

Isolation of genomic DNA fragments corresponding to genes modulated *in vivo* by a transcription factor

Julio Caubín, Teresa Iglesias⁺, Juan Bernal¹, Alberto Muñoz¹, Gabriel Márquez, José Luis Barbero and Angel Zaballos*

Departamento de Investigación, Pharmacia Antibióticos Farma S.A., Antonio López 109, 28026 Madrid and ¹Instituto de Investigaciones Biomédicas (CSIC), Arturo Duperier 4, 28029 Madrid, Spain

Received July 5, 1994; Revised and Accepted September 9, 1994

EMBL accession nos[§]

ABSTRACT

A new methodology for the identification of genes modulated by transcription factors *in vivo* is described. Mouse genomic DNA fragments bound by the thyroid hormone receptor (T3R) were selected and amplified *in vitro*. Subsequent hybridisation with biotinylated cDNA allowed the selection of those DNA fragments containing binding sites for T3R that corresponded to transcribed DNA. Expression analysis of the corresponding genes showed that more than 80% are indeed modulated by thyroid hormones *in vivo* in the liver. Together with the presence of consensus binding sites for T3R this result suggests that the selected DNA fragments may contain T3R transcriptional regulatory elements. This method, extensive to other ligand-modulated transcription factors, might be useful to all transcription factors with slight modifications.

INTRODUCTION

Transcription factors (TFs) are the main regulators of gene expression and thus involved in processes leading to developmental changes, responses to hormones and growth factors and some of the alterations causing tumoral transformation. Systematic search for genes regulated by a given TF is a difficult task. In general, the DNA sequences recognised by TFs are too short and/or degenerated, thus ruling out library screening by nucleic acid hybridisation. Furthermore, finding a consensus binding site in a DNA sequence does not guarantee a biologically functional element. Methods based on subtractive technologies (1–3) or differential screenings (4) are currently used to identify genes which are differentially expressed in cell cultures or tissues. This kind of approaches could be employed with ligand-modulated TFs, such as those belonging to the superfamily of nuclear hormone receptors (5). However, even in those instances, additional genetic or biochemical studies are needed to determine whether the regulation is transcriptionally direct or it involves subordinate responses through other TFs.

New PCR-based strategies take advantage of the TF ability to recognise and bind *in vitro* the DNA elements by which regulation of gene expression proceeds. The general version of these procedures is the whole genome PCR (WGPCR) technique (6,7). In brief, genomic DNA fragments bound by a given TF are selected and amplified by PCR. Iteration of the selection and amplification steps yields a collection of DNA fragments enriched in TF binding sites. Some of such fragments are functional in binding TF *in vitro* and in *cis*-modulation of transiently transfected reporter genes in cell cultures (8–10). However, the physiological significance of these elements is often uncertain and relies on the identification of the genes putatively controlled by them, which requires additional screenings of genomic and cDNA libraries (11). Experience has shown that although theoretically convincing, the isolation of genes via their regulator protein is very difficult (8–10,12). Most likely derived from artifactual protein–DNA interactions, ‘junk’ sites appear to mask the physiological elements, making arduous their selection. In fact, a recently developed method that presumably should select *in vivo* binding sites for TF also isolated irrelevant sites (13).

Here, we have addressed this problem by performing an additional step after the WGPCR process in order to select for TF binding sites close to, or overlapping with, transcribed DNA, i.e. expressed genes. This has been attained by hybridising the population of DNA fragments containing TF binding sites with cDNA derived from a target tissue and selecting the hybridised DNA. The rationale for such a step is that TF binding sites located in the coding or surrounding sequences of a gene would have a higher probability to be functional elements modulating its expression.

To test the adequacy of the proposed approach we have chosen the thyroid hormone receptor, a ligand-modulated TF, as a model. This system offers two important advantages: (i) a high number of genes are expected to be modulated in rodent liver by thyroid hormones (TH), major regulators of vertebrate metabolism, development and homeostasis (14,15), and (ii) manipulations of the thyroidal status allows a suitable analysis of the *in vivo*

*To whom correspondence should be addressed

⁺Present address, Instituto de Investigaciones Biomédicas (CSIC), Arturo Duperier 4, 28029 Madrid

[§]Z32539–Z32551, Z32553, Z32634

regulation of the genes corresponding to the selected fragments. WGPCR using T3R followed by selection with liver cDNA led us to isolate genomic DNA fragments corresponding to genes expressed in the liver. Most of them were in fact regulated *in vivo* by thyroid hormones and also contained putative binding sites for T3R. The usefulness of the described approach for other TFs is discussed.

MATERIALS AND METHODS

Whole genome PCR

Mouse genomic DNA was partially digested with *Sau3AI* and ligated to adaptors to allow subsequent PCR amplification (*Sau3AI* adaptor was composed of the oligonucleotides 5'-GATCCACGTCACCGTGAATTCGTG-3' and 5'-CACGAATTCACGGTGACGTG-3', the latter was the one used in PCR amplification). Ligation products were fractionated in 1×TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3), 5% acrylamide gel electrophoresis and DNA fragments from 250 to 750 bp were purified. This collection of fragments was used as the starting material for the binding/amplification procedure. Equimolar amounts (10–20 fmole) of *in vitro* translated hTR β (human thyroid-hormone receptor beta) and hRXR α (human 9-cis retinoic acid receptor alpha) proteins were incubated in a final volume of 40 μ l with 1 μ g of genomic DNA and 10 μ g of polydI-dC in binding buffer (25 mM Hepes pH 7.8, 100 mM KCl, 50 μ M ZnCl₂, 1 mM DTT, 10% glycerol) during 30 min at room temperature. Protein–DNA complexes were immunoprecipitated with 1 μ g of anti-hTR β antibody and 20 μ l of a 50% protein A-Sepharose (Pharmacia Biotech) suspension. Immunocomplexes were washed five times with 500 μ l of binding buffer, and 1/10 of the final material was used as template for 25 cycles of PCR amplification (30 s at 94°C; 30 s at 60°C; 90 s at 72°C). Amplified DNA was purified (Magic™ PCR Preps, Promega) and subjected to new cycles of binding/amplification, originating the successive collections of DNA fragments.

Selection of transcribed DNA

DNA obtained by the whole genome PCR procedure (also termed TRBS DNA, see 'Results and Discussion') was hybridised with cDNA to select the fraction of transcribed DNA. Biotinylated single stranded cDNA was synthesised from mouse liver poly(A)⁺RNA by random priming in the presence of 0.25 mM Biotin-21-dUTP (Clontech). Repetitive sequences in the driver cDNA (50 ng) were blocked by hybridisation with 10 μ g of denatured mouse genomic DNA in 10 μ l of 5×SSC at 68°C for 2 hours. Then, 1 μ g of denatured TRBS DNA (1 μ g/ μ l) was added and hybridisation continued overnight. The biotinylated complexes were captured by addition of 50 μ l of streptavidin paramagnetic particles (PMPs, Promega), whose non specific DNA binding sites had previously been blocked with single stranded DNA, and washed two times with 500 μ l of 0.1×SSC at 65°C for 10 min. TRBS-cocaptured DNA was finally recovered by boiling PMPs in distilled water and PCR-reamplified as before to obtain the TRGC DNA. After *EcoRI* cleavage, a TRGC library was constructed in the pUC18 vector. Individual clones were further analysed as explained in the text.

Plasmids and probes

pJLAhTR β , an *E.coli* expression vector of hTR β , was constructed by subcloning the hTR β cDNA present in plasmid

peA101 (16) into the *NdeI* site of the expression vector pJLA603 (Medac), in which gene transcription is controlled by the λ P_L promoter and the cIts repressor.

Plasmids for *in vitro* translation of hTR β and hRXR α proteins were prepared by subcloning the corresponding cDNAs, present in plasmids peA101 (16) and in pSGhRXR α (D.Barettino; unpublished data) respectively, into the SP6 promoter vector pSP64polyA (Promega). The oligonucleotide pairs 5'-CCCAA-GCTTGCCGCGCCATGACAGAAAATGGCC3'/5'GCGGATCCTAATCCCTCGAACACTTC-3' and 5'-CCCAAGCTTGA-GTTAGTCGCCACCATGGACACC3'/5'GCTCTAGAGGCCTAAGTCATTTGGTGC-3' were used to PCR-amplify hTR β and hRXR α cDNAs, respectively. In this way, cDNAs with an optimal translation initiation environment were generated (17). In addition, unique restriction sites were introduced to directly clone them into the polylinker region of pSP64polyA. Expression plasmids pSPhTR β and pSPhRXR α were finally obtained.

Probe for cyclophilin was the rat cDNA present in plasmid p1B15 (18). Probe for *SPOT14* was the DNA fragment obtained by PCR from genomic rat DNA using the oligonucleotides 5'-AGCATGCAAGTGCTAACG-3' and 5'-TGATGGAGGCTGGAGAAGTG-3'. The amplified sequence (381 bp in length) covers most of the coding region (19).

hTR β overexpression and antibody production

E.coli DH5 α cells transformed with the pJLAhTR β expression vector were grown at 30°C to an O.D.₅₉₅ = 0.6 in LB-ampicillin medium, and the culture was induced by temperature shift to 42°C for 2 hours. hTR β protein was produced as insoluble material that was extracted with 8 M urea. The solubilised hTR β protein was subsequently dialysed at room temperature against dialysis buffer (50 mM Tris–HCl pH 8, 0.2 M NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA and 5% glycerol) containing first 2 M urea for 2 hours, then 0.5 M urea for 2 hours and finally without urea for 16 hours at 4°C. Soluble hTR β protein was further purified by SDS–PAGE prior to rabbit immunisation. An hTR β –Sepharose column was prepared by coupling soluble hTR β to CNBr-activated Sepharose (Pharmacia) and used to purify specific anti-TR β immunoglobulins from serum.

In vitro transcription and translation

Plasmids pSPhTR β and pSPhRXR α were used as templates in a transcription-translation coupled system (TNT™ lysate, Promega) to synthesise hTR β and hRXR α proteins. Parallel reactions in the presence of [³⁵S]Methionine (Amersham) were analysed in SDS–PAGE to quantify the relative amounts of each protein.

Electrophoretic mobility shift assays (EMSA)

Equimolar amounts (10–20 fmole) of *in vitro* translated hTR β and hRXR α proteins were incubated in 50 μ l of binding buffer with the indicated ³²P-labelled DNA fragments and 1 μ g of polydI-dC as non specific competitor during 30 min at room temperature. For the competition assays, a 125-fold mass excess of the indicated unlabelled competitor was added. Protein–DNA complexes and free DNA were resolved in 0.5×TBE, 5% acrylamide gel electrophoresis.

Preparation of hypothyroid and hyperthyroid mice

Adult postnatal day 60, Swiss mice were used. To induce hypothyroidism, 0.02% of methylmercaptoimidazol was

administered in the drinking water for 25 days. Hyperthyroidism was induced by daily injections of 50 μg of 3,3',5-triiodo-L-thyronine (Sigma) per animal for 5 consecutive days. Animals were sacrificed 90 min after the last injection. The livers from each group (hypothyroid, hyperthyroid and control) consisting of five animals, were pooled for mRNA extraction. The efficacy of the hormonal manipulations were demonstrated by analysing *SPOT14* expression in Northern blots. *SPOT14* mRNA expression has been shown to be regulated by TH (19) and is therefore a reliable indication of TH effects on the liver. Maintenance and handling of the animals were as recommended by the NIH Guide-Lines on The Care and Use of Laboratory Animals.

mRNA extraction, Northern blots and nucleic acid labelling

Total RNA was obtained by the guanidine-phenol-chloroform method (20) and the poly(A)⁺ fraction selected by hybridisation to biotinylated oligo(dT) (Promega) and PMPs capturing. For Northern analysis, 1 μg of poly(A)⁺ RNA was electrophoresed in agarose gels and blotted onto Nylon membranes using standard protocols (21). Inserts from the TRGC clones were *EcoRI*-excised, agarose purified, labelled with [α -³²P]dCTP by the

random-priming method (Megaprime labelling system, Amersham) and used to probe membranes. Hybridisations were carried out in Rapid-Hyb buffer from Amersham as recommended by the supplier. Membranes were washed at high stringency and autoradiographed. DNA fragments used in EMSA were ³²P-labelled by Klenow fill-in reactions using convenient dNTPs.

Nucleic acid sequencing and analysis

Sequences were determined by the dideoxy chain termination method (22) using the Sequenase 2.0 sequencing kit (USB). Sequence analysis were carried out with GCG (23) and Intelligenetics package of programs.

RESULTS AND DISCUSSION

Isolation of transcribed genomic DNA fragments bound by T3R

Figure 1 illustrates the methodology followed. We performed WGPCR basically as described (6, Figure 1A). Mouse genomic DNA was partially digested with *Sau3A1* and modified by adaptor ligation. DNA fragments suitable in size for PCR amplification

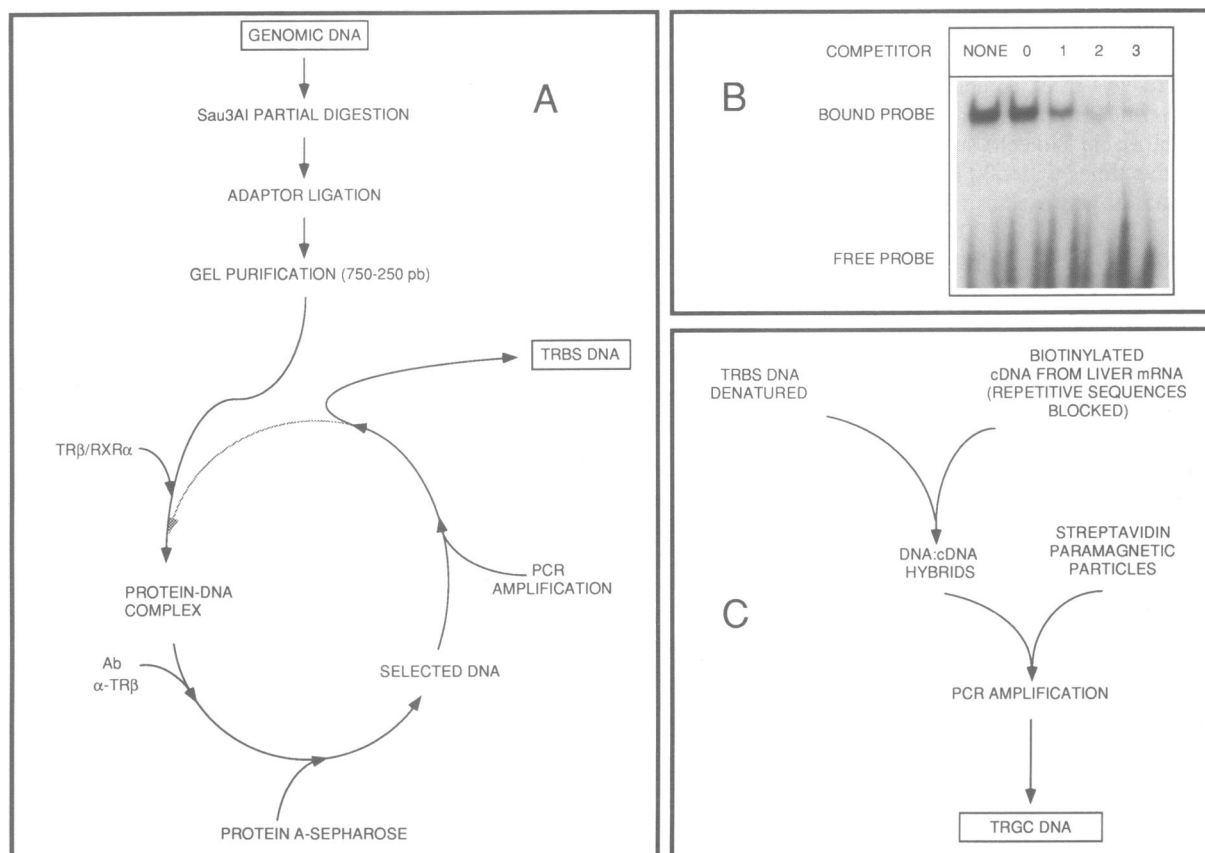


Figure 1. Scheme of the protocol followed to obtain DNA fragments corresponding to TH-regulated genes. **A.** Whole Genome PCR. Genomic DNA was digested, ligated to adaptors to allow PCR amplification and subjected to several cycles of binding to TR β /RXR α and amplification by PCR. DNA enriched in thyroid receptor binding sites was finally obtained. **B.** Competitive EMSA. Binding of TR β /RXR α to the P0 element was competed out with the DNA obtained by WGPCR to confirm enrichment in TRBS through the selection process. Lanes: 0, starting DNA; 1–3, TRBS DNA obtained after 1–3 cycles of WGPCR. **C.** Selection of transcribed DNA. Biotinylated cDNA was used as driver to select transcribed TRBS DNA. Complexes cDNA:DNA were captured through biotin-streptavidin interaction. PCR amplification of this material yielded the TRGC DNA.

(250–750 pb) were purified. This DNA was incubated with T3R (in the form of hTR β /hRXR α heterodimer) and specific protein–DNA complexes were immunoprecipitated with an antibody raised against the hTR β protein (see Materials and Methods). Co-precipitated DNA was amplified by PCR and then used for further rounds of selection and PCR amplification. Two or three of these cycles rendered collections of DNA fragments containing binding sites for T3R (thyroid receptor binding sites, TRBS DNA) as assessed in competition assays with a well characterised binding site for T3R (CTAGATCAGGTCATG-ACCTGATCATG, P0 element, ref. 5, Figure 1B). Since no significant differences in competition efficacy were observed between TRBS DNA from cycles 2 or 3, the former was used in posterior steps.

Next, the novel hybridisation step was carried out (Figure 1C). Repetitive sequences present in biotinylated cDNA made from mouse liver mRNA, were blocked by prehybridisation with sonicated heat-denatured mouse DNA. This step forces the posterior formation of cDNA–TRBS hybrids through single or low copy sequences, favouring the selection of TRBS DNA linked to expressed genes. Then, TRBS DNA was added and hybridisation continued until cDNA–TRBS hybrids were formed. Biotinylated molecules were captured with streptavidin paramagnetic particles and washed at high stringency. Co-captured TRBS DNA was amplified again by PCR, obtaining the final collection of thyroid receptor genomic binding sites linked to cDNA (TRGC DNA). Cleavage with *EcoRI* and cloning in pUC18 originated a TRGC library. In total, 72 clones were checked by restriction analysis. Clones without inserts or containing multiple ones and those whose inserts were shorter than 200 bp were discarded. Finally, 39 clones were chosen for a detailed study.

Characterisation of the selected TRGC DNA fragments

Should the strategy worked as predicted, the TRGC fragments would have three properties: all or part of their sequence should be transcribed DNA, the corresponding genes should be transcriptionally regulated by TH, and one or more units of the T3R binding motif AGGTCA or variants thereof (5), should be present in their sequence. To analyse mRNA expression and TH regulation, Northern blots containing poly(A)⁺RNA from livers of hypothyroid, euthyroid and hyperthyroid mice were probed with individual TRGC DNA fragments (Figure 2). Sixteen clones gave signals after exposure times ranging from hours to one week. More than 80% (14 out of 16) of the transcripts showed regulation by TH *in vivo* as deduced from the Northern analysis (Figure 2). It has been reported that about 5% of transcripts in the liver could be directly regulated by TH (15). The proportion of TH-responsive transcripts we have obtained suggests that there are elements in the selected DNA fragments that modulate gene transcription upon T3R binding and TH action.

Sequences of the Northern-positive TRGC fragments were determined and used in data base searching (Figures 2 and 3). Most of them corresponded to genes described in mouse or in other mammals: protein disulphide isomerase (*PDI*), alcohol dehydrogenase (*ADH*), aldehyde dehydrogenase (*ALDH*), one of the μ -class glutathione-S-transferases (*GST*), alpha-enolase, an endogenous murine retroviral sequence (*MuLV*), NADH ubiquinone oxidoreductase (*CI-23kD*), major urinary protein (*MUP*), ubiquitin activating enzyme E1 (*UAP E1*), a member of the cytochrome p450 testosterone 16-alpha-hydroxylase family

CLONE/SIZE (bp)	NORTHERN			TH <i>IN VIVO</i>	mRNA SIZE(kb)	GENE	FRAGMENT LOCATION
	Ho	C	H3				
TRGC07/531	---	---	---	↓	2,4	<i>PDI</i>	int/ex
TRGC10/674	---	---	---	↓	1,1	<i>ADH</i>	ex/int/ex
TRGC14/422	---	---	---	↓	1,3	<i>GST</i>	int/ex/int/ex
TRGC15/650	---	---	---	↑	1,7	<i>ENOLASE</i>	exon
TRGC19/522	---	---	---	↑↓	7	<i>MuLV</i>	exon
TRGC39/586	---	---	---	↓	0,75	<i>CI-23kD</i>	int/ex/int
TRGC42/535	---	---	---	→	2,8-2,4	<i>ATP SYN</i>	ex/int
TRGC46/555	---	---	---	→	1,6	<i>r18S</i>	transcribed DNA
TRGC52/237	---	---	---	↓	2	<i>ALDH</i>	ex/int
TRGC55/475	---	---	---	↑	0,8	<i>MUP</i>	prom
TRGC67/198	---	---	---	↓	3,9	<i>UAP E1</i>	int/ex
TRGC73/535	---	---	---	↓	1,6	<i>P450</i>	prom/ex/int
TRGC78/414	---	---	---	↑↓	3,7	<i>UNK</i>	---
TRGC79/706	---	---	---	↓	7,9-6,7	<i>UNK</i>	---
TRGC81/700	---	---	---	↓	0,9	<i>L35</i>	ex/3'UTR
TRGC82/348	---	---	---	↑↓	1,7	<i>UNK</i>	---
----	---	---	---	↑		<i>SPOT 14</i>	----
----	---	---	---	→		<i>CYCLOP</i>	----

Figure 2. Expression of mRNAs associated to the TRGC selected DNA fragments. One μ g of poly(A)⁺RNA from livers of hypothyroid (HO), control (C) and hyperthyroid (H3) mice was subjected to Northern analysis and probed with ³²P-labelled TRGC inserts. Exposure times varied depending on the relative expression of each mRNA. Cyclophilin (*CYCLOP*), a TH-non-regulated gene, and *SPOT14*, a positively TH-regulated gene were used as controls. Right to the Northern bands, the observed regulation by TH (↓, down-regulation; ↑, up-regulation; ↑↓, biphasic response; →, no response) apparent mRNA size, corresponding gene (when identified by sequence data base searching; *UNK*, unidentified gene) and location of the fragment along the gene structure (*int*, intron; *ex*, exon; *prom*, promoter, *UTR*, untranslated region) are indicated. Accession numbers for TRGC inserts are: Z32539 (TRGC07), Z32540 (TRGC10), Z32541 (TRGC14), Z32542 (TRGC15), Z32553 (TRGC19), Z32543 (TRGC39), Z32544 (TRGC42), Z32545 (TRGC52), Z32546 (TRGC55), Z32547 (TRGC67), Z32634 (TRGC73), Z32548 (TRGC78), Z32549 (TRGC79), Z32550 (TRGC81) and Z32551 (TRGC82). 18S ribosomal insert in TRGC46 corresponded to 1055–1609 coordinates of the reported sequence (31). Three distinct, but highly similar genes for cytochrome p450 testosterone 16-alpha-hydroxylase have been described (32). Sequence alignments did not allow us to unequivocally assign TRGC73 sequence to any of them. Either it could correspond to a new gene of the family or differences in sequence could be due to mutations introduced by *Taq* polymerase during PCR amplification.

(*P450*) and ribosomal protein L35. Only for *MUP* a direct transcriptional regulation by TH has been described (24), and biochemical data also suggest that *ADH* is modulated by TH (25). Modulation by TH has not previously been described for the remaining genes. Four of these genes encode proteins involved

CLONE, GENE	MAIN MOTIFS		OTHER HALF SITES	
			DIRECT	INVERTED
TRGC07, PDI	ATGTGGG AGGTCTAGCTC AGGCTTT TGACTT GGGGACT	DR1, P5	TAGGCC AGTTC AACCCCTCT TTCTCG AGGTCA TGATTAC GACGAT AGGTG AGTNCNGG	ACAGCT TGACCT TAGACTT ACCC TGCTCT TAGAAATTC
TRGC10, ADH			ACCATCG AGGACA TAGAAGT CGAATT AGGTGACT TACCCT ATATCA AGGTCA AAAGTTTG TGGNAA AGGTA ACCGAAC	GTATGTT TACCT TAATCTCA AGATAA TGACT CTTTGGA CTGGACT TACCT TTAATCC
TRGC14, GST	AATTGAT GGTCA CACA AGTCA CCCGAG	DR4		GATGGAG GGACT CATATCC CCCAT TGACCT TCCTCGC ATCCAG TGCT ACTTAAT
TRGC15, ENOLASE	TGGCGGAT TGACCT ACCT TGACCA ACCGTAA	IR4	GTCTAC AGTCT GGCAAGT TGGCAG AGTTC A TGGT AC GAGACT AGGACA CTTTCAAT	TAGATT TGACCT TAACATTC GC TAC AT TGACCT TGGACTTT CTCCAGAT TGACCT CAACAGG ANCTGG TGACCT GTACAAG CCNC TGACCA ANGACGAC GGCT TGACCA ACGTCTCT TCATCG CGACT TGGTGTG
TRGC19, MuLV	ACCCAG AGTCA AGGTCA AGGTA AAAAAAT	DR4	GAGTAAG AGGACA TAGACCA GTCTCC AGTAA GTCAAGT CCAGGT AGTCA AGTCCGGTG AATGAAT AGGACA ATCAAGG	GGAAA TGACTT TACTGAA AAAA TGACCA AGAAACT GAGACT TGACCA AAATTAAC
TRGC39, CI-23kD	ACTAGG AGTCA AGGTCTCACTCC CTGGCA TGACCT AGGT TGACCT TTTGA CTCTGT AGGCA TC TGCT GCAC AGTCA GTCTGCC	IP0 IP1, IR4 P2, IP4	AGCTGCT AGGTTA TCAGGAC	CACTCC TGACT TGAAAAG CTCCAT TACCT TCCTCC CCCT TGACT TCCTGCAG
TRGC52, ALDH	CGTGGCT TGCTCT CTCC TGCTCA CGG AGTCA GTACTTG	IP4, DR3	CTATAAG TGTC AGGGAGTG ACAGA AGGTCA GATACAG	GCGAGT TGGCT TGCAGGG GACTCC TGCT TGGCTCT
TRGC55, MUP			AATCCAT AGGTA TGAAATT TTTGCA AGGGCA AGGAACA GTATAT AGGACA GCAAAGG	CAATT TGGCT CTAATCA TTGGACT TGACCT TAGTCTG
TRGC67, UAP E1			CTTGGT GGTCA AGGCTGT	TGGACC TGACT TGGAGTCC CCCTACAT TGACCA AGGAACT
TRGC73, P450	CACACAG AGGACA AGGGCAACCCCTGG TTGTACA AGGTCA AGTCTT AGTGGT G	DR1 P1	CTACACC AGGCCA AACAAC TTGCTGC AGGTCA ACATATA GTAA AGGTTA TGATGTTG	CTGGACT TGGCT GTINTATT TGTGCCAT TGGCT TGCTCGG CA TGGCT TGCTCTGGT GACT
TRGC78	CTTGCT AGGACA TGACTT AGTGGT	P0	GGCCAG AGGACA TCAAATC TGGTT AGGACA CAGAAGA GCTCGT AGTCT CTAGTCC	CTGAAGAT TGACCT TGCATGC CAGAGAT TGACCT TGGAACAG
TRGC79	GCCCTGT TGACCT TGCTGT TGACCT TT AGTTC AGTTTCAG	IR5, IP2	CTTNCCT AGGACA GANCNCC	TGAGGAG TGGCT CACCATC
TRGC81, L35			CTGGAA AGTCA CTCGGGC GGAGAG AGTCA TAGTTT	TGAC TGACT TGGGGAAA
TRGC82			AAAATGG AGTCT CTTTCC TTTCC AGTTC AGCTGAA CCAGGT AGTCA AGGGACA TGGCC AGGCCA CAGCATG	CAGAACAT TGCT TATCAGC TCTGCT TGGCT TGGCAT CCACAG TGACCT CAGCGAA

Figure 3. Putative thyroid-hormone receptor binding sites on the TRGC DNA fragments responsive to TH. Sequences corresponding to known genes are presented 5' to 3' relative to the transcription direction. All the AGGTCA (direct half site) and TGACCT (inverted half site) motifs found in each fragment are presented in bold and aligned (allowing one mismatch). Seven nucleotides at each side of the motif are also presented. Only remarkable combinations of these elements (spacing less than 6 bp) are shown in detail. Arrows mark the direction of the motifs. DRx, AGGTCA(x)AGGTCA (direct repeat); IRx, TGACCT(x)TGACCT (inverted repeat); Px, AGGTCA(x)TGACCT (palindrome); IPx, TCACTT(x)AGGTCA (inverted palindrome).

in detoxifying processes (*ADH*, *ALDH*, *GST* and *UAP E1*). Interestingly, they all are down-regulated by TH (actually, acetaldehyde, produced by *ADH* from ethanol is a substrate for *ALDH*). Three other TRGC fragments, whose corresponding genes still could not be assigned, hybridised with transcripts also modulated by TH *in vivo*. Only two TH-non responding genes were isolated in the process, the alpha subunit of the

mitochondrial ATP synthase (*ATP SYN*) and the ribosomal 18S RNA (*r18S*).

As expected from the selection criteria several consensus AGGTCA motifs were found in all TRGC DNA fragments corresponding to TH-responsive genes (Figure 3). A clear relationship between the number of sites, their spacing and orientation, and the extent and type of regulation could however

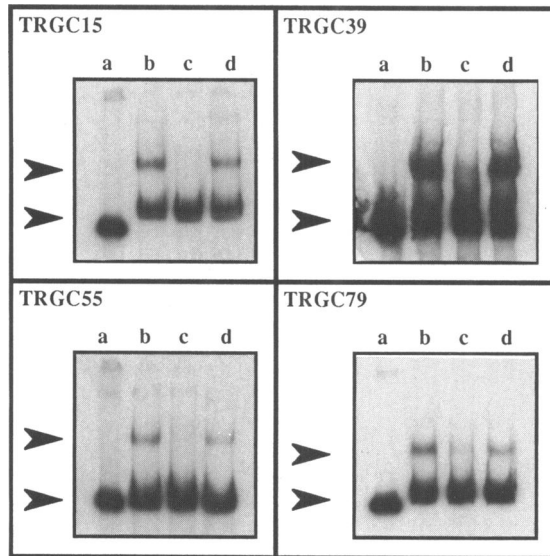


Figure 4. Binding of T3R to TRGC DNA fragments. The indicated ^{32}P -labelled TRGC DNA fragments (5–10 ng; 10,000 cpm) were incubated with the TR β /RXR α heterodimer and analysed in EMSA as described in 'Materials and methods'. Lanes: a, labelled fragment alone; b, no competitor added; c, 1000-fold molar excess of the P0 element as competitor; d, 1000-fold molar excess of the Sau3A1 adaptor as competitor. Upper and lower arrowheads mark the positions of bound and free DNA respectively. All TRGC DNA fragments were analysed giving similar results (data not shown).

not be deduced. Although consensus thyroid response elements, direct repeats of the AGGTCA motif spaced by four nucleotides (26,27), are found in some of our fragments (Figure 3), the presence of other elements and mainly dispersed AGGTCA half sites is more common. Furthermore, T3R was able to bind isolated TRGC DNA fragments. Figure 4 shows the interaction of T3R with some TRGC DNA fragments analysed in EMSA. This binding seemed to be specific, as it was competed out with P0 (Figure 4, lanes c) and not with an irrelevant DNA (Figure 4, lanes d). Obviously a detailed analysis is required to determine which of the putative elements presented in Figure 3 is responsible for such interaction.

All types of TH control, up- and down-regulations and also biphasic responses (down-regulation in both hypo and hyperthyroidism) were observed (Figure 2). Comparing our results with those from others (3,15,28,29), a major discrepancy is observed. We have found a high percentage of down-regulated genes, particularly in the hyperthyroid state, whereas in the mentioned reports, the majority of the genes are up-regulated. Likely explanations for that may arise from particular aspects of our approach: (i) it is unknown whether the TR β /RXR α heterodimer used in the WGPCR selection has some site preference, (ii) the cDNA hybridisation introduced a bias that underscores or discards elements ≥ 500 bp upstream of the transcription start sites and (iii) the only two cycles of selection performed could yield a population of DNA fragments non substantially enriched in *in vitro* canonical high affinity T3R binding sites. All together, the procedure might have led to the preferential selection of thyroid responsive negative elements. It has been recently reported that negative TH elements could be mainly multiple half sites of the AGGTCA consensus (30).

This is the kind of elements we have predominantly detected in the sequence of our TRGC fragments (Figure 3).

CONCLUSIONS

WGPCR is a powerful method to isolate DNA fragments containing *in vitro* binding sites for TFs. However, it is not a straightforward procedure to identify *in vivo* TF-regulated genes. The procedure here described, which includes an additional step of hybridisation with cDNAs from a target tissue, has been useful to identify a subset of genes modulated by T3R. Limitations introduced by effectiveness in PCR amplification restrict the genes potentially selectable to those in which binding sites for TFs fall in or near the spliced mRNA. This new method can be applied to other ligand-modulated TFs for which an easy *in vivo* identification of the transcript regulation is allowed. However, the selection of DNA fragments containing TF binding sites and the cDNA hybridisation are ligand-independent. The method could be extensive to all TFs reproducing the ligand action by, for example, inducing a cell culture to overexpress the TF under study. Another interesting feature of the system resides in the fact that TRBS DNA, or in general TFBS DNA (transcription factor binding sites DNA) could be a common substrate to isolate genes regulated in any other tissue. Cell-, tissue- or developmental-stage specific clones can potentially be obtained from a single TFBS collection of fragments, by simply changing the driver cDNA source.

ACKNOWLEDGEMENTS

We thank Dr R.M.Evans for hTR β cDNA and Drs D.Baretino and H.G.Stunnenberg for hRXR α cDNA; E.Rodríguez for help in the production of anti-hTR β and C.Martínez Caja, M.M.Sánchez and A.Martínez for lab assistance. Drs M.Aracil, I.Barthelemy, L.Carramolino, J.Gutiérrez and E.Mira critically reviewed the manuscript. J.C. and T.I. were predoctoral and postdoctoral fellows from the Spanish Ministerio de Educacin y Ciencia, respectively.

REFERENCES

1. Scott, M.D., Westphal, H.H. and Rigby, R.W.J. (1983) *Cell* **34**, 557–567.
2. Davis, M.M., Cohen, D.I., Nielsen, E.L., Steinmetz, M., Paul, W.E. and Leroy, H. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2194–2198.
3. Wang, Z. and Brown, D.D. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11505–11509.
4. Liang, P. and Pardee, A.B. (1992) *Science* **257**, 967–971.
5. Beato, M. (1989) *Cell* **56**, 335–344.
6. Kinzler, K.W. and Vogelstein, B. (1989) *Nucleic Acids Res.* **17**, 3645–3653.
7. Woodring, E.W. and Funk, W. (1993) *TIBS* **18**, 77–80.
8. Inoue, S., Kondo, S., Hashimoto, M., Kondo, T. and Muramatsu, M. (1991) *Nucleic Acids Res.* **19**, 4091–4096.
9. El-Deiry, W.S., Kern, S.E., Pietenpol, J.A., Kinzler, K.W. and Vogelstein, B. (1992) *Nature Genetics* **1**, 45–49.
10. Costa-Giomi, M.P., Gaub, M.P., Chambon, P. and Abarzúa, P. (1992) *Nucleic Acids Res.* **20**, 3223–3232.
11. Inoue, S., Orimo, A., Hosoi, T., Kondo, T., Ikegami, A., Ouchi, Y., Orimo, H. and Muramatsu, M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11117–11121.
12. Kinzler, K.W. and Vogelstein, B. (1990) *Mol. Cell Biol.* **10**, 634–642.
13. Gould, A.P., Brookman, J.J., Strutt, D.I. and White, R.A.H. (1990) *Nature* **348**, 308–312.
14. Glass, C.K. and Holloway, J.M. (1990) *Biochim Biophys. Acta* **1032**, 157–176.
15. Seelig, S., Liaw, C., Towle, H.C. and Oppenheimer, J.H. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4733–4737.

16. Weinberger, C., Thompson, C.C., Ong, E.S., Lebo, R., Gruol, D.J. and Evans, R.M. (1986) *Nature* **324**, 641–646.
17. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8132.
18. Danielson, P.E., Forss-Petter, S., Brow, M.A., Calavetta, L., Douglass, J., Milner, R.J. and Sutcliffe, J.G. (1988) *DNA* **7**, 261–267.
19. Liaw, C.W. and Towle, H.C. (1984) *J. Biol. Chem.* **259**, 7253–7260.
20. Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159.
21. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbour Laboratory Press, Cold Spring Harbour.
22. Sanger, F., Miklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
23. Devereux, J., Haerberli, P. and Smithies, O. (1984) *Nucl. Acids Res.* **12**, 387–395.
24. Kuhn, N.J., Woodworth-Gutai, M., Gross, K.W. and Held, W.A. (1984) *Nucleic Acids Res.* **12**, 6073–6090.
25. Dipple, K.M., Qulali, M., Ross, R.A. and Crabb, D.W. (1993) *Hepatology* **17**, 701–706.
26. Näär, A.M., Boutin, J.M., Lipkin, S.M., Yu, V.C., Holloway, J.M., Glass, C.K. and Rosenfeld, M.G. (1991) *Cell* **65**, 1267–1279.
27. Umesoro, K., Murakami, K.K., Thompson, C.C. and Evans, R.M. (1991) *Cell* **65**, 1255–1266.
28. Buckbinder, L. and Brown, D.D. (1992) *J. Biol. Chem.* **267**, 5786–5791.
29. Shi-Y.B. and Brown, D.D. (1993) *J. Biol. Chem.* **268**, 20312–20317.
30. Carr, F.E. and Wong, N.C.W. (1994) *J. Biol. Chem.* **269**, 4175–4179.
31. Raynal, F., Michot, B. and Bachelier, J.P. (1984) *FEBS Lett.* **167**, 263–268.
32. Wong, C., Itakura, T., Kawajiri, K., Skow, L. and Negishi, M. (1989) *J. Biol. Chem.* **264**, 2920–2927.