RESEARCH COMMUNICATION

Cloning of a functional promoter of the human sodium/iodide-symporter gene

Michael BEHR, Thomas L. SCHMITT, Celia R. ESPINOZA and Ulrich LOOS¹

Abteilung Innere Medizin I, University of Ulm, Universitätsklinikum Ulm, Robert-Koch-Strasse 8, 89081 Ulm, Federal Republic of Germany

We have cloned and sequenced genomic DNA from a human library extending 1300 bp upstream the 5'-untranslated sequence of the cDNA coding for the sodium/iodide symporter. In transient transfection assays this sequence exhibited promoter activity, which could be confined to nucleotides –443 to –395

relative to the ATG start codon. This minimal promoter, including a putative GC- and TATA- box, was preferentially activated in the rat thyroid cell line FRTL-5, but was also active in non-thyroidal cells, such as COS-7 and Chinese-hamster ovary, albeit to a markedly lower extent.

INTRODUCTION

The uptake of iodide into the thyroid gland is a prerequisite for the synthesis of thyroxine. In addition, this process is made use of in radioiodine therapy for thyroid carcinomas. Iodide influx into the thyroid cell is actively accomplished by the sodium/iodide symporter (NIS), a protein for which the genes of the rat and the human form have been cloned recently [1,2]. Expression of the human NIS gene was found primarily in thyroid tissue, but to a much lower extent also in mammary gland and ovary [3]. The expression level of NIS is greatly decreased in tumours and in many transformed rat and human thyroid cell lines. Correspondingly, many thyroid carcinomas do not respond to radioiodide therapy. Iodide uptake and NIS content in thyroid cells are increased by thyroid-stimulating hormone (TSH)/cAMP, which seems to be due to enhanced gene expression [4]. Almost nothing is known about the transcriptional regulation of the human NIS gene. In order to understand the underlying molecular mechanisms of NIS gene expression, we have cloned, sequenced and characterized the 5'-putative promoter of the human NIS gene. Transient transfection assays were carried out in thyroid and non-thyroid cell lines to determine the tissue specificity of NIS gene expression.

MATERIALS AND METHODS

PCR and DNA-sequence analysis

In order to obtain the 5'-promoter sequence of the human NIS gene we used the PCR-based gene-walking human genomic library system of Clontech (Heidelberg, Germany). Two subsequent PCRs were performed on five libraries of uncloned, adapter-ligated genomic DNA fragments respectively, in that an aliquot of one first finished PCR was diluted and used as a template in a second, nested, PCR to get NIS-gene-specific DNA fragments. Beside the adapter-specific outer (AP1) and nested 5'-primer (AP2), the NIS-gene specific 3'-PCR primer (5'-CCTCCCTGTCCCCGCTGTCTGTCTCTGC-3') (NIS1) used in the first PCR was located at nucleotide position 142 of the 5'-untranslated region of the NIS gene, whereas the nested primer

for the second PCR (5'-TGCTGTCTGTCTGTCTGCCACC-TCCCC-3') (NIS2) localized to position 105 of the NIS cDNA sequence, previously published [2]. Except cycling parameters, amplification reactions were performed according to the protocols of Clontech by using Tth polymerase mixes, as well as all other reaction components of Clontech.

A PCR-amplified genomic NIS-gene DNA fragment from the *SspI* library was cloned in the TA-cloning vector pCR2.1 (Invitrogen, Leek, The Netherlands) resulting in pMB1. This fragment was then automatically sequenced in both directions by Eurogentec (Seraing, Belgium). To scan for putative eukaryotic promoter elements, computational sequence analysis on that DNA fragment was performed by using the EUKPROM program of PC/GENE version 6.5 software package (IntelliGenetics, Edegem, Belgium). Nucleotide homology analysis was done by using the GCG (Genetics Computer Group) software (University of Wisconsin, Madison, WI, U.S.A.).

Expression-vector constructions

On the basis of computer analysis, three regions of the PCR fragment in pMB1 were further subcloned into luciferase (LUC) reporter vector pGL3-Basic, a promoter- and enhancer-less vector (Promega, Mannheim, Germany): (a) a 1459 bp *EcoRI* fragment containing the complete sequence of 1402 NIS-specific nucleotides (pGL-N1), (b) an 850 bp *EcoRI/XbaI* fragment comprising the approx. 5'-half with 804 NIS-specific nucleotides (pGL-N2) and (c) a 612-bp *XbaI/EcoRI* fragment comprising the 3'-half with 602 NIS-specific nucleotides (pGL-N3) near to the translational start codon of the NIS gene. These fragments were cut from pMB1 and blunted by Klenow polymerase following cloning into the *SmaI* site of pGL3-Basic in sense and antisense orientation relative to the LUC-coding sequence respectively. The orientation of the respective DNA fragments was proved by restriction analysis. pGL-N3 was further used for a series of 5'-terminal deletion constructs by using the following restriction enzymes: (a) *SacI* (–117 bp, pGL-N4), (b) *NheI/BlnI* (–408 bp, pGL-N5), (c) *SmaI* (–436, pGL-N7), (d) *NheI/EspI*

Abbreviations used: NIS, sodium/iodide symporter; TSH, thyroid-stimulating hormone; SV40, simian virus 40; CHO, Chinese-hamster ovary; LUC, luciferase; TTF, thyroid-specific transcription factor.

¹ To whom correspondence should be addressed (e-mail Ulrich.Loos@medizin.uni-ulm.de).

The nucleotide sequence reported in this paper will appear in the EMBL, GenBank[®] and DDBJ Nucleotide Sequence Databases under the accession number AJ224353.

```

-1644 ATTTGTATTTTTCATACAGACAAGATCTCACTATGTTGCTCAGGGTAGTCTCAAATCTGGGACTCAAAT
-1574 GATCCTCCCACCTTCAGCCTCCCAAAGTGCTGGGATTACAGGCATAAGCCATCATGCCCCGCTCTGACGC
-1504 TGTTTCTTTCAACCCCAAGATTTCAGATTCCACCAGCTTATGGAGAAGGGAACCAAGTTTGAGATGCGT
-1434 GATTGCCAGAAAGTTGGAGGCTGAGCTGAGACTTGAACCCAGAGACCAGAACCTCCAGAGGTCAAAGTC
-1364 CTCCTGGGTCCCCCAGAGAAGGGCCCTGAGATGACAGCTCGTTGGTCTCATGGAAGCGTGACCCCCCA
-1294 GTAGACTTTCTCCACACCCAACCTTGGTTTCCTCATCTATATGATAAGGACAAGCCAGACTCTACCTCC
-1224 CTGGTGGTCATGGTCTCCGCTTATTCGGGTTTCATAACCTTAAAGGCCCTCGCACCACCTCAGTGAGCCA
-1154 TTTATGCCCTGGCACAGGGCCAACCTCTCAGTGCATATCTGCAAAGGAACCAATGAATGAGTGAATGAAGTG
-1084 ACAAATGAATAAAGGAATAAATGAATGAGGCACTTATCATGTACCAGGCTTTCGTTACCACGTCCCATTT
-1014 ATTCCTCTGAGGCAGGGTCTATTTTATCCTTGTGTACAGATGGGGAACCTAAGGCCAGGGAGGAGCAAA
-944 TCTTCCCCAAGTATGTACCCACTCAGAACTTGAGCTCTGAATGTCTCCACCCAGCTTAGCCCAAGAGCG
-874 GGGTTCAGTGATGCCACCCCTAAGGCTCTAGAGAAAGGGGTAGGCCACATGCCAGTTTGGGGGTGG
-804 TAAAGCCAGGTAAGTTTCTTTATGGGTCCCTGAAACCTGAAAGTGAACCCAGTCTGCATGAAAGT
-734 GAGCTCCCATAGCTCAAGGTATTCAGCAACAATACGGCTTTGAGTGCTGAAGCAGGCTGTGCAGGCTTG
-664 GATAGTGACATGCCCTCTCTGAGCCTCAATTTCCCACTGTCAACAGCAGACAGTGACAGCTGTGATCA
-594 GGGGATCAGTGATGCCAGGGATGGGTGGGTGCATGGGGATGGAGGGGCATTTGGGAGCCCTCCCGATAC
-524 CACCCCTGCAGCCACCCAGATAGCCTGTCTGGCCTGTCTGTCCAGTCCAGGGCTGAAAGGGTGCGGG
-454 TCCTGCCGCCCTAGGTCTGGAGGGCGAGTCGCGGTGACCCGGGAGCCCAATAAAATCTGCAACCCACAA
-384 TCACGAGCTGCTCCCGTAAGCCCCAAGGCGACCTCCAGCTGTCAGCGCTGAGCACAGCGCCAGGGAGAG
-314 GGACAGACAGCCGGCTGCATGGGACAGCGGAACCCAGAGTGAGAGGGGAGGTGGCAGGACAGACAGACAG
-244 CA

```

Figure 1 Nucleotide sequence of the 5'-flanking region of the human NIS gene

Numbering refers to the translational start codon. The putative TATA-box is enclosed and the GC-box (binding site for SP-1) is underlined. Two putative TTF-1-binding sites are wavy-underlined. The sequence in *italics* shows the overlap to the published cDNA [2], which starts at -347 relative to the start codon. The broken underlining (---) shows a potential binding site for AP-1.

(-513 bp, pGL-N8). In addition, a *PvuII* fragment was removed, resulting in an internal deletion of 256 bp (pGL-N6). All expression vector constructs are schematically represented in Figure 2 (below).

Cell culture and transient transfection assay

FRTL-5 rat thyroid cells (a gift from Dr. M. Derwahl, Medizinische Universitätsklinik Bergmannsheil, Bochum, Germany) were grown in Coon's modified Ham's F12 medium (Biochrom, Berlin, Germany) [5] containing 5% calf serum (Gibco Life Technologies, Eggenstein, Germany) and a mixture of hormones and growth factors (Sigma, Deisenhofen, Germany), including glycyl-L-histidyl-L-lysine acetate (10 ng/ml), cortisol (3.6 ng/ml), somatostatin (10 ng/ml), insulin (10 µg/ml), transferrin (5 µg/ml) and bovine TSH (10 m-units/ml). Antibiotics (Gibco) used were gentamicin (50 µg/ml) and amphotericin B (1.25 µg/ml). COS-7 and Chinese-hamster ovary (CHO) cells were grown in Dulbecco's Modified Eagle's Medium (Gibco), containing 5% fetal-calf serum (Gibco), gentamicin (50 µg/ml) and amphotericin B (1.25 µg/ml).

FRTL-5 (2×10^5) and COS-7 and CHO (1×10^5) cells were seeded in 35-mm-diameter plates before transfection (48 and 24 h respectively). Each transfection, performed transiently by the calcium phosphate co-precipitation method, contained 6 µg each of genomic NIS-fragment pGL3-Basic vector or 6 µg fragmentless pGL3-Basic vector respectively, and 1 µg of pSV-βGAL vector. A pGL3 construct containing simian-virus-40 (SV40) promoter and enhancer sequences served as a positive control. After transfection, cells were incubated for 48 h. Cell-extract preparations and LUC and β-galactosidase assays were carried out as previously described [6]. Experiments were performed three or four times in duplicate or triplicate, and results are means ± S.E.M. The stimulating effect of various genomic NIS-

specific fragments in pGL3-Basic on the expression of the LUC gene was calculated as the increase (fold) in LUC activities over the basic LUC activities obtained by the fragmentless pGL3-Basic vector.

RESULTS

Cloning and sequence analysis of the putative NIS-gene promoter

A 1.5 kbp PCR product containing the 5'-flanking region of the human NIS gene was amplified from a human genomic library by using 3'-primers complementary to the untranslated region of the human NIS cDNA [2]. After cloning in pCR2.1, double-stranded sequencing of this DNA region yielded 1402 gene-specific nucleotides with an overlapping stretch of 105 nucleotides identical with the NIS cDNA (which extends up to nt -347 from the translational start codon) demonstrating that the sequence is specific to the NIS gene. Thus the known sequence of this gene was extended by about 1300 bp further upstream up to -1644 from the ATG start codon. The complete nucleotide sequence of the NIS gene 5'-flanking region is presented in Figure 1. The sequence is rich in GC content (54%), shows a non-canonical TATA-box motif AATAAAT (nt -404 to -398), preceded by a GC-box (nt -436 to -423) spaced by 18 nucleotides to the putative TATA-box. Notable is the presence of four of the otherwise very scarce CpG dinucleotides in this short sequence. Another copy of the non-conventional TATA-box is located at nt -1079, but surrounded by a different context. Furthermore, we could detect two sites with similarity to the consensus sequence for the thyroid-specific transcription factor (TTF)-1 (ACTCAAGTG/A), with seven and eight matches out of nine nucleotides respectively (Figure 1). No thyroid-hormone-responsive-element or cAMP-responsive-element sequences were found.

Figure 2 Schematic representation of expression vector constructs used in transfection experiments.

The 5'-flanking region of the human NIS gene and various deletions were fused to the LUC gene in expression vector pGL3-Basic. The numbering of nucleotides refer to their positions relative to the ATG start codon of the NIS gene.

Transcriptional properties of reporter-gene constructs

We analysed promoter activity of various genomic DNA fragments by transient-transfection methods. The full 1402 bp sequence, and two fragments of it, were fused to a promoterless LUC expression vector pGL3-Basic. All expression vector constructs used are depicted in Figure 2. These constructs were transiently transfected into the rat thyroid cell line FRTL-5, and the transcriptional activity was measured. The cloned sequence exhibited considerable promoter activity, which resides exclusively in the 3'-half of the molecule (pGL-N3), and was capable of stimulating LUC-gene expression about 17-fold compared with fragmentless pGL3-Basic (Figure 3A). This transcriptional activity was specific, since the same sequences cloned in antisense orientation resulted in no activity. The data for the 5'-half of the NIS-flanking region (pGL-N2) suggest that this region contains

Figure 4 Stimulation of LUC-gene transcription in FRTL-5 and COS-7 cells by deletion constructs of NIS genomic DNA sequences

pGL-N3 and various deletions thereof (cf. Figure 2) were transfected in COS-7 and FRTL-5 cells and the relative transcriptional efficiency was determined. The experimental design and the calculation of data were the same as for Figure 3.

a silencing element, because even basic transcription mediated by pGL3-Basic was markedly suppressed. Very similar results were obtained when the above fragments were cloned into an enhancer-containing LUC-expression vector (results not shown).

Comparative promoter analysis in different cell types

In order to elucidate tissue-specific regulation of the NIS gene, we also transfected COS-7 and CHO cells with the above-mentioned constructs (Figure 3B). Compared with FRTL-5 cells, the promoter activity of pGL-N1 and pGL-N3 was also observed in COS-7, but to a lower extent. The activity in CHO cells in

Figure 3 Stimulation of LUC-gene transcription by 5'-upstream-located genomic DNA sequences of the human NIS gene

(A) Transfections of FRTL-5 cells. FRTL-5 cells were transfected with 6 μ g of pGL3 derivatives containing various genomic fragments of the human NIS gene in sense and antisense orientation relative to the LUC gene, or 6 μ g of pGL3-vector lacking any foreign sequences, and 1 μ g of pSV- β GAL control vector respectively. After transfection, cells were incubated for a minimum of 48 h and assayed for LUC activity, which was normalized for β -galactosidase activity. The LUC activity obtained by fragmentless pGL3-Basic vector alone was calculated to basic transcriptional levels of unity, and activities induced by NIS fragments were expressed as a fold increase over that basic value. (B) Transfections of COS-7 and CHO cells. Cells were transfected with pGL-N vectors containing 5'-flanking regions cloned in the sense direction relative to the LUC gene.

```

-193  GGACCGAAAGGGTCGGGGTCGCGCCCGCCT
      ||| ||||| ||||| |||||
-472  GGGCTGAAAGGGTCGGGTCCTGCCCGCCC

CCCAGGCTGGGGGCGGAGCTGCGCTGACCCCGGAG
|   | |||| ||||| ||| ||||| |||||
CTAGGTCTGGAGGCGGAGTCGCGGTGACCCGGGAG

TTCAATAAATCCGCAGCGCGAAAAGCGAGTG  -98
    ||||| ||| | |
CCC[AATAAAT]CTGCAACCCACAATCACGAGC  -377

```

Figure 5 Comparison of the human (lower) and rat (upper) 5'-flanking sequences of the NIS gene

The functional elements are denoted as in Figure 1. A potential AP-1 site is underlined with a broken line. Transcription in the rat sequence was determined to start at the last nucleotide shown (−98 relative to the ATG start codon).

contrast with FRTL-5 cells was rather poor. The relative ratios between the fragments and their orientation-dependency were similar in all three cell lines tested. In comparing transcriptional activities between these cell lines, different transfection efficiencies should be taken into consideration. For comparison, we included a positive control vector containing SV40 promoter and enhancer sequences. Although we normalized LUC values with respect to β -galactosidase activities, a large difference between FRTL-5 and CHO compared with COS-7 cells was obvious for this control vector (25-, 34- and 2865-fold relative to pGL3-Basic respectively). This is not surprising, since COS-7 cells contain the SV40 genome. In summary, strong promoter activity was observed in FRTL-5 cells, but the promoter was also active in non-thyroidal cells, albeit to a lesser extent.

Further delineation of promoter sequences

A series of deletions of the 5'-region of the human NIS gene (cf. Figure 2) was constructed to further delineate NIS-genomic DNA sequences showing promoter activity (Figure 4). Deletions up to nucleotide position −443 (pGL-N5) still had considerable promoter activity in FRTL-5 cells, whereas these deletions showed correspondingly less promoter activity in COS-7 cells. In transfections by using pGL-N7, in which a putative GC-box near to the putative TATA-box was deleted, and also by using pGL-N8, which lacks the GC- as well as the TATA-box, promoter activity was no longer detectable. This indicated that the sequence at least in the region of nucleotides −443 to −395 is essential to confer promoter activity to the NIS gene. It should be mentioned that the latter two constructs still showed some activity in COS-7 cells, pointing again to some unspecificity of the transcriptional activity in these cells.

It is interesting to compare the human sequence with that of the flanking region of the rat NIS gene that has been published very recently [7,8]. Whereas overall similarity is low (41.5%), it reaches 74% in the region delineated above, including full identity of the functional motifs TATA- and GC-box (Figure 5). Compared with the rat, the human NIS cDNA shows a very long 5'-untranslated region. By analogy to the rat sequence, where the transcriptional start site was mapped at nt −98 [7], transcription in the human NIS gene would start at position −377 relative to the ATG start codon.

DISCUSSION

The present study describes the sequencing and characterization of the 5'-flanking region of the human NIS gene. By use of a PCR-based approach, the sequence was extended up to −1644 bp relative to the ATG start codon. The 3'-half of this sequence near the translational start codon exhibited strong promoter activity when fused to a LUC expression vector and transfected transiently in a rat thyroid cell line. By using a series of deletion constructs we were able to confine this activity to a rather short stretch of DNA beginning 40 bp upstream of the putative TATA box and comprising a total of about 50 bp. This minimal region showed strong promoter activity in thyroid cells, but was also weakly active in non-thyroidal cell lines. Beyond the elements already mentioned (TATA- and GC-box) we could find only one potential binding site for AP-1 in this sequence (cf. Figure 5). Interestingly, this region is very similar to the corresponding region in the rat NIS gene, which shows much less sequence similarity elsewhere. We could not detect consensus sequences for TTFs, except for two putative TTF-1-binding sites at nt −1582 and −715. The significance of these potential binding sites is questionable, since they are not necessary for thyroid-specific transcriptional activity.

The features described above are different from the properties of the 5'-flanking region of other thyroid-specific genes, reviewed in [9]. In both the human thyroglobulin and thyroid peroxidase gene promoters, three TTF-1 binding sites are present within 170 bp upstream the transcriptional start site. Additionally, in the thyroid peroxidase gene, two further TTF-1 sites 5.5 kbp upstream function as tissue-specific enhancer elements. Both genes also contain consensus sites for TTF-2 and PAX-8, two other TTFs. In contrast, in the rat thyrotropin receptor promoter, only one TTF-1 site exists, which exerts both positive and negative transcriptional regulation [10].

The question of tissue-specific gene expression is not easy to address because of the lack of an accepted cellular model for the human thyroid gland. In the rat, a model is available with the continuous but differentiated cell line FRTL-5. Nevertheless, the two investigations on the rat NIS gene promoter came to contrary results regarding the cell-specific transcription [7,8]. Non-specificity could result from the presence of tissue-specific enhancers around the NIS gene or from the influence of chromatin structure and/or methylation on gene expression. The presence of numerous CpG dinucleotides, which form the CpG islands found at the 5'-end in 40% of tissue-specific regulated genes [11], in the minimal promoter region described herein suggests a role for methylation in transcriptional regulation of the NIS gene. These processes cannot be studied in reporter-gene assays with exogenous DNA added. Notwithstanding that, the present results demonstrate the functional characterization of the promoter of the human NIS gene and opens an avenue to further studies on genetic regulation of this important molecule.

We thank Dr. W. Minich (Abteilung Innere Medizin, University of Ulm) for critical reading of the manuscript before its submission. Celia Espinoza was a recipient of a short-term scholarship from the Deutscher Akademischer Austauschdienst. This work was supported by Ortho-Clinical Diagnostics (Neckargemünd, Germany).

REFERENCES

- 1 Dai, G., Levy, O. and Carrasco, N. (1996) *Nature* (London) **379**, 458–460
- 2 Smanik, P. A., Liu, Q., Furminger, T. L., Ryu, K., Xing, S., Mazzaferri, E. L. and Jhiang, S. M. (1996) *Biochem. Biophys. Res. Commun.* **226**, 339–345
- 3 Smanik, P. A., Ryu, K.-Y., Theil, K., Mazzaferri, E. L. and Jhiang, S. M. (1997) *Endocrinology* (Baltimore) **138**, 3555–3558
- 4 Saito, T., Endo, T., Kawaguchi, A., Ikeda, M., Nakazato, M., Kogai, T. and Onaya, T. (1997) *J. Clin. Endocrinol. Metab.* **82**, 3331–3336

-
- 5 Ambesi-Impiombato, F. S., Parks, L. A. M. and Coon, H. G. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3455–3459
 - 6 Behr, M., Ramsden, D. B. and Loos, U. (1997) *J. Clin. Endocrinol. Metab.* **82**, 1081–1087
 - 7 Tong, Q., Ryu, K. and Jhiang, S. M. (1997) *Biochem. Biophys. Res. Commun.* **239**, 34–41
 - 8 Endo, T., Kaneshige, M., Nakazato, M., Ohmori, M., Harii, N. and Onaya, T. (1997) *Mol. Endocrinol.* **11**, 1747–1755
 - 9 Damante, G. and Di Lauro, R. (1994) *Biochim. Biophys. Acta* **1218**, 255–266
 - 10 Ohmori, M., Shimura, H., Shimura, Y., Ikuyama, S. and Kohn, L. (1995) *Endocrinology (Baltimore)* **136**, 269–282
 - 11 Cross, S. and Bird, A. (1995) *Curr. Opin. Genet. Dev.* **5**, 309–314

Received 4 February 1998/19 February 1998; accepted 20 February 1998