# Distinct DNA binding preferences for the c-Myc/Max and Max/Max dimers

# David L.C.Solomon, Bruno Amati and Hartmut Land\*

Growth Control and Development Laboratory, Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Received August 27, 1993; Revised and Accepted October 6, 1993

# ABSTRACT

The transcription factor c-Myc and its dimerisation partner Max are members of the basic/helix-loophelix/leucine-zipper (bHLH-Z) family and bind to the DNA core sequence CACGTG. Using a site-selection protocol, we determined the complete 12 base pair consensus binding sites of c-Myc/Max (RACCACGTG-GTY) and Max/Max (RANCACGTGNTY) dimers. We find that the c-Myc/Max dimer fails to bind the core when it is flanked by a 5'T or a 3'A, while the Max/Max dimer readily binds such sequences. Furthermore we show that inappropriate flanking sequences preclude transactivation by c-Myc *in vivo*. In conclusion, Max/Max dimers are less discriminatory than c-Myc/Max and may regulate other genes in addition to c-Myc/Max targets.

# INTRODUCTION

The c-Myc (Myc) protein is known to be a key regulator of cell proliferation, differentiation and apoptosis (1-4). Myc and its partner protein Max dimerise via helix-loop-helix-leucine zipper interactions and bind the core DNA sequence CACGTG (5-15); for reviews, 16, 17). Two major max mRNA splice variants encode the proteins Max1 and Max2, which are equally abundant in cells and differ by the presence of a nine amino-acid insert preceding the basic domain in Max2 (5, 6). Other variants have been described, in which the carboxy-terminus is truncated ( $\Delta$ Max1 and  $\Delta$ Max2) or replaced by alternative sequences (18, 19). Whereas Max proteins form homodimers which bind to the same DNA sequence (5, 7, 8, 11, 12, 20), Myc alone neither forms homodimers nor binds DNA except at very high concentrations *in vitro* (7, 12 and references therein).

It was recently demonstrated that Myc is a transcriptional regulator (7, 13, 21-25). Dimerisation with Max is required for all tested biological activities of Myc, including transactivation of promoters containing Myc/Max binding sites (7), cooperative transformation of primary cells (14) as well as induction of cell-cycle progression and apoptosis in non-transformed cells (25b). Furthermore, these activities all require the Myc amino-terminal transactivation domain (4, 26, 27). Thus, the Myc/Max dimer is a transcription factor responsible for the biological activities of Myc.

Our studies in yeast showed that Max/Max dimers can competitively antagonise Myc/Max *in vivo* (7). Max can also act as a dose-dependent antagonist of Myc function in transactivation and co-transformation assays in mammalian cells (14, 18, 21-23, 28, 29). Thus, specific target genes may be regulated by both Myc/Max and Max/Max dimers.

It was previously shown that dimers of bHLH-containing proteins specifically recognise three base pairs on either side of the CANNTG core (30). Studies using purified Myc proteins which bound DNA as homodimers identified related 12 base pair motifs (31, 32). *In vitro* translated Myc bHLH-Z peptides, which bind DNA in association with Max proteins present in the reticulocyte lysate (12), also recognise a similar motif (33). However, complete consensus sequences for Myc/Max or Max/Max dimers have not been determined. Such determination requires a binding site-selection from an initially random oligonucleotide pool (34), or comprehensive testing of all possible binding site variants. In this work, we use the former approach to define and compare the DNA binding specificity of the Myc/Max and Max/Max dimers.

# METHODS

# Proteins

Myc and Max2 proteins were synthesised as previously described (12). Synthesis was controlled for by immunoprecipitating <sup>35</sup>S labelled proteins with the appropriate antibody followed by SDS gel electrophoresis. DNA binding activity of unlabelled protein was monitored in a trial mobility-shift assay using <sup>32</sup>P labelled wt12 oligonucleotide (see below).

## Oligonucleotides and probes

Double stranded probes were prepared by annealing synthetic oligonucleotides and were end-labelled with T4 Polynucleotide Kinase with <sup>32</sup>P  $\gamma$ -ATP under standard conditions. The double-stranded wt12 probe (12) (derived from the E<sub>MS</sub> sequence; 33, 35) is composed of wt1: 5'-TCGACGCCGAC<u>CACGTGGT</u> CCCTC-3' and wt2: 5'-TCGAGAGGGAC<u>CACGTGGT</u> CGGCG-3'. For site selection, we used the following oligonucleotides (34): Primer F: 5' GCTGCAGTTGCACTGAATT-CGCCTC 3', Primer R: 5' CAGGTCAGTTCAGCGGAT-

<sup>\*</sup> To whom correspondence should be addressed

#### Binding site selection and mobility shift assays

Binding site selection was carried out as described (34), by immunoprecipitation of DNA sequences bound to Myc/Max and Max/Max complexes. Immunoprecipitation was carried out as described (12) using the peptide specific antisera PM (anti-Myc) (36) and MX (anti-Max) (12). Each antibody is specific for its substrate and can be blocked by its cognate immunogenic peptide. There are no detectable Myc/Myc dimers in our system and immunoprecipitation of a labelled oligonucleotide by PM is specific for material complexed to Myc/Max dimers (12). Selected DNA sequences were cloned into the vector plasmid Bluescript KS+ (Stratagene) and sequenced as plasmids from the T3 or T7 primers using the Sequenase kit (United States Biochemicals).

Mobility shift assays were performed as described in (12) and complexes were quantified on a Molecular Dynamics phosphor imager using ImageQuant v. 2.0. The background values, obtained with unprogrammed reticulocyte lysate, were subtracted from all others.

## Cell culture and transfection

Rat-1 MycER fibroblasts expressing the  $\beta$ -oestradiol inducible MycER chimaera (37) were cultured as described (4). The MycER chimaera was activated by exposure of the cells to 2  $\mu$ M  $\beta$ -oestradiol. Insulin was added at 5  $\mu$ g/ml to protect against Mycinduced apoptosis (E.Harrington and G.Evan, personal communication). Stable transfections were performed with calcium phosphate using 10  $\mu$ g of reporter plasmid and 1  $\mu$ g of J6 $\Omega$ puro (38), allowing selection of transfected cells with puromycin.

# **Reporter constructs and CAT assays**

We made CAT reporter constructs by replacing the HindIII–PstI fragment of pBLCAT2 (39) with the sequence 5' AGCTTGCCG-ACCACGTGGGTCGACCACTTCGCATATTAAGGTG-ACGCGTGTGGGCCTCGAACACCGAGCGACCCTGCAG 3' to yield *CcoreG*. The plasmids *TcoreG*, *TcoreA* and *GTA* differ only as indicated in Figure 5A. We assayed CAT enzymatic activity in lysates of transfected cells by liquid scintillation using standard procedures (40).

# RESULTS

# **Binding site selection**

We determined the DNA binding preferences of the Myc/Max and Max/Max dimers using a previously described selection strategy (34). We used the Max2 form for these studies (hereafter termed Max) since it binds DNA more efficiently than Max1 both as a homodimer and as a heterodimer with Myc (12) (B.A. and H.L., unpublished data).

A mixture of *in vitro* translated full-length Myc and Max proteins or Max alone was exposed to a random oligonucleotide pool. Bound oligonucleotides were then co-immunoprecipitated with either Myc or Max-specific antibodies, as appropriate. This permits the isolation of DNA specifically bound to Myc/Max or Max/Max complexes respectively (12), whereas oligonucleotides that do not efficiently bind these dimers are not precipitated. The selected DNA was amplified by PCR and used for a subsequent

#### After 7 rounds:

GAATAGTAGA	CACGTG	GTCCACGAGGC
.GCAGAACAGTAATTTTAG	CACGTG	GT
.GCAGAACAGTAATTTTAG	CACGTG	GT
. TAGAATCAAAACGAGAG	CACGTG	GTA
	CACGTG	
TAAAATCCCAATGAACG		GT
AGCCGAATTAAATTGAC	CACGTG	GTT
TTATAANAAAAAAATTAA	CACGTG	G
ATAGTGATGGTATATAA	CACGTG	СТ
.CCGCCTATCCTTAAAAGC	CACGTG	G
TGAC	CACGTG	GTCGAAGTAACTGACC
GAC	CACGTG	GTCATACCGCATCAT
GATGTGGAGAC	CACGTG	TTTGAAAAA
ATTATTTAAAC	CACGTG	GTCGACACAT
GAC	CACGTG	GCATAATACGACGTTAG
AAGCAC	CACGTG	GTGTAATTTCTGCC
ACAGTCGAC	CACGTG	GTTAACTTATT
	CACGTG	TCGTTGACCGC
	CACGTG	GTCGCTGATAGA
	CACGTG	GCGGGTAA
· · · · · · · · · · · · · · AGGAAAAAA	CACGIG	GCGGGTAA
		AAAATT
		<b>AAATT</b>
		CAAATTCAAAT
		rcatcgta
		<b>FTTATT</b>
		<b>TTGTTTCGGATTCACAGGATCC</b>
GAATTCCTGCA	SCCCGGG.	
After 8 rounds:		
. CTTTTAGGTCTCAAAAAG	CACGTG	GC
GGCTAAATAAGTAG	CACGTG	GTTAT
	CACGTG	ACTAATTCC
AACTATGGAAACGTAGAG	CACGTG	GT
AACTATGGAAACGTAGAG	CACGTG	GAAAATA
GAG	CACGTG	GTCGTCGATGATTGCAT
TACTTGAG	CACGTG	GCCGAAACTCAC
AAGTAC	CACGTG	GCCGACTACATAAT
ATGCACAAC	CACGTG	CTATATTCTG
ACAAACGAATGGC	CACGTG	GTCAGAA
CAGAAGTCCGAC	CACGTG	GTTACTAT
ACATTCGAG	CACGTG	GCATTAAAATG
GAAACGGCTGAG	CACGTG	GTTCATTAT
GAAACGGCTGAG	CACGTG CACGTG	GTTCATTAT GTTCATTACAA
GAAACGGCTGAG		

Figure 1. Nucleotide sequence of DNA molecules selected by the Myc/Max complex. Only the selected bases are shown. The common flanking sequences are: 5'-AGCGGAG and CGACGTC-3'. Sequences are grouped according to the round of selection after which they were cloned (see text) and to the presence or absence of a CACGTG core.

round of selection. Repeated rounds of selection and amplification were performed until specific binding to the selected oligonucleotide pool was evident. Specific DNA binding was measured after each round by immunoprecipitating radiolabelled amplified probe with the Myc or Max-specific antibodies in the presence and absence of the respective cognate immunogenic peptide, and selection was continued for one round after recording a significant level of peptide-sensitive immunoprecipitation (five rounds with Max, eight with Myc/Max). As a positive control we performed immunoprecipitations (with and without peptide) using radiolabelled wt12 probe which specifically binds to both Myc/Max and Max/Max complexes (12). In addition, a parallel selection was performed with unprogrammed reticulocyte lysate to control for any endogenous DNA binding activity, and no enrichment was observed.

## Myc/Max and Max/Max selected sequences

To determine the binding site preferences of the two dimers, we cloned and sequenced the DNA molecules selected in the above experiments. To preserve low affinity sites, oligonucleotides from the penultimate as well as the final rounds of selection were taken (Figures 1 and 2).

#### 5374 Nucleic Acids Research, 1993, Vol. 21, No. 23

After 4 rounds:

. AAAGCTTCGGAAATTAC	CACGTG	ACA
	CACGTG	ACTTCCGTGTT
	CACGTG	ACTTCCGGCGCACC
ACAA	CACGTG	CTCCACTCGAAGTGAC
.CATCAAAGAGTAAAGGA	CACGTG	GCT
ATCGCTGAA	CACGTG	ATTCTGAGGCA
ACGAACCGCGCC	CACGTG	GACAATTG
GAGGAATAGAGGC	CACGNG	TCAGACAGATA
ATGATATCACTAA	CACGTG	ΤΤ
.GAACCTGGAACGGTGAA	CACGTG	ACT
	CACGTG	TTTATACTTACNCAA
CTATGAAT	CACGTG	TNNACTTAAAGTTTGGT
TATGAAG	CACGTG	GTAGTGTTGTTTGC
GCGTAAAGT	CACGTG	GTCCATACAGCCGATAGCGG
TGCATTAACTG	CACGTG	GTCAATAAT
GGCAGTACAGAAAAAG	CACGTG	ΤΤΑ
	CACATG	CACTGATGTT
	CACATG	TAGTTTTTAAAAATT
CAGT	CACATG	CCATTATCACCCCGC
GCCGAAGC'	FAAATTT	AAAATTACCA
CGATAGGATA	<b>FGAAGTT</b>	CAGATTTCGA
TAATTACATAG		
		TATTTGT
ICAIGGCAATA	TAAGCGG	IATTIGI
		ACAGAATGCT
		ICCCTTTGATAAT
AGCTAAA'	<b>FCATTGA</b>	ΓΑΑΑGGTAAATAT
	ACATA	AATAGGTATATATCCT
	TACCGCA	TTTATCTCACA
CCACAAT	<b>FACCGCA</b>	TTTATCTCACA
CCACAAT	<b>FACCGCA</b>	ATTATCTCACA
CCACAAT' GAGATGA	<b>FACCGCA</b>	TTTATCTCACA
CCACAAT	<b>FACCGCA</b>	TTTATCTCACA
CCACAAT" GAGATGA.	FACCGCA' ACTCGTG	TTTATCTCACA ATTTGCGCC
CCACAAT' GAGATGA	<b>FACCGCA</b>	ГТТАТСТСАСА АТТТGCGCC СТ
CCACAAT" GAGATGA. After 5 rounds: GAAATCCATTGAGTGGGT	TACCGCA' ACTCGTG CACGTG	ГТТАТСТСАСА АТТТGCGCC СТ
CCACAAT GAGATGA After 5 rounds: GAAATCCATTGAGTGGGGT	TACCGCA' ACTCGTG CACGTG CACGTG	CT
GAGATGA GAGATGA After 5 rounds: GAAATCCATTGAGTGGGGT TTTCGACTTGGT .TCAAAACACCANCTGAA	TACCGCA ACTCGTG CACGTG CACGTG CACGTG CACGTG	TTTATCTCACAATTTGCGCC
	TACCGCA ACTCGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT TCGTGCGCCC TCGTGCTGCCGTCGCT GT GAGTAATCGAGATAAA
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT TCGTGTGCCCCTCGCT GT GAGTAATCGACGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCTAA
CCACAAT 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT TCGTGTGCCCCTCGCT GT GAGTAATCGACGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCTAA
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT TCGTGTGCCGCCGCT GT GAGTAATCGAGATAAA CTCTACGGGCTAA ACTTCA ACTTCA ACTTCA
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT TCGTGTGCCGCCGCT GT GAGTAATCGAGATAAA CTCTACGGGCTAA ACTTCA ACTTCA ACTTCA
CCACAAT GAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TCAAAACACCANCTGAA CATCGCTATGACTAGTA AAAC GATCGCTATGATACTCA GTACTTAGTAGAC CATCGGTAGTAGAA	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CTGGTGGCCGCCGCCGCT GAGTAATCGAGATAAA CTCTACGTGGCTAA ACTTCA ACATCAT GTGATAAG
CCACAAT 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CTGTGTGCCCGCCC GT GAGTAATCGAGATAAA CTCTACGTGGCCTAA ACTTCA GTGTTAAGA GTAATAGA GTAATAGA ACCTQACGTGA
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CT.GTGTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGGCGCTAA ACTTCA ACTTCA GTGTTAAGA GTAATAGA GTGATAAGA ACCTGCACGGGA ACTGGACGTGA
CCACAAT GAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TTTCGACTTGGT 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CTGTGGCGCCGCGCGCT GAGTAATCGAGATAAA CTCTACGTGGCTAA ACTTCA ACATCGAT GTGATAGG ACCTCGACGTGA ACTGAC ACTCGACGTGA ACTCGACGTGA ACTGCACGTGGA ACTGC CACGTG CAAT A
	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CT.GTGTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGGCGCTAA ACTTCA ACTTCA GTGTTAAGA GTAATAGA GTGATAAGA ACCTGCACGGGA ACTGGACGTGA
CCACAAT CAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT .TCAAAACACCANCTGAA AAC AAC CATCGTAGTACTCA 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT. CT. CT. GT. GT. GT. GT. GTAATCGACGTCGCT. CTCTACGGCGTAAA. ACTTCA. ACATGGAT. GTGTTAAGA. GTGTTAAGA. GTGTTAAGA. ACCTGACCGTGA. ACCTG CACGTG CAAT. A. TTTCT. CTTTCT.
CCACAAT 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT TCGTGGCCGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCCGACAA ACTTCA ACTTCA GGTAATAGA GTGTTAAGA GTGTAAGA GTCATAGA ACTGCA
CCACAAT 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT
CCACAAT CAGAATGA After 5 rounds: GAAATCCATTGGACTGGGT TTCGACTTGGT .TCAAAACACCANCTGAA AAC GATCGCTATGTACTCA GTACTATATATGATAC TTTAATATGATACA CAGTCGGTAGAA AGCACAATAGTCAGCC CCACAATAGTCAGCC CACAATAGTCAGCC CCACAATAGTCAGCC CCACAATAGTCAGCC CCACAATAGTCAGCC CCACAATAGTCAGCC CCACAATAGTCAGCC CCACAATAGTCAGCC 	CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CT CGT.GTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCTAA ACTTCA ACTTCA GTGATAACA GTGATAACA ACTGCA ACTGCA ACTGCA ACTGCA ACTGCA ACTGCA CTCA CTCA CTCA CTTCTCT ACTG CTTCTCTATT
CCACAAT GAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT CCATAGCACCACTGGA CAAACACCANCTGAA CATCGCTATGTACTCA AAAC GATCGCTATGTACTCA GATCGCTATGATAGTAGA CCACTCGGTAGAA ACCTCAAGC GTACAATAGTTCACGC CCAAAATAGTTCACGC CCAAAATAGTTCACGC GAGTGTAANCNGGAA CCCAACAGAGAG TTCCAGCAAAGAG	CACGGCA ACTCGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT
CCACAAT CCACAAT After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT AAACACCANCTGAA AAC AAC GTACTAGTAGTACTCA TTTAATATTGATACC GTACTTAGTAGAC ACCTCAGTAGAA ACGCGAGAAA ACCTCAAGC ACATCAGTAGTCAGC 	CACGTG CA	CT. CT. CT. GAGTGACGCCGCCGCT. GT. GTATCGACGATAAA. CTCTCACGGCTAAA. ACTTCA. ACATGAT. GTGATAAGA. GTGATAAGA. ACTG CACGTG CAAT. A. ACTG CACGTG CAAT. A. CTCA. CTCATAA. ACTG CACGTG CAAT. A. CTCAA. CTCACTCATT. TTTGGACC. TTTGAAC.
CCACAAT CCACAAT After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TTTCGACTTGGT AAAC GATCGCTATGTACTCA 	CACGGCA ACTCGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG CACGTG	CT CT CT.CGTGTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGGGGCTAA ACTTCA ACTTCA GTGTTAAGA GTGATAAGA GTGATAAGA ACTGCGACGTGA ACTGCGACGTGA ACTGCCACT ACTACA CTTCTCT ACTAC GACT GAC CTTCTCTATT TTTGGACC TTTAANCAC GTCCGATNGAC
CCACAAT CAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TTTCGACTTGGT AAAC AAC CATCGCTATGTACTCA CATCGTAGTACTCA AAC 	CACGTG CA	CT. CT. CT. GAGTGACGCCGCCGCT. GT. GTATCGACGATAAA. CTCTCACGGCTAAA. ACTTCA. ACATGAT. GTGATAAGA. GTGATAAGA. ACTG CACGTG CAAT. A. ACTG CACGTG CAAT. A. CTCA. CTCATAA. ACTG CACGTG CAAT. A. CTCAA. CTCACTCATT. TTTGGACC. TTTGAAC.
CCACAAT 	CACGTG CA	CT CT CT.CGTGTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGGGGCTAA ACTTCA ACTTCA GTGTTAAGA GTGATAAGA GTGATAAGA ACTGCGACGTGA ACTGCGACGTGA ACTGCCACT ACTACA CTTCTCT ACTAC GACT GAC CTTCTCTATT TTTGGACC TTTAANCAC GTCCGATNGAC
CCACAAT CCACAAT After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TTTCGACTTGGT AAAC .GATCGCTATGTACTCA 	CACGTG CA	CT
CCACAAT CCACAAT After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT TTTCGACTTGGT 	CACGGCA ACTCGTG CACGTG	CT CT CT CGT.GTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCTAA ACTTCA ACTTCA GTGTTAAGA GTGTTAAGA GTGTTAAGA TTTCT ACTG CACGTG CAAT ACTG CACGTG CAAT ACTGACGTGACAT CTTCTCTTATT TTTCGACC CTTCTCTATT TTTGGACC GTCGGTCAATATCT GTCGGTCCAGC GTTGGTCCCAGC
CCACAAT CAGATGA After 5 rounds: GAAATCCATTGAGTGGGT TTTCGACTTGGT .TCAAAACACCANCTGAA AAC TTTAATATTGATACC TTTAATATTGATACC TTTAATATTGATACC TTTAATATTGATACC TTTAATATTGATACC ACAAAGAC ACAAAGAC 	CACGTG CA	CT CT CT CGT.GTGCCGTCGCT GAGTAATCGACGATAAA CTCTACGTGGCTAA ACTTCA ACTTCA GTGTTAAGA GTGTTAAGA GTGTTAAGA TTTCT ACTG CACGTG CAAT ACTG CACGTG CAAT ACTGACGTGACAT CTTCTCTTATT TTTCGACC CTTCTCTATT TTTGGACC GTCGGTCAATATCT GTCGGTCCAGC GTTGGTCCCAGC

Figure 2. Nucleotide sequence of DNA molecules selected by the Max/Max complex. See Figure 1 legend for details.

Both the Myc/Max and the Max/Max complexes selected sequences containing the previously described hexanucleotide C-ACGTG (73% and 90% for Myc/Max in the penultimate and final rounds respectively, and 55% and 91% for Max/Max). To investigate the base composition at positions flanking the core we must allow for the fact that the orientation of the selected molecules is indeterminate since the core sequence CACGTG is a palindrome. To account for this we considered all half-sites independently. For example, the sequence AAAG CACGTG TTTG was considered as one AAAG CAC half site and one CAAA CAC half site. All the half sites were aligned and the number of occurrences for each base at each position was determined as shown in Figure 3 and used to derive a half-site composed of two mirror-image half-sites as indicated.

Surprisingly the two complexes exhibited very different preferences at the -1 position. Of the sequences containing the CACGTG motif, only one of the 76 selected by Myc/Max had a T residue at this position, demonstrating that the TCAC half site is strongly disfavoured. In contrast, the Max homodimers show no significant preference at this position.

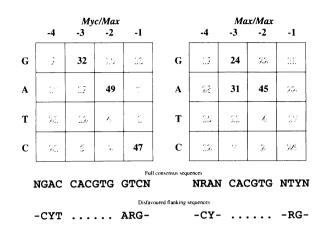


Figure 3. Consensus sequences for positions flanking the CACGTG core in binding sites selected by the Myc/Max and Max/Max complexes. The compilation is from the sequences shown in Figures 1 and 2. The sequences containing a CACGTG motif from both the final and penultimate rounds were considered in each case. The table shows the number of occurrences for each base at successive positions from the core; thus, -1 refers to the position adjacent to the core, -2 to that two residues away and so on.

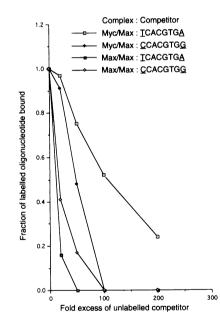


Figure 4. Myc/Max but not Max/Max discriminates against TCACGTGA. In vitro translated Myc and Max proteins were incubated together with both the radiolabelled probe CGCCGA<u>C</u>CACGTG<u>G</u>TCCCTC and increasing amounts of either identical unlabelled oligonucleotide or a similar oligonucleotide differing only at the underlined positions, as indicated. Myc/Max and Max/Max complexes were resolved by gel electrophoresis as previously described (12). The amount of radioactive probe in each complex was quantified and normalised to the amount of probe in the asme complex in the absence of competitor.

Both complexes have a clear preference for purines (R) at positions -2 and -3. The -2 position is especially well defined: A is the preferred residue but a G is allowed while either pyrimidine (Y) is strongly disfavoured. At -3, the preference appears less strong and only a C is significantly disfavoured. The consensus binding site of the Myc/Max dimer is <u>RACC</u>-

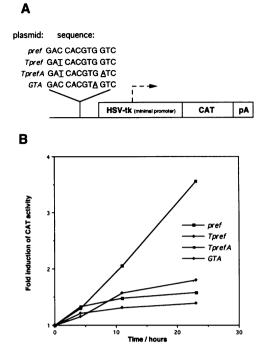


Figure 5. Discrimination against core-flanking sequences by Myc in transactivation assays. (A) Schematic representation of the CAT-reporter plasmids used. (B) CAT activity from the various reporters in Rat-1 MycER cells following MycER activation. Cells stably transfected with each reporter construct were grown to confluence and arrested for 48 hours in 0% serum prior to activation of MycER with  $\beta$ -oestradiol for the indicated periods of time. CAT activity was assayed from equal numbers of cells. The fold activation of CAT activity after hormone stimulation is indicated.

ACGTG<u>GTY</u>, while the consensus for Max homodimers is <u>RANCACGTGNTY</u>. Neither complex shows any preference for particular bases at the -4 position (Figure 3) nor at positions further out.

Among the sequences lacking a CACGTG, the only obvious feature was a CACATG core found in three out of 52 sequences selected by Max, none of them in the final round of selection (Figure 1). Max/Max and Myc/Max dimers can bind to this sequence in gel retardation experiments, although with lower affinity than for the CACGTG sequence (11 and data not shown). Moreover, multiple CACATG sites can mediate Myc-induced transactivation (21). Computer analysis of those sequences lacking either CACGTG or CACATG motifs failed to reveal any recurrent feature other than AT-rich sequences similar to those brought down by the unprogrammed lysate control (data not shown). Specifically, the sequence TCTCTTA reported as being associated with Myc (41, 42) was not found.

To confirm the difference in binding preference of the Myc/Max and Max/Max complexes, we analysed the DNA binding activities of these complexes in gel retardation assays as previously described (12) and performed competition experiments. When using a radiolabelled <u>CCACGTGG</u> probe, a 100-fold excess of <u>TCACGTGA</u> was only able to compete weakly for the Myc/Max complex in comparison with competition by the <u>CCACGTGG</u> sequence itself, while both sequences competed efficiently for the Max/Max complex (Figure 4). Moreover, when using <u>TCACGTGA</u> as a probe, we only

detect the Max/Max complex and not the Myc/Max complex (data not shown).

## Sequences flanking the hexanucleotide core affect transactivation by Myc/Max in vivo

We next wanted to determine whether the difference in binding specificity of Myc/Max for CCACGTGG (termed CcoreG) and TCACGTGA (termed TcoreA) in vitro may be reflected in Myc/Max dimer activity in vivo. Therefore, we assessed the ability of Myc to activate transcription from promoters driven by these sequences. We measured Chloramphenicol Acetyl Transferase (CAT) expression from reporter plasmids (Figure 5A) in which the minimal Herpes Simplex Virus-Thymidine Kinase promoter was linked to either CcoreG, TcoreA, the single half-site change TcoreG, or the non-binding core mutant CC-ACGTAG (12) (termedGTA). We stably co-transfected the reporters with a puromycin resistance gene into Rat-1 cells expressing MycER, a  $\beta$ -oestradiol-inducible Myc-oestrogen receptor fusion protein (37) and pooled approximately 1500 puromycin resistant colonies for each reporter. Exposure of the transfected cells to  $\beta$ -oestradiol results in an increase in the rate of CAT synthesis in the pool containing the CcoreG reporter, while the GTA, TcoreG and TcoreA reporters were not significantly transactivated by MycER (Figure 5B).

## DISCUSSION

We report the complete dodecameric DNA binding sites for the Myc/Max and Max/Max complexes derived from selectionamplification experiments with a random oligonucleotide pool and full-length in vitro-translated Myc and Max2 proteins. We find that the binding preferences of the two complexes are distinct, although they both select the same CACGTG core. The sequences selected by Myc/Max form a much more restricted set than those for Max/Max since the former discriminates against T residues at the -1 position relative to the hexanucleotide core while the latter does not. It remains to be determined whether the other forms of Max have the same DNA binding specificity as Max2. However, the discrimination against T at position -1 by Myc/Max heterodimers is most likely true with all foms of Max expressed in cells, since Myc activates reporter genes containing CcoreG, but not TcoreG or TcoreA sequences in vivo (see Figure 5).

One might expect that a difference in DNA site preference between the Myc/Max and Max/Max complexes would be due to a difference in the half-site preference of Myc compared to Max. If this were the case then sequences with a T 5' to the hexanucleotide core (or an A 3') such as TCACGTGG (or CC-ACGTGA) should still be efficiently recognised since there would still be one half-site recognised by Myc. Only sequences with both a 5' T and a 3'A which exclude Myc from both sides should be discriminated against. However in the sequences selected by Myc/Max, those with a single Myc-incompatible half-site are strongly discriminated against (only 1 out of 76). Furthermore, the TcoreG reporter is no more able than the TcoreA reporter to mediate Myc-driven transcription in vivo. These results suggests that the proteins in a dimer do not recognise half-sites independently of one-another. Confirmation of this prediction must await resolution of the structure of DNA-bound Myc/Max dimers, for comparison with the solved structure of the DNAbound Max/Max bHLH-Z domain (20).

## 5376 Nucleic Acids Research, 1993, Vol. 21, No. 23

We also show that the preferred binding sites of Myc/Max and Max/Max complexes are very similar at the flanking positions -2 and -3, where purines are strongly favoured. Moreover, it becomes clear from the sequences shown in Figure 3 that certain residues are particularly disfavoured. These are T and C at the -2 position and C at position -3. According to these rules only 8% of CACGTG containing sequences are likely to be high affinity Myc/Max targets and 14% likely to be high affinity Max/Max targets.

Interestingly, our binding consensus for Myc/Max dimers is in agreement with studies with *in vitro* translated truncated Myc protein (33), which most probably forms dimers with endogenous Max in the lysate (12). Additionally, bacterially expressed GST-Myc bHLH-Z fusion peptides (31) and Myc expressed in baculovirus or purified from Chinese Hamster Ovary cells (32) (all of which probably bind DNA as homodimers) also bind to similar sequences. Therefore the Myc homodimer seems to have DNA binding preferences similar to those of Myc/Max. However Myc homodimers are highly unlikely to exist *in vivo* (7,12) and, when formed, are non-functional in transcriptional activation (7) and cellular transformation assays (14) owing to a decreased DNA binding efficiency.

There is a precedent for discrimination against a flanking 5'T or 3'A being used to determine the specific targets for different transcription factors in yeast (43). Two unrelated bHLH proteins (CPF1and PHO4) both recognise the sequence CACGTG, yet PHO4 fails to substitute for a CPF1 null mutation in vivo since it cannot bind to the CPF1 target sites in which the CACGTG is flanked by a 5' T or a 3' A (or both). It has been suggested (43) that inhibition of PHO4 DNA binding by a flanking T is due to steric hindrance between a glutamic acid residue in the basic region and the methyl group of the T, while CPF1 has a smaller aspartic acid residue at this position. The validity of this hypothesis for Myc and Max remains to be demonstrated. However, in apparent agreement, Myc has an arginine residue with a very long carbon side-chain at the corresponding position while Max has an alanine (amino acid position +3 according to Fisher's nomenclature (44)).

In a similar fashion, the human E12 and E47 proteins discriminate against 5' T residues while their binding partner MyoD does not (30). Thus discrimination between nucleotides flanking the core sequence recognition motif may be a common regulatory mechanism in bHLH-containing transcription factors.

Gene regulation by Max remains an open question. High levels of Max in transfections inhibit transactivation as well as cellular transformation by Myc (7, 14, 18, 21-23, 28, 29). In mammalian cells, this may due to competitive displacement of Myc/Max from DNA sites by Max, either as a homodimer or as a heterodimer with one of its recently discovered alternative partners Mad and Mxi-1 (45, 46). However, our studies in yeast demonstrate that Max homodimers can indeed compete with Myc/Max *in vivo* (7). Thus, our data suggest that as well as repressing genes targeted by Myc/Max, Max homodimers, being less discriminatory, may also repress genes not regulated by Myc.

# REFERENCES

- Penn,L.J.Z., Laufer,E.M. and Land,H. (1990) In Jones,N. (eds), Seminars in Cancer Biology. Saunders, London, Vol. I, pp. 69-80.
- 2. Luscher, B. and Eisenman, R.N. (1990) Genes Dev. 4, 2025-2035.
- Askew,D.S., Ashmun,R.A., Simmons,B.C. and Cleveland,J.L. (1991) Oncogene 6, 1915-1922.

- Evan,G.I., Wyllie,A.H., Gilbert,C.S., Littlewood,T.D., Land,H., Brooks,M., Waters,C.M., Penn,L.Z. and Hancock,D.C. (1992) Cell 69, 119-128.
- 5. Blackwood, E.M. and Eisenman, R.N. (1991) Science 251, 1211-1217.
- 6. Prendergast, G.C., Lawe, D. and Ziff, E.B. (1991) Cell 65, 395-407.
- Amati, B., Dalton, S., Brooks, M.W., Littlewood, T.D., Evan, G.I. and Land, H. (1992) Nature 359, 23-426.
- Berberich, S., Hyde-DeRuyscher, N., Espenshade, P. and Cole, M.D. (1992) Oncogene 7, 775-779.
- 9. Blackwood, E.M., Luscher, B. and Eisenman, R.N. (1992) Genes De. 6, 71-80.
- 10. Davis, L.J. and Halazonetis, T.D. (1993) Oncogene B, 125-132.
- 11. Kato,G.J., Lee,W.M.F., Chen,L. and Dang,C.V. (1992) Genes Dev. B, 81-92.
- 12. Littlewood, T.D., Amati, B., Land, H. and Evan, G.I. (1992) Oncogene B, 1783-1792.
- Reddy,C.D., Dasgupta,P., Saikumar,P., Dudek,H., Rauscher,F. and Reddy,E.P. (1992) Oncogene B, 2085-2092.
- Amati, B., Brooks, M.W., Levy, N., Littlewood, T.D., Evan, G.I. and Land, H. (1993) Cell B, 233-245.
- Crouch, D.H., Fisher, F., Clark, W., Jayarama, P., Goding, C.R. and Gillespie, D.A.F. (1993) Oncogene B, 1849-1855.
- Blackwood, E.M., Kretzner, L. and Eisenman, R.N. (1992) Curr. Opin. Genet. Dev. 2, 227-235.
- Torres, R., Schreiber, A.N., Morgenbesser, S.D. and DePinho, R.A. (1992) Curr. Opin. Cell Biol. 4, 468-474.
- Makela, T.P., Koskinen, P.J., Västrik, I. and Alitalo, K. (1992) Science 256, 373-377.
- Västrik, I., Koskinen, P.J., Alitalo, R. and Makela, T.P. (1993) Oncogene 8, 503-507.
- Ferré-d'Amaré, A., Prendergast, G.C., Ziff, E.B. and Burley, S.K. (1993) Nature 363, 38-45.
- 21. Kretzner, L., Blackwood, E.M. and Eisenman, R.N. (1992) Nature 359, 426-429.
- Gu, W., Cechova, K., Tassi, V. and Dalla-Favera, R. (1993) Proc. Natl. Acad. Sci. USA 90, 2935 – 2939.
- Amin,C., Wagner,A.J. and Hay,N. (1993) Mol. Cell Biol. 13, 383-390.
   Reisman,D., Elkind,N.B., Roy,B., Beamon,J. and Rotter,V. (1993) Cell
- Growth Differ. 4, 57-65.
  25. Benvenisty, N., Leder, A., Kuo, A. and Leder, P. (1992) Genes Dev. 6, 2513-2523.
- 25b. Amati, B., Littlewood, T.D., Evan, G.I. and Land, H. (1993) *EMBO J.* in press.
- Stone, J., de Lange, T., Ramsay, G., Jakobvits, E., Bishop, J.M., Varmus, H. and Lee, W. (1987) Mol. Cell Biol. 7, 1697-1709.
- Kato,G.J., Barrett,J., Villa,G.M. and Dang,C.V. (1990) Mol. Cell Biol. 10, 5914-5920.
- 28. Mukherjee, B., Morgenbesser, S.D. and DePinho, R. (1992) Genes Dev. 6, 1480-1492.
- Prendergast, G.C., Hopewell, R., Gorham, B.J. and Ziff, E.B. (1992) Genes Dev. 6, 2429-2439.
- 30. Blackwell, T.K. and Weintraub, H. (1990) Science 250, 1104-1110.
- Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N. and Weintraub, H. (1990) Science 250, 1149-1151.
- Papoulas, O., Williams, N.G. and Kingston, R.E. (1992) J. Biol Chem. 267, 10470-10480.
- 33. Halazonetis, T.D. and Kandil, A.N. (1991) Proc. Natl. Acad. Sci. USA 88, 6162-6166.
- 34. Pollock, R. and Treisman, R. (1990) Nucleic Acids Res. 18, 6197-6204.
- 35. Prendergast, G.C. and Ziff, E.B. (1991) Science 251, 186-189.
- Evan, G., Lewis, G., Ramsay, G. and Bishop, J. (1985) Mol. Cell Biol. 5, 3610-3616.
- 37. Eilers, M., Picard, D., Yamamoto, K.R. and Bishop, J.M. (1989) *Nature* **340**, 66–68.
- 38. Morgenstern, J.P. and Land, H. (1990) Nucleic Acids Res. 18, 1068.
- 39. Luckow, B. and Schâtz, G. (1987) Nucleic Acids Res. 15, 5490.
- 40. Seed, B. and Sheen, J.Y. (1988) Gene 67, 271-277.
- 41. Negishi, Y., Iguchi, A.S. and Ariga, H. (1992) Oncogene 7, 543-548.
- 42. Iguchi-Ariga,S.M.M., Okazaki,T., Itani,T., Ogata,M., Sato,Y. and Ariga,H. (1988) Embo J. 7, 3135-3142.
- 43. Fisher, F. and Goding, C.R. (1992) EMBO J. 11, 4103-4109.
- 44. Fisher, F., Jayaraman, P.S. and Goding, C.R. (1991) Oncogene 6, 1099-1104.
- 45. Ayer, D.E., Kretzner, L. and Eisenman, R.N. (1993) Cell 72, 211-222.
- 46. Zervos, A.S., Gyuris, J. and Brent, R. (1993) Cell 72, 223-232.