RESEARCH COMMUNICATION

A network of specific minor-groove contacts is a common characteristic of paired-domain–DNA interactions

Lucia PELLIZZARI*, Dora FABBRO*, Renata LONIGRO*, Roberto DI LAURO† and Guiseppe DAMANTE*

*Dipartimento di Scienze e Tecnologie Biomediche, Università di Udine, via Gervasutta 48, 33100 Udine, and †Stazione Zoologica 'A. Dohrn', Villa Comunale 1, 80123 Napoli, Italy

Pax proteins are a family of transcription factors conserved during evolution and able to bind specific DNA sequences through a domain called a 'paired domain'. The DNA-binding specificity of the Pax-8 paired domain was investigated. Site-selection experiments indicate that Pax-8 binds to a consensus sequence similar to those bound by Pax-2 and Pax-5. When consensus sequences of various paired domains are observed in light of recent structural studies describing paired-domain–DNA interaction [Xu, Rould, Jun, Desplan and Pabo (1995) Cell **80**, 639–650], it appears that base-pairs contacted in the minor groove are conserved, while most of the base-pairs contacted in the major groove are not. Therefore a network of specific minor groove contacts is a common characteristic of paired-domain–DNA interactions. The functional importance of such a network was successfully tested by analysing the effect of consensus-based mutations on the Pax-8 binding site of the thyroglobulin promoter.

INTRODUCTION

The Pax gene family encodes several transcription factors showing sequence similarity to the *Drosophila* segmentation gene *paired* (Prd), containing a highly conserved paired box [1]. Pax genes have been found in vertebrates [2] and encode proteins important for development which are expressed in a tissue and time-specific manner during embryogenesis [3]. Several mutations of mouse and human Pax genes, within the paired box, are known to be associated to congenital disorders [4]. Pax genes also appear to promote oncogenesis [5].

The paired box encodes the paired domain capable of sequence-specific DNA recognition [6-8]. The crystal of the Prd paired domain–DNA complex revealed that paired domains fold in two separate subdomains (N and C domains), each of which possesses a helix–turn–helix (H–T–H) motif [9]. The N-domain H–T–H recognizes DNA through major-groove contacts, while a β-turn (amino acids 13–16 in the paired domain of Prd) as well as the linker to the C domain recognize DNA through minor-groove interactions [9]. Although in the Prd–DNA crystal the C domain does not contact DNA, several results suggest that this domain could play a role in the interaction with target sites ([9] and references cited therein).

The Pax-8 paired domain is very similar to those of Pax-2 and Pax-5 [10]. Pax-8 is expressed in kidney, thyroid gland and, only during development, in some areas of the central nervous system [11]. In co-transfection experiments, Pax-8 is able to activate promoters of thyroglobulin (Tg) and thyroperoxidase genes [12].

In the present study we define the consensus for the DNA sequences recognized by the Pax-8 paired domain. The superimposition of consensus sequences of different paired domains to the DNA-binding mode of these proteins, found by crystallographic analysis [9], reveals the conservation of a network of specific minor-groove interactions. Mutants of the Pax-8-binding site of the Tg promoter demonstrate the relevance of such a specific contacting network.

EXPERIMENTAL

Preparation of Pax-8- and Prd-paired-domain–glutathione S-transferase (GST) fusion protein

Escherichia coli XL-1 Blue cells were transformed with the expression vectors pGEX Pax-8 Pb and pGEX Prd Pb. Production of the fusion protein was induced by isopropyl thiogalactoside (1 mM). Total bacterial extracts were incubated with GSH–agarose beads as described by Smith and Johnson [13]. Proteins were not eluted from the beads. Once prepared, the bead-bound proteins could be stored for several days at 4 °C.

Binding-site-selection procedure

The 64 bp oligonucleotide CD1 contains 18 randomly degenerate bases and was used as an initial template to perform the binding-site-selection procedure [sequence: 5′-CATGAATTCTCCTATACTGACTC$(N)_{18}$AGAACTGTATCGATGAATTCCAC-3′]. CD1 sequences were made double-stranded by annealing the CD3 oligonucleotide (sequence: 5′-GTGGAATTCATCGATACAGT-3′) and extending it by using Klenow polymerase. Once rendered double-stranded, CD1 was used for the initial round of binding-site selection. This round was performed by mixing 300 ng of the double-stranded template with 10–50 ng of fusion protein linked to 20 μl of GSH–agarose beads. The binding reaction was performed in a 100 μl final volume of 20 mM Tris (pH 7.5)/75 mM KCl/1 mM dithiothreitol/50 μg/ml BSA/10% glycerol/2 μg/ml poly(dI-dC). Samples were gently rocked at 4 °C for 1 h, and then centrifuged for 1 min at 12000 rev./min ($r_{av.} = 5$ cm). The bead pellet was washed with 0.8 ml of ice-cold buffer [20 mM Tris (pH 7.5)/0.1 mM EDTA/75 mM KCl] with gentle shaking for about 1 min. The fusion-protein beads were then resuspended in 30 μl of water, and bound DNA was eluted by heating at 98 °C for 10 min. After heating, the samples were quickly centrifuged and the supernatant was collected. A 10 μl portion of supernatant

Abbreviations used: Tg, thyroglobulin; GST, glutathione S-transferase; H–T–H, helix–turn–helix; Prd, *Drosophila* segmentation gene *paired*.

* To whom correspondence should be addressed.

was used as template in the PCR reaction. Primers for PCR reactions were CD3 and CD2 (sequence: 5′-CATGAATTC-TCCTATACTGA-3′). PCR reactions were performed in 10 mM Tris (pH 8.3)/50 mM KCl/2 mM $MgCl_2$/10 % DMSO, containing 200 μM of each dNTP and 100 pmol of each primer (CD2 and CD3) in total volume of 100 μl. 'Hot start' was performed by heating samples for 5 min at 94 °C before the addition of 2.5 *Taq* polymerase units/sample. In all, 20 cycles of 94 °C (45 s), 44 °C (30 s) and 72 °C (30 s) were performed with an additional extension of 10 min at 72 °C at the end. A 20 μl sample of the amplified material was used in the further rounds of selection. Seven rounds of selection and amplification were performed. After the last PCR, the DNA was phenol/chloroform-extracted, ethanol-precipitated and cloned in the PCR cloning vector PCR-Script (Stratagene). Plasmid DNA clones were sequenced. In order to measure the relative binding activity of each of the selected clones to Pax-8 protein, the insert sequences were separated from the vector sequence by *Eco*RI digestion (site located in the primer sequences) and gel purification. Equal amounts of purified sequences were end-labelled by Klenow fill-in and used in gel-retardation assays with an equal amount of crude bacterial extracts with or without Pax-8–GST protein. The binding activity to Pax-8–GST of each selected sequence was expressed as fraction of the C/Pax-8–GST binding activity, considered arbitrarily as 1.0.

Gel-retardation assay, orthophenanthroline footprinting and methylation interference

Gel-retardation assay was performed, incubating protein and DNA in a buffer containing 20 mM Tris/HCl (pH 7.6)/75 mM KCl/0.25 mg/ml BSA/5 mM dithiothreitol (DTT)/5 μg/ml poly(dI-dC)/10 % glycerol for 30 min at room temperature. Protein-bound DNA and free DNA were separated on native 7 % polyacryamide gel run in 0.5 × TBE (45 mM Tris/borate/45 mM boric acid/1 mM EDTA) for 2 h at 4 °C. Gels were dried, exposed to X-ray films and the bands were quantified by densitometric scanning of the autoradiogram using a LKB laser densitometer.

Orthophenanthroline-copper (Cu^{2+}) footprinting was carried out as described by Kuwabara and Sigman [14].

Methylation interference experiments were performed as described in [15], using as probes dimethyl sulphate-treated oligonucleotides. Protein-bound and free DNAs were separated by preparative PAGE, identified by autoradiography and eluted from the gel. After chemical cleavage the products were separated on 16 % denaturing polyacrylamide gels and revealed by autoradiography.

RESULTS AND DISCUSSION

In order to identify the optimal sequence recognized by the Pax-8 paired domain, this domain was expressed as a fusion protein with GST and used to select specific sequences from a pool of 64-bp oligonucleotides each containing a core of 18 random nucleotides. Sequences that bound specifically were amplified by PCR and subjected to further rounds of selection. Conditions of selection were chosen on the basis of the binding activity of model oligonucleotides that in gel-retardation assay are bound (C sequence of Tg promoter; [12]) or not (BS2 sequence; [15]) by Pax-8 (results not shown). After seven cycles, selected oligonucleotides were cloned, and both the sequence and the relative binding affinity were determined [16]. Results are shown in Figure 1(a). Most of the selected sequences show a high affinity for Pax-8. In fact 23 out of 35 sequences are recognized with an affinity higher than, or equal to, that observed for the C sequence of the Tg promoter. Sequences were manually aligned

(a)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Relative Binding Affinity
1	c	t	t	c	T	t	c	T	c	C	a	a	G	A	A	c	t	g	0,2
2	t	g	a	c	T	c	a	T	G	C	G	T	a	g	A	g	a	t	10,0
3	t	c	a	c	T	g	g	T	G	C	a	T	t	g	t	t	c	g	1,0
4	g	t	t	g	a	a	t	a	a	C	G	T	G	g	A	g	t	c	5,4
5	a	g	c	c	c	a	g	T	G	C	G	T	c	t	G	c	t	c	7,0
6	t	c	t	a	T	t	t	C	t	a	t	T	G	A	t	c	t	a	0,2
7	t	c	a	g	T	t	a	g	t	C	a	T	t	A	G	a	a	c	5,0
8	t	a	c	t	T	g	c	C	G	a	G	T	G	A	G	t	c	a	0,2
9	t	t	t	a	g	a	a	C	c	C	G	g	G	A	G	t	c	a	0,2
10	c	t	c	t	c	g	g	C	G	g	G	T	G	A	G	t	t	g	0,9
11	c	t	a	t	T	g	a	a	G	C	G	T	G	A	A	g	t	t	18,0
12	t	g	a	c	T	c	a	a	G	C	a	T	G	A	G	a	g	g	7,0
13	g	c	a	a	c	a	c	T	G	C	G	T	a	A	G	a	c	g	6,0
14	a	c	a	g	T	t	c	T	G	C	G	T	G	A	A	t	t	c	3,0
15	t	t	g	a	c	a	c	g	G	g	G	T	G	A	A	t	c	a	0,5
16	a	c	g	g	T	c	a	T	t	C	G	a	G	t	A	g	a	a	7,0
17	a	c	c	g	T	a	t	a	G	a	t	T	G	c	t	a	t	a	1,0
18	t	a	g	a	T	g	t	C	G	t	G	g	G	A	t	c	g	a	0,8
19	a	c	t	g	a	c	t	C	G	g	G	T	G	g	G	c	a	g	7,0
20	t	t	c	t	a	c	t	C	G	a	a	T	G	A	c	c	g	t	6,0
21	a	g	t	c	c	t	c	T	t	C	a	T	G	A	G	t	c	a	4,7
22	t	c	a	a	T	c	g	a	G	a	G	T	G	A	t	t	a	c	5,8
23	g	a	c	t	c	c	a	T	G	g	t	T	G	c	A	t	t	g	3,6
24	a	c	a	g	T	t	c	T	G	t	G	T	a	c	c	a	t	c	0,8
25	t	c	t	g	c	a	a	T	G	a	G	T	G	g	t	t	t	a	2,0
26	c	c	a	t	g	g	t	T	G	g	a	T	t	A	c	c	t	c	2,0
27	a	c	a	t	T	a	a	T	G	C	c	c	G	A	G	t	c	a	2,0
28	a	c	t	g	a	c	t	C	G	a	G	T	G	g	A	a	t	c	2,0
29	g	g	c	c	a	t	t	C	c	C	a	g	G	A	G	t	c	a	1,5
30	t	g	a	c	T	c	g	a	G	C	G	T	G	t	G	t	a	g	0,3
31	a	g	a	t	T	g	a	T	c	C	a	T	G	A	c	a	t	g	1,5
32	t	g	g	t	T	t	t	a	a	a	c	T	G	A	G	t	c	a	0,5
33	t	a	g	a	T	g	t	C	G	t	G	g	G	A	t	c	g	a	1,0
34	t	a	a	g	a	a	t	C	c	a	a	T	G	A	G	a	g	t	0,7
35	c	c	a	g	a	t	a	g	a	g	G	T	t	A	A	g	a	a	0,8
Frequency:	T		7	8	9	**19**	9	12	**14**	4	3	3	**28**	4	3	7	14		
	C		15	7	8	7	9	7	**11**	5	**17**	2	1	1	3	4	8		
	G		8	5	11	2	8	5	3	**23**	6	**20**	4	**27**	6	**14**	5		
	A		5	15	7	7	9	11	7	3	9	10	2	3	**23**	**10**	8		
Consensus:		_	_	_	T	_	_	T C	G	C	G	T	G	A	G A	_	_	_	

(b)

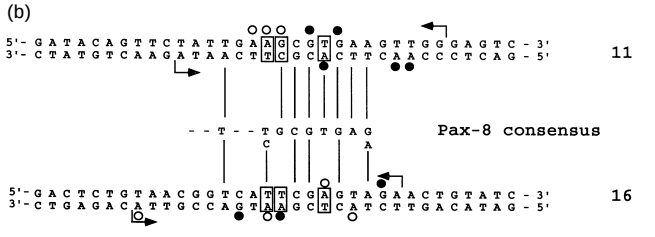

Figure 1 Determination of the consensus sequence for the Pax-8 paired domain

(**a**) Pax-8 selected sequences are aligned for the best fit. At the bottom, the consensus sequence derived by statistical analysis of the base frequency of each position is reported. Base numbering is used according to that of the paired-domain–DNA crystal structure description [9]. Capital letters indicate identity with the consensus. On the right the binding affinity of each sequence is expressed in comparison with the C sequence, considered arbitrarily as 1.0. (**b**) Schematic representation of results of orthophenanthroline footprinting and methylation interference experiments, carried out on two selected sequences (11 and 16). Arrows indicate protection borders obtained by orthophenanthroline footprinting. Contacts mapped by methylation interference are indicated by circles (○, indicate low-strength interactions; ●, indicate high-strength interactions). Boxed base-pairs indicate positions showing contacts in both sequences. Between the two sequences analysed, Pax-8 consensus is reported. Bars indicate the homologies existing between the consensus and the selected sequences.

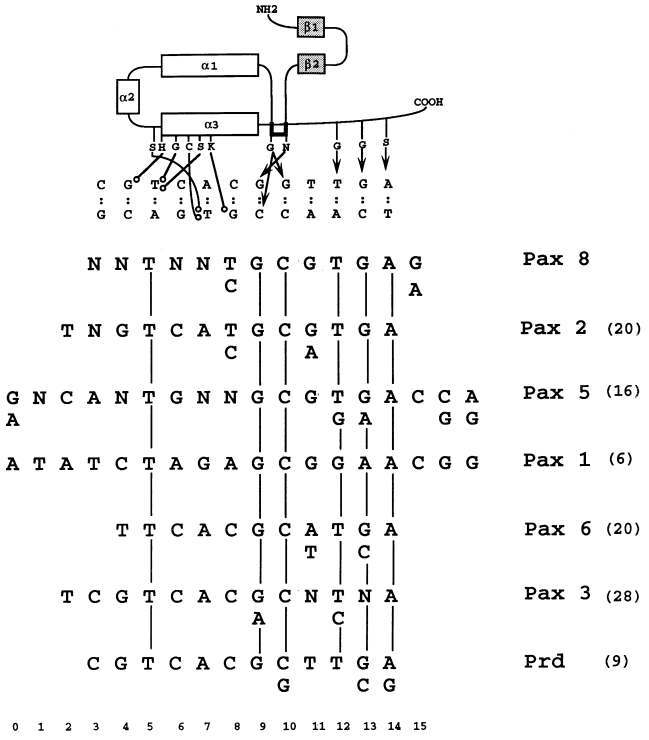

Figure 2 Comparison of consensus sequences of different paired domains

At the top of the Figure is drawn a schematic representation of protein–DNA contacts described in the crystallographic analysis of the Prd-paired-domain–DNA complex [9]. Empty boxes indicate α-helices, shaded boxes indicates β-sheets and a thick line indicate a β-turn. Contacting amino acids are shown by single-letter code. Only direct amino acid–base contacts are shown. Empty circles indicate major groove contacts while filled arrows indicate minor groove contacts. This scheme is aligned to all known consensus sequences for paired-domain proteins (top strands only are shown). Vertical lines between consensus sequences indicate conserved base-pairs. Numbering of the positions is shown at the bottom of the Figure and it is the same as that used in [9].

for the best fit and, though a variability was observed, a consensus was obtained after statistical evaluation of the base frequency at each position. The high binding affinity of sequences that only partially conform to the consensus confirms that Pax proteins possess a versatile DNA sequence recognition [17]. In order to demonstrate that the consensus obtained is due to the protein binding, sequences 11 and 16 were used to test whether the area bound by the protein (assessed by the phenanthroline footprint and methylation interference) corresponds to the area that fits the consensus. Figure 1(b) shows a schematic representation of the results, confirming that the sequence corresponding to the consensus is indeed the area where the protein binds. Protein–DNA contacts mapped by methylation interference indicate the presence of interactions both inside and outside the consensus area. Moreover, the distribution of contacts observed in sequence 11 is significantly different from that observed in sequence 16, with only three positions contacted in both sequences. Interestingly, all of these three positions are located inside the consensus area. These results are similar to those obtained with natural Pax-8-binding sequences [12], stressing the versatile DNA sequence recognition of Pax proteins.

In Figure 2 the consensus sequence for Pax-8 is aligned with that of the *Drosophila* Prd protein and with those found for other Pax proteins. The aligned sequences are superimposed in the manner of DNA docking of the Prd paired domain, recently determined by crystallography [9]. As shown in the Figure, base-pairs at the 3′ half of consensus sequences (from position 9 to position 14) are conserved and, contrarily, base-pairs at the 5′ half appear to be quite heterogeneous. According to the crystal of the Prd–DNA complex [9], base-pairs of the 3′ half are bound only in the minor groove and base-pairs of the 5′ half are bound only in the major groove (these latter mainly by the amino acids of the recognition helix: α3 in Figure 2). Thus, among Pax proteins, the minor groove contacts appear much more conserved than the major groove contacts. Accordingly, amino acids that in the Prd paired domain–DNA complex establish the minor groove contacts are invariant among all the paired domain-containing proteins [9]. The differential binding specificity of Pax proteins [17,18] would be mostly due to the major groove contacts established by recognition helices of the N- and C- terminal H–T–H motifs. The high homology of bases contacted in the minor groove contradicts the generally accepted idea that interactions established in the minor groove hardly discriminate among different base-pairs [19]. A possible explanation of this phenomenon could be that Pax proteins are able to bend DNA [20], therefore introducing spacing constraints that allow a preference for particular base-pairs in the network of the minor-groove contacts. It should be noted that the high binding affinity to Pax-8 of some sequences in which only a fraction of bases contacted in the minor groove are conserved (see, for instance, sequences 2, 5 and 16 in Figure 1) indicates a flexibility of the interaction, maybe due to 'induced fit' mechanisms [21]. Induced fit mechanisms may well occur during the paired-domain–DNA interaction. In fact, it has been demonstrated that secondary-structure changes upon DNA binding occur during the Pax-6–DNA interaction [22]. Nevertheless, the functional relevance of the network of base-specific minor-groove contacts is suggested from the existence of mutants of Pax protein with changes in amino acids contacting the minor groove, in which the biological function is severely impaired [9,23].

In order to provide a direct evidence of the importance of base-specific minor groove contacts, the effect of consensus-based mutations on the Pax-8 binding sequence of Tg promoter (C site) was analysed. The C site was aligned to the Pax-8 consensus and the best fit was observed when the sequences were matched as previously done by Epstein et al. [22] by using the Pax-2 consensus. The sequences of the C site mutants that we have analysed are shown in Figure 3(a). Pax-8 and Prd paired domain binding to these sequences is shown in Figure 3(b). When the C sequence is mutated to obtain a perfect match for base-pairs contacted in the minor groove (Cβ mutant), an increase in the strength of interaction with Pax-8, with respect to wild-type sequence, is observed. Mutants Cβ_{12m} and Cβ_{14m} contain an inosine–5′-methylcytosine base-pair in place of the A–T base-pair at position 12 and 14 respectively (Figure 3a). This mutation affects the distribution of substituents in the major groove but not in the minor groove (Figure 3c). Pax-8 recognizes mutants Cβ_{12m} and Cβ_{14m} with the same efficiency observed for Cβ (Figure 3b), indicating that the distribution of substituents in the major groove does not play a role in the Pax-8–Cβ interaction. Thus these data support the view that, at base-pairs 12 and 14, Pax-8 interacts through minor-groove contacts. In the C_{ant2} mutant, the change of C–G base-pair to A–T replaces in the minor groove the amino group of G with the carbonyl oxygen of T. In the Prd–DNA crystal, Gly_{15} (conserved in Pax-8) interacts using its carbonyl oxygen with the amino group of G. Therefore the lack of Pax-8 binding to the C_{ant2} sequence could be easily explained by the lack of the hydrogen bond because of the

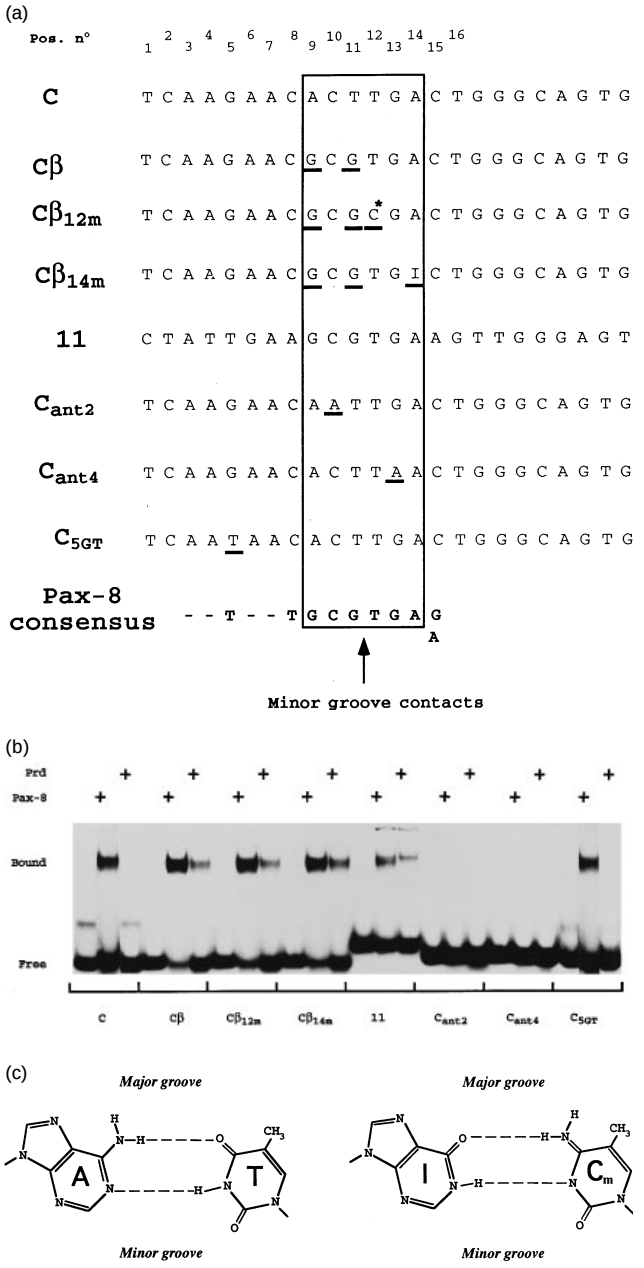

Figure 3 Binding activity of Pax-8 and Prd proteins on different sequences

(**a**) Sequences used in binding study. Only the top strand is indicated. Mutants of the C sequence are described in the text. Mutations introduced are underlined. The asterisk over Cβ_{12m} indicates a methylated cytosine. I in the sequence Cβ_{14m} indicates inosine. Sequences are aligned to the Pax-8 consensus, which is shown at the bottom. A box delimits the region contacted in the minor groove. (**b**) Gel-retardation assay demonstrative of the strength of the interactions of sequences shown in (**a**) with Pax-8 and Prd proteins. (**c**) Schematic representation of major-groove modifications introduced by substituting A with I and T with 5-methylcytosine (C_m) in sequences Cβ_{12m} and Cβ_{14m}.

presence of two acceptor sites (carbonyl groups of Gly_{15} and T). Thus this finding supports the relevance of contacts established in the minor groove. Interestingly the perturbation of Pax-1–DNA interaction at this position explains the *undulated* mutation in the mouse [6]. In C_{ant4} the G:C base-pair at position 13 has been substituted by an A–T base-pair. In this manner the G↔Atransition, in the minor groove removes an amino group on one strand. The abolition of Pax-8 binding induced by this mutation reveals the importance of this amino group for a proper contact. This finding complements the crystallographic data, in which precise interactions were not clear in this region. On the basis of the comparison of the Pax protein consensus sequence (Figure 2), the only base-pair contacted in the major groove (and conserved) is the T–A at position 5. In the Prd–DNA crystal the methyl group of thymine is contacted through van der Waals interactions by Gly^{48} and Ser^{51}. In the C site, at position 5, a G in place of T is present, supporting the notion that the hydrophobic interaction described in the Prd–DNA crystal does not occur in the Pax-8–C interaction. In mutant C_{5GT} a thymine is present at position 5 (Figure 3a). However, this mutant is recognized by Pax-8 exactly in the same way as the wild-type C sequence (Figure 3b). This finding indicates that, in the Pax-8–C interaction the hydrophobic contacts at the level of base-pair 5 do not have a relevance in terms of binding strength, supporting the view of versatile DNA recognition by Pax proteins [17]. All together these data demonstrate that specific minor-groove contacts appear to play a fundamental role in the Pax-8–C interaction, indicating a role of this conserved network in the interactions between Pax proteins and natural binding sites. This observation is in good agreement with the findings of Czerny et al. [17]. In fact, in that study sequences referred to as 'class II binding sites' show high homology to the consensus of bases contacted in the minor groove and are efficiently recognized by a wide spectrum of Pax proteins. In order to support this view, the binding activity of the Prd paired domain with sequences shown in Figure 3a was evaluated. Albeit with a much lower affinity (about 10-fold less; results not shown), the Prd paired domain efficiently recognized only sequences in which the consensus for the network of minor-groove contacts was fully preserved (Cβ, Cβ_{12m}, Cβ_{14m} and 11) (Figure 3b).

In addition to the paired domain, the homeodomain is another eukaryotic DNA-binding domain able to establish specific contacts in the minor groove [24]. Homeodomains contact DNA in the minor groove through the N-terminal arm [24] but, in contrast with what is observed for paired domains, a variability of the minor groove contacting amino acid is present among different members of this class of proteins [25]. This variability plays a major role in determining the differential DNA-binding specificity observed among homeodomains [26–27]. Therefore distinct classes of DNA-binding proteins use specific minor-groove contacts according to different strategies: homeodomains use minor-groove contacts to obtain a differential DNA-binding specificity, while, in paired-domain proteins, minor-groove contacts are conserved and represent a common characteristic for all members of the class.

Plasmids pGEX Pax-8 Pb and pGEX Prd Pb were provided by C. Desplan. This work was funded by the Associazione Italiana per la Ricerca sul Cancro (A.I.R.C.). L. P. is supported by the Ph.D. program 'Diagnosi e terapie cellulari e molecolari' of Udine University. D. F. is supported by an A.I.R.C. Fellowship. We thank Elio Biffali for the synthesis of high-quality oligonucleotides and S. Guerra for comments about manuscript before its submission.

REFERENCES

1 Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. and Noll, M. (1986) Cell **47**, 1033–1040
2 Burri, D., Tromvoukis, Y., Bopp, D., Frigerio, G. and Noll, M. (1989) EMBO J. **8**, 1183–1190
3 Noll M. (1993) Curr. Opin. Genet. Dev. **3**, 595–605
4 Stuart, E. T. and Gruss P. (1995) Hum. Mol. Genet. 4, 1717–1720

5 Maulbecker C. C. and Gruss P. (1993) EMBO J. **12**, 2361–2367
6 Chalepakis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M. and Gruss, P. (1991) Cell **66**, 873–884
7 Dressler, G. R. and Douglass, E. C. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 1179–1183
8 Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R. and Gruss, P. (1991) EMBO J. **10**, 1137–1147
9 Xu, W., Rould, M. A., Jun, S., Desplan, C. and Pabo, C. O. (1995) Cell **80**, 636–650
10 Kozmik, Z., Kurzbauer R., Dorfler P. and Busslinger, M. (1993) Mol. Cell. Biol. **13**, 6024–6035
11 Plachov, D., Chowdhury, K., Walther, C., Simon D., Guenet, J. L. and Gruss P. (1990) Development **110**, 643–651
12 Zannini, M., Francis-Lang, H., Plachov, D. and Di Lauro, R. (1992) Mol. Cell. Biol. **12**, 4230–4241
13 Smith, D. B. and Johnson, K. S. (1988) Gene **67**, 31–40
14 Kuwabara, M. D. and Sigman, D. S. (1987) Biochemistry **26**, 7234–7238
15 Affolter, M., Percival-Smith, A., Muller, M., Leupin, W. and Gehring, W. J. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 4093–4097
16 Pierrou, S., Enerback, S. and Carlsson, P. (1995) Anal. Biochem. **229**, 99–105
17 Czerny, T., Schaffner, G. and Busslinger, M. (1993) Genes Dev. **7**, 2048–2061
18 Czerny, T. and Busslinger, M. (1995) Mol. Cell. Biol. **15**, 2858–2871
19 Seeman, N. C., Rosenberg, J. M. and Rich, A. (1976) Proc. Natl. Acad. Sci. U.S.A. **73**, 804–808
20 Chalepakis, G., Wijnhlds, J. and Gruss, P. (1994) Nucleic Acids Res. **22**, 3131–3137
21 Spolar, R. S. and Record, Jr., M. T. (1994) Science **263**, 777–784
22 Epstein, J., Jiexing, C., Glaser, T., Jepeal, L. and Maas, R. (1994) J. Biol. Chem. **269**, 8355–8361
23 Balling R., Deutsch U. and Gruss P. (1988) Cell **55**, 531–535
24 Gehring, W. J., Qian, Y. Q., Billeter, M., Furukubo-Tokunaga, K., Schier, A. F., Resendez-Perez, D., Affolter, M., Otting, G. and Wuthrich, K. (1994) Cell **78**, 211–223
25 Gehring, W. J., Affolter, M. and Burglin, T. (1994) Annu. Rev. Biochem. **63**, 487–526
26 Ekker, S. C., Jackson, D. G., von Kessler, D. P. Sun, B. I., Young, K. E. and Beachy, P. A. (1994) EMBO J. **13**, 3551–3560
27 Ades, S. E. and Sauer, R. T. (1995) Biochemistry **34**, 14601–14608
28 Chalepakis, G. and Gruss P. (1995) Gene **162**, 267–270

Received 15 January 1996/20 February 1996; accepted 22 February 1996