to these more distant homologies or could reflect the binding of other tissue-specific factors to the region, a binding not reflected in vitro.

Some of the protections and enhancements that we see are outside the four octamer sequences. One protection (G545), which maps outside of these clusters, is located in the sequence TAATTTCGAT. This sequence is homologous to one of the two short sequences found upstream of all immunoglobulin variable region genes and within a region, as defined by transfection experiments, that is essential for correct transcription of those genes (21).

**Altered G reactivity correlates with enhancer function.** By treating whole cells in vivo with dimethyl sulfoxide and analyzing their DNA by genomic sequencing, we have shown tissue-specific contact of factors to the mouse immuno-

---

Fig. 5. Summary of the protections and enhancements. The complete sequence of the Xba I–Xba I fragment which contains the mouse immunoglobulin heavy chain enhancer is shown. Residues whose reactivities to DMS have been examined are indicated by capital letters; G residues 14 through 996 on the top strand and on the lower strand, G residues 275 through 574 and 818 through 951. A fragment spanning bases 376 to 683 has approximately full enhancer activity (4). B lineage–specific protections (○) and enhancements (●) are indicated above upper strand and below lower strand residues. Sequences homologous to those found upstream of all immunoglobulin variable region gene segments (21) are located at 451 to 465 and 539 to 548. The protection at 545 is within one of these sequences. A cross at 708 marks the G which is protected in all differentiated cells examined.
cells make no detectable immunoglobulin mRNA (16).

In most cells, the enhancer exists in two copies which, in the case of B-lineage cells, are often located in different regions of the genome. In J558L, for instance, only one copy of the enhancer is appropriately located within the large intron of a productively rearranged α gene. The other copy has been translocated from the immunoglobulin heavy chain locus to the c-myc locus on chromosome 15 (22). Since some of the protections in these cells are complete, it is likely that the reactivities of these G's in both copies of the immunoglobulin gene are being affected by the same factors in the same way.

In double-stranded DNA, as opposed to single-stranded DNA, the sequence 5′GGAGA 3′ consistently yields an under-reacted middle G in the reaction with DMS (12). There are five 5′GGAGA 3′ sequences distributed within the enhancer. The reactivities of the middle G's in all 5′GGAGA 3′ sequences in vivo are characteristic weak, like those of double-stranded naked DNA. Thus the immunoglobulin enhancer is for the most part double-stranded.

These patterns are probably not a simple consequence of transcription since the amount of transcription in the cell lines varies. A more likely interpretation is that these patterns are a consequence of binding of proteins that cause the enhancer activation.

There are other examples of activation of eukaryotic genes by binding proteins. In yeast, elements called upstream activator sequences (UAS) are regulated by trans-acting positive regulatory proteins and have some properties similar to those of enhancers (23). The glucocorticoid receptor, which can stimulate the transcription of mouse mammary tumor virus (24, 25), binds in vitro (26) to a fragment of DNA called the glucocorticoid response element (GRE) (27). The GRE, reminiscent of enhancers, has been shown to confer hormone inducibility to heterologous promoters in a distance and orientation-independent fashion.

These experiments demonstrate tissue-specific contacts, most likely made by a binding protein, that correlate with enhancer function. In general, the tissue-specific action of the enhancer could be due to factors positively acting on it in B cells or to negatively acting factors in non-B cells. Our findings support the first model.

References and Notes

Mutation in LDL Receptor: Alu-Alu Recombination Deletes Exons Encoding Transmembrane and Cytoplasmic Domains

Mark A. Lehman, Wolfgang J. Schneider, Thomas C. Südhof, Michael S. Brown, Joseph L. Goldstein, David W. Russell

Approximately one person in every 500 inherits one mutant gene for the low density lipoprotein (LDL) receptor, a cell surface protein that carries the cholesterol transport protein LDL from plasma into cells by receptor-mediated endocytosis in coated pits (1). When the receptors are defective, LDL cannot be removed normally from plasma, it accumulates to high levels, and premature atherosclerosis ensues. Mutations in the gene for the receptor are important clinically because a single copy produces heterozygous familial hypercholesterolemia (FH), a common cause of heart attacks in middle-aged people (J). About one in million persons inherits two copies of a mutant LDL receptor gene. In such individuals, called FH homozygotes, plasma levels of LDL-cholesterol are even higher than in heterozygotes, and heart attacks occur before 20 years of age (1).

Mutations in the LDL receptor are important subjects of study because they provide insights into the general mechanisms of receptor-mediated endocytosis (2). These mutations were originally defined through functional assays of LDL receptor activity in cultured fibroblasts and freshly isolated lymphocytes from FH patients (3). Subsequently, the LDL receptor was purified, and antibodies were generated. These advances led to the demonstration that some mutant alleles produce reduced amounts of receptor protein, whereas others encode receptor proteins with a demonstrably altered structure (4).

Among the most informative mutations have been those in which the LDL receptor reaches the cell surface but does not cluster in coated pits, a defect that prevents the receptors from carrying LDL into cells (5–7). Study of these mutations demonstrated the essential role of coated pits in receptor-mediated endocytosis (2, 5). As a result of kinetic and genetic complementation studies, we proposed that mutations that block internalization reside in the gene for the LDL receptor, and not in some other gene necessary for receptors to be incorporated into coated pits (5, 7).

Recently it has become possible to study the molecular genetics of these mutations as a result of the cloning of a full-length complementary DNA (cDNA) for the human LDL receptor (8, 9) and the subsequent isolation of genomic clones covering most of the LDL receptor gene, which spans at least 60 kilobases (kb) of DNA (10). The amino acid sequence of the LDL receptor was deduced from the nucleotide sequence of its cDNA (9). This sequence, coupled with protein chemistry studies (8), revealed that the normal receptor consists of at least five structural and functional domains. Listed in order from the amino terminus, the domains are (i) a cysteine-rich region of 322 amino acids that contains the binding site for LDL; (ii) a region of ~350 amino acids that is homologous to the precursor for epidermal growth factor; (iii) a serine and threonine-rich region of 48 amino acids that is the site for O-linked glycosylation; (iv) a stretch of 22 hydrophobic amino acids that spans the plasma membrane; and (v) a carboxyl terminal region of 50 amino acids that projects into the cytoplasm (9).

The mature messenger RNA (mRNA) for the receptor is unusual in that its 3′ untranslated region contains several copies of a repetitive DNA sequence of the Alu family (9).

We now report the use of our cDNA and genomic clones to characterize a mutation in the structural gene for the LDL receptor in a family with FH. The index case is a young man (B.H.), hereafter designated FH 274, who has all of the clinical features of homozygous FH (7). Previous functional studies revealed that cultured fibroblasts from FH 274 bound about one-third of the normal amount of 125I-labeled LDL. However, the receptors in FH 274 did not cluster in coated pits and hence did not transport their bound LDL into the cell (7). Thus, FH 274 was categorized functionally as an "internalization-defective" mutation.

Studies of fibroblasts from the relatives of FH 274 revealed that he had inherited two different mutant alleles; the allele encoding the internalization-defective