IDENTIFICATION OF A REPRESSOR IN THE FIRST INTRON OF THE HUMAN
α2(I) COLLAGEN GENE (COL1A2)*
Taras T. Antoniv‡§, Shizuko Tanaka‡§, Bayan Sudan‡, Sarah De Val¶, Ke Liu*, Lu Wang‡,
Dominic J. Wells*, George Bou-Gharios¶ and Francesco Ramirez‡ + **
§ Joint first authors
From the ‡Laboratory of Genetics and Organogenesis, Research Division of the Hospital for Special
Surgery, and Department of Physiology and Biophysics, Weill Medical College of Cornell
University, New York, NY 10021 (USA); *Gene Targeting Unit, Division of Neuroscience and
Mental Health, Imperial College School of Medicine, Charing Cross Hospital, London W6 8RP,
(UK); ¶ Renal Medicine, Imperial College School of Medicine, Hammersmith Campus, London W12
ONN (UK) and ‡CEINGE-Biotecnologie Avanzate, 80131 Naples (IT)
Running title: Intronic inhibitor of COL1A2 expression
** To whom correspondence should be addressed: Laboratory of Genetics and Organogenesis
Hospital for Special Surgery, 535 East 70th street, New York, NY 10021
Tel: 212-774-7554; Fax: 212-774-7864 E-mail: ramirezf@hss.edu

The human and mouse genes that code for the α2 chain of collagen I (COLIA2 and Calla2,
respectively) share a common chromatin structure and nearly identical proximal promoter
and far-upstream enhancer sequences. In spite of these homologies, species-specific differences have
been reported regarding the function of individual cis-acting elements, such as the first
intron sequence. In the present study, we have investigated the transcriptional contribution of
the unique open chromatin site in the first intron of COL1A2 using a transgenic mouse model.
DNase I footprinting identified a cluster of three distinct areas of nuclease protection (FI1-3) that
span from nucleotides +647 to +759, relative to the transcription start site, and which contain
consensus sequences for GATA and IRF transcription factors. Gel mobility shift and
chromatin immunoprecipitation assays corroborated this last finding by documenting
binding of GATA-4 and IRFs 1 and 2 to the first intron sequence. Moreover, a short sequence
encompassing the three footprints was found to inhibit expression of transgenic constructs
containing the COL1A2 proximal promoter and far-upstream enhancer in a position-independent
manner. Mutations inserted into each of the footprints restored transgenic expression to
different extents. These results therefore indicate that the unique open chromatin site of COL1A2
 corresponds to a repressor, the activity of which seems to be mediated by the concerted action of
GATA and IRF proteins. More generally, the study reiterates the existence of species-specific
difference in the regulatory networks of the mammalian α2(I) collagen coding genes.

The collagens represent a large family of extracellular proteins that impart specific physical
properties to tissues, in addition to playing important roles during morphogenesis and growth and in tissue
homeostasis and repair (1). Collagen I is the most abundant and most widely distributed collagen type,
with high prevalence in bone, skin, teeth, ligaments and tendons (2). It consists of two α1(I) chains and
one α2(I) chain which are produced by two fairly large genes that reside on different chromosomes in
both the human and mouse genome (3). Structural mutations in the collagen I chains give rise to
manifestations that affect the integrity of multiple organ systems in patients with osteogenesis
imperfecta and Ehlers-Danlos syndrome (3). Likewise, excessive deposition of collagen I
resulting from dysregulated expression of the corresponding genes (COL1A1 and COL1A2)* is the
hallmark of many fibrotic disorders that impair the function of affected organs (4–6). Expression of the
collagen I genes is tightly controlled during development and in a discrete subset of
mesenchymal cell types. That collagen I transcripts are found in the same 2:1 ratio as the corresponding chains has been interpreted to suggest that common regulatory programs coordinate expression of the two genes (7). However, multiple studies have failed to reveal a common organization of cis-acting elements and cognate trans-acting factors that would be consistent with the notion of shared regulatory programs between the collagen I genes (6). In point of fact, transgenic studies have revealed that the regulatory networks of the mammalian collagen I genes are organized very differently. On the one hand, tissue-specific production of α1(I) collagen chains is under the control of distinct and separate DNA elements scattered throughout the 3.2 kb immediately upstream of the start site of transcription (8-11). On the other hand, proper α2(I) collagen synthesis is the result of combinatorial interactions amongst nuclear factors that bind to overlapping DNA motifs clustered within the proximal promoter, as well as between them and those binding to a far-upstream enhancer (12-15).

Species-specific differences have also emerged with respect to the organization of the regulatory network of the human COL1A2 and mouse Col1a2 gene (12-15). Chromatin analyses have shown that COL1A2 and Col1a2 share five DNase I hypersensitive sites (HS) within nearly identical sequences of the proximal promoter (HS1) and 2.3 kb (HS2) and 20 kb upstream of from the transcription start site (HS3-5) (12, 13, 16). Furthermore, studies in transgenic mice have demonstrated that high and tissue-specific expression of both COL1A2 and Col1a2 proximal promoters requires interaction with the upstream sequence containing HSs 3-5, also know as the far-upstream enhancer (12,13). Deletion experiments have however shown that the region around HS5 is dispensable in the mouse, but absolutely required in the human transgene (13,14). Additional analyses have documented that the proximal promoter or the far-upstream enhancer of the human but not of the mouse gene, can by themselves drive transgenic expression in osteoblasts (15).

The transcriptional contribution of the first intron sequence is another potential difference between the two species. First, we have identified an open chromatin site, HS(In), that is unique to the first intron of the human gene (13). Second, earlier cell transfection experiments had assigned an enhancing activity to the first intron of Col1a2 and an inhibitory role to the COL1A2 counterpart (17,18). As part of our effort to delineate the full anatomy of the COL1A2 regulatory network, we have revisited these early studies using the transgenic mouse model in conjunction with DNA-binding assays and guided by the knowledge that the intronic sequence contains an open chromatin site (13). The results indicate that the sequence harboring HS(In) acts as a strong inhibitor of COL1A2 transcription, thus supporting the earlier contention of Sherwood et al. (17). Our investigations also mapped relevant cis-acting elements within the HS(In) sequence, identified the cognate trans-acting factors and demonstrated that full HS(In) repressing activity requires the concerted action of GATA and IRF transcription factors. This study therefore extends the characterization of the major functional elements of the COL1A2 regulatory network, in addition to identifying another species-specific difference between the human and mouse genes.

EXPERIMENTAL PROCEDURES

DNA Binding Assays- Nuclear extracts were purified from cultured WI-38 human lung fibroblasts, NIH-3T3 mouse fibroblasts or Jurkat T cells according to the previously published protocol (19). For DNase I footprinting assay, a plasmid DNA containing the intron 1 sequence that spans from nucleotides +524 to +895 was cleaved internally with Hind III, end-labeled by filling-in 3'-recessed ends with the Klenow enzyme, excised from the plasmid backbone with EcoRI , and incubated with nuclear extracts with or without addition of DNase I as previously described (19, 20). Likewise, the EMSA conditions for oligonucleotide end-labeling and incubation with nuclear extracts were essentially the same as previously described (19, 20). In some experiments, nuclear extracts were pre-incubated with commercial antibodies against GATA or IRF proteins (Santa Cruz Biotechnology, Santa Cruz, CA) or molar excesses of unlabelled mutant or wild-type oligonucleotides were added to the nuclear extract incubation. Footprinted sequences or DNA-protein complexes were resolved by polyacrylamide gel electrophoresis and visualized by autoradiography. The chromatin immunoprecipitation (ChIP) assay was performed on WI-38 cells using a commercial kit (Upstate, Lake Placid, NY) and according to the published protocol (21). Oligonucleotide primers corresponding to
+680/+705 and +1074/+1047 (intron-specific) or corresponding to -2472/-2450 and -2358/-2339 (negative control) were employed to PCR amplify sequence potentially bound by various nuclear proteins. The PCR reaction was performed for a total of 38 cycles after initial denaturation at 93°C for 3 min.; amplification conditions included denaturation at 93°C for 45 sec., annealing at 55°C (intron-specific) or 47°C (negative control) for 1 min. and elongation at 72°C for 2 min. One seventh of the immunoprecipitated genomic fragment was used as a template for amplification, except for the input sample in which 0.01% of the total DNA was used. Results of the ChIP analysis were visualized by Southern blot hybridization to the intron or upstream sequences of the amplification products separated on a 2% agarose gel. Intron sequences from different vertebrate organisms were derived from the Ensemble database (www.ensemble.org) and aligned using the program GeneDoc (www.psc.edu/biomed/genedoc).

RESULTS AND DISCUSSION

Early cell transfection experiments have shown that the first intron of Col1a2 and COL1A2 stimulates and inhibits transcription, respectively (17,18). These studies were however performed using long intronic sequences (1.2-1.7 kb) and without prior knowledge of the precise location of relevant cis-acting element(s). Subsequent analyses of chromatin structure located a unique DNase sensitive site, termed HS(In), at about +730 in the first intron of COL1A2 and within an evolutionarily divergent sequence (13,17). The present study was designed to characterize this putative regulatory element of COL1A2 using transgenic mice in combination with DNA-binding assays. Accordingly, the DNase I footprinting assays were first performed on a genomic fragment spanning from nucleotides +524 to +801 in order to map sites of nuclear protein-binding sites of the intronic sequence. Preparation of linearized plasmid DNA for microinjection was according to the standard protocol (22).

Transgenic Constructs- The control LacZ reporter constructs harboring the far-upstream enhancer and proximal promoter of COL1A2 had been already described (13). Mutant constructs were engineered using PCR amplification to insert single nucleotide substitutions into the various nuclear protein-binding sites of the intronic sequence. Preparation of linearized plasmid DNA for microinjection was according to the standard protocol (22).

Transgenic Embryos- Transgenic embryos were produced by the standard pronuclear injection of DNA into fertilized C57Bl/10 x CBA/Ca F1 eggs (22). Plasmid DNA was digested with appropriate enzymes, purified from agarose gel and microinjected at a concentration of 2-4 ng/ml in 10 mM Tris (pH 7.4) and 0.1 mM EDTA. Embryos were collected from the recipient females mainly at 15.5 days post coitum (E15.5) for whole-mount fixation and staining. This stage was chosen because it is characterized by high Col1a2 activity and to avoid decreased skin permeability due to increased keratinization (12). Southern blot hybridization and/or PCR amplification of placental DNA were used to assess transgene integration as previously described (12). After cutting open the thorax and abdomen, embryos were placed in cold phosphate-buffered saline and fixed for 45-60 min in 0.2% gluteraldehyde, 2% formalin in 0.1 M phosphate buffer ph 7.3 containing 2 mM MgCl2 and EGTA. After three washes of 1 hr each in the same buffer supplemented with 0.1% sodium deoxycholate and 0.2% Nonidet P-40, embryos were stained overnight at room temperature in 1 mg/ml of 5-bromo-4-chloro-3-indolyo-β-D-galactosidase solution (X-gal) containing 5 mM potassium ferrocyanide and 5 mM ferricyanide. For histology, X-gal positive embryos were dehydrated and wax-embedded, and 6-μm tissue sections were prepared, de-waxed and counterstained with eosin. The data presented are from embryos with comparable numbers of transgenic inserts, 2 to 5.
Cross-species sequence alignment of the relevant intron elements revealed only a modest level of sequence homology and loss of most of the GATA and IRF binding sites identified in the human gene (Fig. 1C).

Additional EMSAs using specific antibodies confirmed that GATA and IRF proteins indeed bind to the FI1 and FI2 sequences and to the FI3 sequence, respectively. Specifically, the assays showed that the FI1 and FI2 probes bind GATA-4 and not GATA-1, whereas probe FI3 recognized both IRF-2 and IRF-1 (Fig. 2). That the IRF-1 antibodies reduced FI3 complex formation without yielding a supershift could be accounted for by unspecific antibody interference. However, lack of IRF-1 antibody interference with Jurkat nuclear extracts excluded this possibility (Fig. 3A). Consistent with the differential distribution of the two GATA proteins (23), the specificity of the FI2 complex was indirectly corroborated by the finding that GATAs 2 and 3 or GATA-4 bind to FI2 in Jurkat T-cells and fibroblasts, respectively (Fig. 3A). The same results were obtained with the IF1 probe (data not shown). These in vitro binding assays were confirmed in vivo by ChIP analysis of human lung fibroblasts. Sequence-specific PCR amplification of genomic DNA immunoprecipitated with antibodies against GATA-4, IRF-1 or IRF-2, but not with unspecific antibodies, yielded reproducibly positive signals when hybridized to the HS(In) probe (Fig. 3B). Furthermore, specificity of in vivo binding was independently confirmed by lack of positive signals in a parallel control sample in which the ChIP assay was performed with an upstream sequence of COL1A2 (Fig. 3B). Preliminary evidence also suggests that a potential CREB/AP1 recognition sequence in FI2 binds c-Jun (data not shown).

Having determined the precise location of the HS(In) element and the identity of the interacting nuclear factors, the next experiments examined its functional contribution to COL1A2 transcription. Transient transfections using a 372 bp intronic sequence (+524 to +895) inclusive of the HS(In) element showed a slight downregulation of the –378 sequence (+524 to +895) inclusive of the HS(In) element showed a slight downregulation of the –378 sequence (+524 to +895) inclusive of the HS(In) element within the HS(In) probe (data not shown) (17). Based on this preliminary evidence, the transgenic model was then employed to examine the HS(In) element within the in vivo context and in relationship to the interaction between the proximal promoter and far-upstream enhancer. Transgenic constructs included the original 21.1/18.8pLAC plasmid, which contains the core sequence of the far-upstream enhancer and the -378 proximal promoter (13), and a modification of 21.1/18.8pLAC in which the wild-type 372 bp segment containing the HS(In) element had been inserted downstream of the reporter gene (21.1/18.8pLAC-In), or between the far-upstream enhancer and proximal promoter (21.1/18.8(In)pLAC) (Fig. 4). Several 21.1/18.8pLAC-In founders were generated and all showed lower β-galactosidase staining in most tissues compared with embryos harboring the intronless 21.1/18.8pLAC transgene (Fig. 5A and B). Histological sections demonstrated that β-galactosidase staining in different transgenics, albeit variable in intensity, was always confined to collagen I-producing cells (Fig. 5D-I). Overall, the intensity and distribution of X-gal staining in the 21.1/18.8pLAC-In transgene was reminiscent of the pattern previously observed with the proximal promoter transgene without the far-upstream enhancer (13). Limited staining was seen in some ossification centers of intramembraneous bones (Fig. 5H and I), and in patches of skin fascia and tendon (Fig. 5G) as well as in a few internal organs, such as the forming kidney and spleen (Fig. 5D and F). By contrast, no staining was detected in the lung, heart, gut and blood vessels (Fig. 5E). Similar results were obtained with 21.1/18.8(In)pLAC, the construct in which the intronic sequence had been inserted between the enhancer and the promoter (Fig. 5C). Together, these findings demonstrated the inhibitory effect of the HS(In) sequence on COL1A2 transcription.

Based on the above findings, we assessed the contribution of individual nuclear protein-binding sites to HS(in) inhibition of 21.1/18.8pLAC expression by examining β-galactosidase activity in transgenic mouse embryos harboring mutated versions of FI1, FI2 or FI3. The mutations included nucleotide substitutions in the binding sites of each footprint (21.1/18.8pLAC-Inm1, 21.1/18.8pLAC-Inm2 and 21.1/18.8pLAC-Inm3) and in the GATA-binding sites of both FI1 and FI2 (21.1/18.8pLAC-Inm1,2) (Fig. 4). X-gal staining revealed comparable β-galactosidase levels that were similar to that of the intronless 21.1/18.8pLAC transgene (Fig. 6A-D). They also identified a few interesting differences amongst the four mutant transgenes. First, X-gal staining in the skin and other tissues of the
21.1/18.8pLAC-Inm1 and 21.1/18.8pLAC-Inm2 transgene was lower than in 21.1/18.8pLAC-Inm3 embryos (Fig. 6A, B and D). Sole exception was the unusually strong β-galactosidase activity in 21.1/18.8pLAC-Inm1 and 21.1/18.8pLAC-Inm2 bones (Fig. 6K and L). In point of fact, this was the strongest X-gal staining ever recorded in bone with any of the COL1A2 constructs examined in this and previous studies (13, 15). Second, LacZ gene expression in internal organs of the 21.1/18.8pLAC-Inm2 transgene was consistently higher than the 21.1/18.8pLAC-Inm1 construct (data not shown). Third, the combination of both GATA mutations (21.1/18.8pLAC-Inm1,2) yielded the same level of X-gal staining as the mutation of only the IRF-binding element is a strong repressor of COL1A2 transcription in vivo. This conclusion was based on the ability of a short HS(In) containing sequence to inhibit the activity of an experimental model that closely replicates the expression pattern of COL1A2 in transgenic mice. Within the limitations of this in vivo model, our result supports Sherwood et al. (17) cell transfection data and reiterates the existence of functional differences in the organization of the regulatory network of the Col1a2 and COL1A2 genes. As such, it underscores the peril of extrapolating functional conclusions from one mammalian species to another.

The EMSA and the ChIP assay have correlated the inhibitory activity of the HS(In) sequence with the specific binding of GATA and IRF proteins to three nearly juxtaposed elements. Moreover, transgenic experiments have demonstrated that inhibition requires the full complement of nuclear protein binding sites. They have also raised the possibility that each of the cis-acting HS(In) elements imparts slightly different properties to the inhibitory protein complex. GATA proteins represent a small family of zinc finger transcriptional regulators that are expressed in hematopoietic stem cells (GATAs 1, 2 and 3) and in a variety of mesoderm and endoderm-derived tissues (GATAs 4, 5 and 6) (23). Consistent with the tissue distribution of GATA family members, we observed binding of GATA-4 and GATAs 2 and 3 with nuclear extracts from NIH-3T3 and Jurkat cells, respectively. GATA proteins have been reported to modulate tissue-specific gene expression across various cell types by interacting with a large array of transcriptional activators and repressors (23). Along these lines, we recently found that the HS2 element of COL1A2 is another GATA-binding site that represses transcription from the -378 promoter (24). It is therefore conceivable to argue that combinatorial interactions among GATA proteins and co-factors at different COL1A2 sites may orchestrate expression of this mesenchyme-specific gene within different cellular contexts.

Originally identified as transcriptional repressors or activators of interferon-β (IFN-β) and of INFγ-inducible genes, IRFs have later emerged as broader regulators of other biological processes, such as cell growth, in conjunction with other nuclear proteins (25). A case in point is IRF-2 which has been shown to repress IFN-β activation by coactivator repulsion (26). In this novel regulatory mechanism, incorporation of IRF-2 into the enhanceosome prevents recruitment of the CBP-polII holoenzyme complex through a specific protein domain. Similar to the relatively higher activity of 21.1/18.8pLAC-Inm3 compared with the other mutant transgenes in most collagen I-producing tissues, inactivation of IRF-2 in mice has been shown to expand the number of cells that respond to viral infection by inducing INF-β gene transcription (26). Although our study did not address whether a similar mechanism may operate in COL1A2, it nonetheless suggested that other factors which normally counteract HS(In) repression in collagen I-producing cells are not present in the transgenic model used as our experimental read-out. Amongst others, probable modulating factors include additional sequences and cognate trans-acting factors and/or appropriate spacing of the interacting domains in the regulatory network (27). Work in progress is characterizing the
precise mechanism of the transcriptional repression by the intronic elements and the identity of the trans-acting factors involved in this process. It is also exploring the possibility that HS(In) may contain another interferon-responsive element of the COL1A2 gene.

REFERENCES

1. Hay E.D. (1995) in Cell Biology of the Extracellular Matrix (Hay E.D., ed.) 2nd Ed., pp. 419-462, Plenum Press, New York, N.Y.
2. Van der Rest M., and Garrone R. (1991) FASEB J. 5, 2814-2823
3. Myllyharju J., and Kivirikko K.I. (2001) Annu. Med. 33, 7-21
4. Varga, J. and Jimenez, S.A. (1995). Ann. Intern Med. 122,60-62.
5. Trojanowska, M. (2002). Front. Biosci. 7,d608-d618.
6. Ghosh, A.K. (2002). Exp. Biol. Med. 227,301-314.
7. Vuust, J., Sobel, M.E., Martin, G.R. (1985) Eur. J. Biochem. 151, 449-453
8. Rossett, J., Eberspaecher, H., de Crombrugghe, B. (1995) J.Cell Biol. 129, 1421-1432
9. Rossett, J.A., Chen, S.S., Eberspaecher, H., Smith, C.N., de Crombrugghe, B. (1996) Proc. Natl. Acad. Sci. U S A 93, 1027-1031.
10. Terraz, C., Toman, D., Delauche, M., Ronco, P., Rossert, J. (2001) J. Biol. Chem. 276, 37011-37019.
11. Terraz, C., Brideau, G., Ronco, P., Rossert, J. (2002) J Biol. Chem. 277, 19019-19026.
12. Bou-Gharios, G., Garrett, L.A., Rossert, J., Niederreither, K., Eberspaecher, H., Smith, C., Black, C. and de Crombrugghe, B. (1996) J. Cell. Biol. 134, 1333-1344.
13. Antoniv, T.T., DeVal, S., Wells, D., Denton, C.P., Rabe, C., de Crombrugghe, B., Ramirez, F. and Bou-Gharios, G. (2001) J. Biol. Chem. 276, 21754-21764.
14. De Val, S., Ponticos, M., Antoniv, T.T., Wells, D.J., Abraham, D., Partridge, T., Bou-Gharios, G. (2002) J. Biol. Chem. 277, 9286-9292.
15. Tanaka, S., Antoniv, T.T., Lui, K., Wang, L., Wells, D.J., Ramirez, F. and Bou-Gharios, G. (2004) *J. Biol. Chem.* **279**, 56024-56031.

16. Dickson, L., de Wet, W.J., Di Liberto, M., Weil, D. and Ramirez, F. (1985) *Nucleic Acids Res.* **13**, 3427-3438.

17. Sherwood, A.L., Bottenus, R.E., Martzen, M.R., Bornstein, P. (1990) *Gene* **89**, 239-244.

18. Rossi, P., de Crombrugghe, B. (1987) *Proc. Natl. Acad. Sci. U S A* **84**, 5590-5594.

19. Truter, S., Di Liberto, M., Inagaki, Y., Ramirez, F. (1992) *J. Biol. Chem.* **267**, 25389-25395.

20. Inagaki, Y., Truter, S., and Ramirez, F. (1994) *J. Biol. Chem.* **269**, 14828-14834.

21. Smaldone, S., Laub, F., Else, C., Dragomir, C. and Ramirez, F. (2004) *Mol. Cell. Biol.* **24**, 1058-1069.

22. Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. (2003) in *Manipulating the Mouse Embryo: A Laboratory Manual*, 3rd Ed., pp 289-358. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

23. Molkentin, J.D. (2000) *J. Biol. Chem.* **275**, 38949-38952.

24. Wang, L., Shizuko, T. and Ramirez, F. (2005) GATA-4 binds to an upstream element of the human α2(I) collagen gene (COL1A2) and inhibits transcription in fibroblasts. *Matrix Biology (In press).*

25. Mamane, Y., Heylbroeck, C., Genin, P., Algarte, M., Servant, M.J., LePage, C., DeLuca, C., Kwon, H., Lin, R., Hiscott, J. (1999) *Gene* **237**, 1-14.

26. Senger, K., Merika, M., Agalioti, T., Yie, J., Escalante, C.R., Chen, G., Aggarwal, A.K., Thanos, D. (2000) *Mol. Cell.* **6**, 931-937.

27. Ogbourne, S., Antalis, T.M. (1998) *Biochem. J.* **331**, 1-14.
FOOTNOTES

We thank Karen Johnson for organizing the manuscript and Graham Reed for the photographic reproductions.

*This work was supported by National Institutes of Health Grant AR386481, the St. Giles Foundation, the James D. Farley Family and by the Arthritis Research Campaign. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program Grant Number C06-RR12538-01 from the National Center for Research Resources, National Institutes of Health.

1The abbreviations used are: COL1A2 and Col1a2, human and mouse α2(I) collagen gene, respectively; EMSA, electrophoretic mobility shift assay; HS, DNase I hypersensitive site; IRF, interferon regulatory factor.

FIGURE LEGENDS

Fig. 1. DNase I footprinting of the COL1A2 intron. A, an end-labeled genomic fragment spanning from nucleotides +542 to +801 was incubated with increasing amounts of DNase I in the presence (+) or absence (-) of WI-38 nuclear extracts. The numbers on the side of the protected areas correspond to the footprinted sequences (FI1-3) shown in panel B, where the locations of the GATA and IRF consensus sequences are also indicated along with the nucleotide substitutions (underlined) of the mutant sequences used in the transgenic experiments and EMSAs. C, Cross-species sequence alignment of the relevant intronic elements shown in panel B and analyzed in Fig. 2.

Fig. 2. Gel mobility shift assays of the HS(In) footprints. Labeled oligonucleotides corresponding to the sequences of FI1, FI2 and FI3 were incubated with nuclear extracts purified from NIH-3T3 with or without molar excesses (50 to 100 fold) of unlabeled oligonucleotides corresponding to wild-type (wt) and mutant (mt) versions of the original probes, and the wild-type or mutant GATA (Gwt and Gmt) and IRF (Iwt and Imt) consensus sequences, as well as pre-incubation with antibodies (α) against the indicated transcription factors. Symbols indicate retarded (arrow) or supershifted (arrowhead) complexes, and non-specific bands (asterisk).

Fig. 3. GATA and IRF proteins binding to HS(In) elements. A, antibody interference of FI2 and FI3 complex formation using the indicated antisera (α) and nuclear extracts purified from fibroblasts (F) or Jurkat cells (T). B, ChIP analysis of chromatin from WI-38 cells before (lane 1 and 2) and after immunoprecipitation (in duplicate) with antibodies against IRF-1 (lanes 3 and 4), IRF-2 (lanes 5 and 6) and GATA-4 (lanes 7 and 8). Other controls include IgG treated sample (lane 9) and input DNA (lane 10). Top (intron), PCR amplified products of the +680 to 1074 intron segment bearing the GATA and IRF binding sites. Bottom (control), PCR amplified products of the -2472 to -2339

upstream sequence of COL1A2. Amplified products were electrophoresed in a 2% agarose gel and visualized by Southern blot hybridization to the original intron or upstream probes.

Fig. 4. Transgenic constructs. Schematic representation of the LacZ reporter constructs employed in the transgenic analyses and containing the intronic elements shown in Fig. 1B.

Fig. 5 Functional analysis of HS(In) in transgenic mice. Whole-mount X-gal staining of illustrative E15.5 embryos harboring the 21.8/18.8pLAC (A), 21.1/18.8pLAC-In (B) or 21.1/18.8(In)pLAC transgenes (C). X-gal staining of tissue sections from 21.1/18.8pLAC-In transgenic embryos showing β-galactosidase activity in a few collagen I-producing cells, such as those in the splenic premordium (D, s) and around the vertebra but not in blood vessels (E, s). Positive mesenchymal cells of the developing kidney and tendons are shown in F and G, respectively. A few osteoblasts are positively stained (arrows) in the parietal bone (H) and maxillary gland (I).

Fig. 6. Effects of HS(In) mutations on transgene activity. Whole-mount X-gal staining of illustrative E15.5 embryos harboring the 21.1/18.8pLAC-In\textsuperscript{m1} (A), 21.1/18.8pLAC-In\textsuperscript{m2} (B), 21.1/18.8pLAC-In\textsuperscript{m1,2} (C) and 21.1/18.8pLAC-In\textsuperscript{m3} (D) transgene. Tissue sections from embryos representative of mutations m1, m2 and m3 (E-J) or m1 and m2 (K, L) showing β-galactosidase activity in skin fibroblasts (E) but not keratinocytes (arrow), gut muscular layers (F, m), pancreas (G), blood vessels (H, arrow), tendon (H, t), lung fibroblasts (I), and splenic premordium (J, s) but not kidney (k). Intense staining was also seen in all osteoblasts, including those in growth plates (K, arrows) and ribs (L, arrows).
Fig. 1
Figure 3
Fig. 4
Identification of a repressor in the first intron of the human α2(I) collagen gene (COL1A2)
Taras T Antoniv, Shizuko Tanaka, Bayan Sudan, Sarah De Val, Ke Liu, Lu Wang, Dominic J. Wells, George Bou-Gharios and Francesco Ramirez

J. Biol. Chem. published online August 8, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M502681200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts