Volume 14 Number 16 1986 Nucleic Acids Research

Oligonucleotide duplexes containing inosine, 7-deazainosine, tubercidin, nebularine and 7-deazanebularine as substrates for restriction endonucleases *Hin*dII, *Sal*I and *Taq*I

Josef Jiricny, Steven G.Wood, Doris Martin and Aiko Ubasawa

Friedrich Miescher Institut, PO Box 2543, CH-4002 Basel, Switzerland

Received 11 June 1986; Accepted 14 July 1986

ABSTRACT. Synthetic hexadecanucleotide duplexes containing a single purine nucleotide analogue in the recognition sites of the restriction endonucleases HindII, SalI and TaqI were used to investigate the restriction site determinants required by these enzymes for sequence recognition and phosphodiester bond cleavage. The enzymes were, in general, unaffected by changes introduced into the minor groove of the helix. SalI was found to be inhibited by the major groove modifications introduced into the fourth position of its recognition sequence GTCGAC. HindII and TaqI were, by contrast, able to cleave the sites containing the analogues at this position. TaqI and, to a lesser extent, HindII could also be shown to tolerate "mismatch analogues" at this site.

INTRODUCTION.

Restriction endonucleases type II, over 600 of which have been isolated to date (1) are capable of cleaving DNA at precisely-defined recognition sequences. The multitude of restriction sites makes these enzymes indispensible tools of DNA analysis and manipulation. They are also highly versatile models for studying protein/DNA interactions.

The restriction enzymes HindII, SalI and TaqI recognize and cleave the palindromic duplex sequences GTYRAC (R=purine, Y=pyrimidine), GTCGAC and TCGA, respectively. All three should thus be able to cleave the 16-mer duplex shown in Fig. 1 (R/Y=G/C) at the indicated sites. HindII should, in addition, cleave the duplex (R/Y=A/T).

As we recently reported, HindII and TaqI, unlike SalI, are also able to cleave DNA containing a single purine/pyrimidine mismatch within their respective restriction sites (2). Thus, at 37°C, HindII was shown to digest the duplexes R/Y = G/C, G/T, A/T, A/C, and TaqI cleaved the sequences R/Y = G/C, G/T and A/C.

These unexpected findings indicate that despite the similarity of their restriction sites, all three enzymes require a different set of structural determinants for the recognition of these sites.

Bodnar et al. (3) investigated the effect of twelve nucleoside analogues on the rate of DNA cleavage by AluI, DdeI, HinfI, RsaI and TaqI. This study indicated the importance of major groove determinants, such as the nitrogen atom at the 7-position of purine residues, for recognition by these enzymes.

Using synthetic decanucleotide duplexes, Ono et al. (4) have substantiated these findings by demonstrating the inhibition of enzyme action by the presence of tubercidin (7-deazaadenosine) in the recognition sites of BglII and Sau3AI. EcoRI, probably the most widely-studied restriction endonuclease, was shown to be sensitive to the presence of 7-deazaguanosine (DG) in the octamer duplex G DG A A T T C C (5). The same enzyme was, by contrast, found to tolerate thymine analogues within its restriction sequence (6).

We now wish to report the effect on the rate of cleavage of a synthetic 16-mer duplex by HindII, SalI and TaqI, when the central guanine moiety of the palindrome GTCGAC is substituted with guanosine analogues inosine (I) and 7-deazainosine (DI), and with adenosine analogues tubercidin (Tu), nebularine (N) and 7-deazanebularine (DN) (Fig.1). The effect of "mismatch analogues" at this site on the rate of DNA cleavage by these enzymes is also described.

MATERIALS AND METHODS.

All restriction enzymes described in this study were purchased from Boehringer Mannheim and used according to the manufacturers instructions.

Synthesis of 2'-deoxynucleosides: 2'-deoxytubercidin (Tu) was prepared from tubercidin (Sigma) by the method of Robins (7). 2'-Deoxy-7-deazainosine (DI) was prepared from 2'-deoxytubercidin by the action of nitrous acid (8). 2'-Deoxy-7-deazanebularine (DN) was prepared by chlorination of 2'-deoxy-7-deazainosine, followed by catalytic reduction of the resulting 4-chloro-pyrrolo[2,3-d]-pyrimidine-2'-deoxyriboside using published methods (9). 2'-

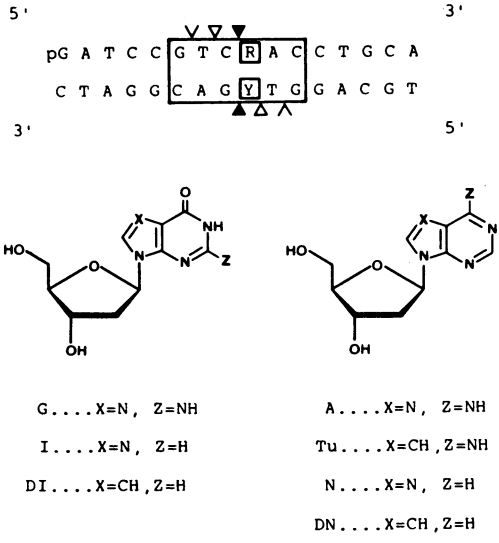

Figure 1: The structures of the 16-mer oligodeoxyribonucleotide duplex, and of the nucleoside analogues used in this study. **p** denotes a [^{32}P]-phosphate group, **R** = purine or purine analogue, **Y** = C or T, the symbols ▼, ⋁ and ▽ indicate the cleavage sites of HindII, SalI and TaqI resp.; the 2'-deoxynucleosides are: G = guanosine, I = inosine, DI = 7-deazanebularine, A = adenosine, Tu = tubercidin, N = nebularine, DN = 7-deazanebularine. [Note: For the sake of clarity, pyrrolo[2,3-d]pyrimidines (DI, Tu, DN) are referred to as 7-deazapurines throughout this report; the purine numbering system is used for both types of compounds.]

Deoxynebularine (N) was prepared by 2'-deoxygenation (7) of 6-chloropurine riboside (Sigma), followed by catalytic reduction of the resulting 6-chloropurine-2'-deoxyriboside. 2'-Deoxyinosine was purchased from Sigma. All the above nucleosides were converted to the suitably protected phosphoramidite building blocks for oligonucleotide synthesis by published methods (10).

Oligonucleotide synthesis: The oligonucleotides were synthesized by the phosphoramidite method using an Applied Biosystems Model 308A synthesizer, and purified by polyacrylamide gel electrophoresis (11). The position and integrity of the analogues within the hexadecanucleotides was established by sequencing, using the chemical modification procedure of Maxam & Gilbert as described by Zoller & Smith (12), and by complete hydrolysis with

snake venom phosphodiesterase and alkaline phosphatase, followed by HPLC analysis of the resulting nucleoside mixture (13, and Wood et al., in preparation).

Annealing of the oligonucleotides: The 5'-labelled oligonucleotide (20 pmole) was hybridized to its complementary unlabelled hexadecanucleotide (25 pmole) in annealing buffer (10mM Tris pH 8.3, 10mM $MgCl_2$) in a total volume of 10ul (80°C, 5 min., then 20°C, 20 min.).

Duplex melting experiments: The thermal dissociation of the 16-mer duplexes was measured as a function of the UV absorbance at 260nm versus temperature, using a Gilford 2600 spectrophotometer fitted with a model 2527 thermo-programmer and a microcuvette Peltier heater assembly. The results are shown in figure 3.

Restriction enzyme digests: The annealed oligonucleotide mixture (2ul, 4 pmole), BSA (1ul, 50mg/ml), 10x restriction enzyme assay buffer (2ul), water (14ul) and one unit of the enzyme were incubated at 37, resp. 20°C for up to 4 hours. 2ul aliquots were withdrawn at regular time intervals, mixed with 2ul of formamide loading dye and heated for 1min at 90°C. 2ul of this mixture were immediately loaded onto a denaturing 20% polyacrylamide gel (20x20cm). Electrophoresis was carried out at 40V/cm, until the bromophenol blue marker dye migrated approximately 10cm. Following autoradiography, the radioactive bands were cut out and counted by Cerenkov method for 0.5min. The autoradiograms were additionally scanned with a Shimadzu C-930 densitometer. The relative yields of product were determined from the relationship

[%Product] = cpm(product)/[cpm(product) + cpm(substrate)].

The results of digests, in which the analogue-containing oligonucleotide was labelled, are shown in Table 1. Experiments involving the digests of duplexes with labelled complementary strands were also carried out, in order to establish that both strands of the duplex were cleaved. HindII and SalI digested the two strands with different rates, while TaqI showed no strand preference (results not shown). This phenomenon was discussed earlier (2) and is therefore not mentioned below.

RESULTS AND DISCUSSION

Recent data from our laboratory (2) showed that SalI was unable to cleave the hexadecanucleotide duplexes where R/Y = A/C,

Table 1: SalI, HindII and TaqI digests of the 16-mer oligonucleotide duplexes. The figures quoted in this table represent the average values obtained from three separate experiments. The observed deviations between two successive experiments were typically within +/- 8% of the quoted values.

	Yield of product (%)								
	SalI			HindII			TaqI		
R/Y	10'	30'	180'	10'	30'	180'	10'	30'	180'
G/C	36	64	84	26	48	94	14	31	97
I/C	37	63	81	26	66	90	16	26	87
DI/C	1	2	6	38	85	91	11	29	91
A/C	0	0	0	26	53	88	8	19	86
Tu/C	0	0	0	2	4	13	10	23	77
N/C	0	0	0	23	61	86	6	14	70
DN/C	0	0	0	2	6	29	8	19	66
G/T	0	0	0	4	13	61	10	21	77
I/T	0	0	0	7	20	56	5	7	21
DI/T	0	0	0	3	6	19	2	3	6
A/T	0	0	0	21	41	88	0	0	0
Tu/T	-	-	-	24	57	73	-	-	-
N/T	-	-	-	10	16	65	-	-	-
DN/T	-	-	-	10	26	81	-	-	-

A/T or G/T. These findings are substantiated by the results of our current experiments. Of the three enzymes studied, SalI was shown to be the most sensitive to changes introduced into its recognition sequence GTCGAC. However, the substitution of the central guanosine of this site with inosine (I) had no demonstrable effect on the rate of DNA cleavage (Table 1, R/Y=I/C). This suggests that the exo-amino group of this guanine, a minor groove determinant, is not important for the recognition/cleavage process. By contrast, the introduction of 7-deazainosine (DI) at this site resulted in a drastic reduction in the rate of digest (Table 1, R/Y=DI/C). In this instance, the N-7 nitrogen atom of the purine moiety was substituted with a methine (CH) group, a major groove modification. The question whether this rate reduction is brought about by the lack of enzyme binding at this site, or whether the bound enzyme merely fails to cleave the DNA could be answered by the following inhibition experiment: To

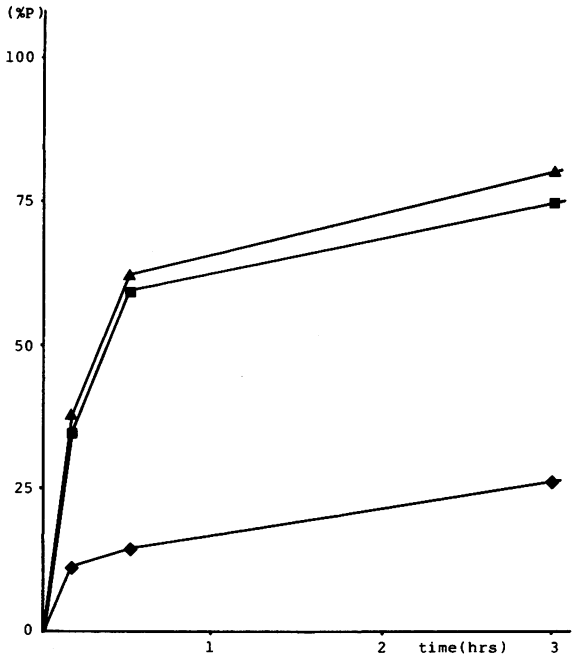

Figure 2: The inhibition of SalI I/C digest (▲) with EcoRI linker CGGAATTCCG (■) and with D/I duplex (◆).

the radioactively-labelled I/C digest mixture was added a 20-fold excess of unlabelled DI/C duplex. The rate of digest of the I/C sequence was significantly reduced as compared to a control experiment, in which a decamer EcoRI linker was added in place of DI/C (Fig. 2). These results suggest that SalI binds to the DI/C duplex, but fails to cleave it. We may conclude that the potential hydrogen-bonding site at the N-7 position of the purine moiety plays an important role in the recognition and cleavage of this site by SalI, the latter being more strongly affected by its absence.

HindII was shown to have somewhat less stringent sequence recognition requirements within the central two base-pairs of its hexamer restriction site GTYRAC, as demonstrated by its ability to cleave mismatch-containing DNA duplexes (2). This enzyme is also relatively insensitive to major and minor groove modifications within the fourth base-pair (Table 1). The duplexes I/C,

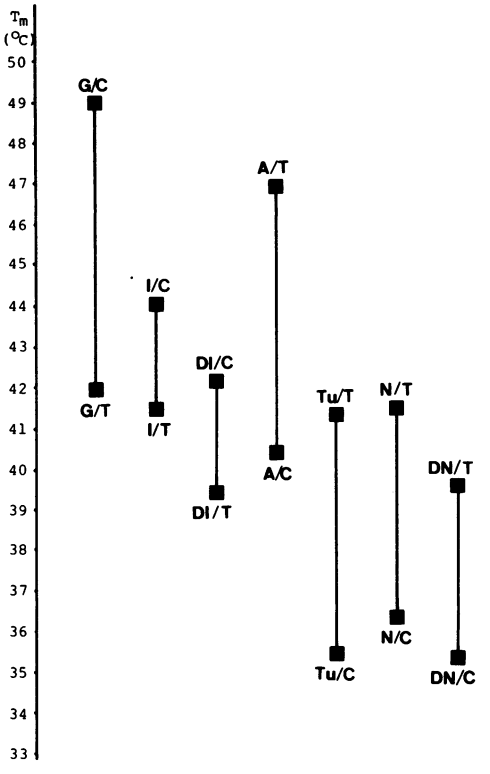

Figure 3: Duplex dissociation temperatures T_m estimated from an $Absorbance_{260}$ vs temperature plot in 10mM Tris pH 8.3, 15mM NaCl.

DI/C, Tu/T, N/T and DN/T, which contained the analogues base-paired with their respective complementary pyrimidines, were cleaved with kinetics similar to the two bona fide duplexes, G/C and A/T. However, significant rate differences became apparent during the digestions of the "mismatch analogues" I/T, DI/T, Tu/C, N/C and DN/C. The mismatched duplexes containing 7-deazapurines DI, DN and Tu were cleaved more slowly than those containing the purine/pyrimidine mismatches I/T and N/C (Table 1). These results raise an interesting question: why should the same major groove change influence the recognition/cleavage process of the mismatched duplexes, without affecting the rates of digest of the matched ones? Melting curves of the duplexes (Fig. 3) show that the destabilization of the double-stranded

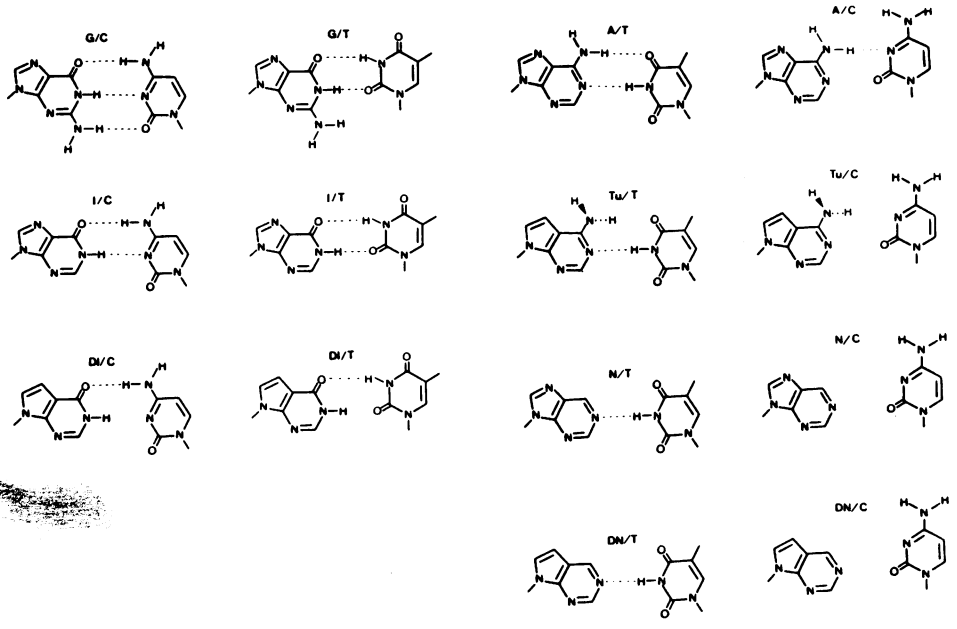

Figure 4: Proposed structures of the matched and mismatched base-pairs used in the study.

structure, brought about by the presence of a mismatch, lowers the T_m by 2-7°C. At 37°C, the assay temperature, all the duplexes should still possess a considerable degree of double-stranded character, as demonstrated by the comparable rates of digest of the duplexes N/C and N/T for example. Furthermore, when the digests were carried out at room temperature, the rates of digest of all the duplexes were reduced to a similar extent.

The answer to our question lies most likely in the structures of the different matched and mismatched base-pairs (Fig.4). The matched base-pairs are held together by three (G/C), two (A/T, Tu/T, I/C, DI/C) or one (N/T,DN/T) hydrogen bond, the bond between the N-1 position of the purine system and the N-3 of the pyrimidine being present in all cases. The G/T mismatch (14,15) is stabilized by two H-bonds, the N-1 and O-6 of guanosine being bound to the O-2 and N-3 positions of the thymidine,respectively. This has the effect of pushing the O-4 group of T further into the major groove and, at the same time, rotating the purine

moiety into the minor groove. The structures of the G/T and I/T mismatches would be expected to be similar. The DI/T mismatch could potentially also form two H-bonds, however, the T_m of the duplex containing this mismatch is approximately 2°C lower than those of the G/T and I/T duplexes. This suggests that the base-pairing in this mismatch is substantially weakened, presumably as a result of the different electron delocalization in the deazapurine ring system. Further evidence for this theory comes from the melting experiments (Fig.3). The DI/T duplex dissociated at a temperature characteristic for duplexes in which the central base-pair is stabilized by only one H-bond, such as the matched duplex DN/T. A single hydrogen bond may not be sufficiently strong to stabilize a mismatched base-pair; a duplex carrying such a mismatch might thus be "melted" at this position. Furthermore, as the analogue lacks a nitrogen atom at the 7-position, a possible stabilization through H-bonding with solvent molecules or, conceivably, with the restriction enzyme itself, would not be possible. The base-pair "melting" could thus be sufficient to adversely affect the recognition/cleavage process of HindII.

A similar argument could apply to the mismatch analogues DN/C and Tu/C. The former cannot form any H-bonds with C at neutral pH. The latter, in a fashion analogous to the A/C mismatch (16), could form a bond between the exo-amino group of tubercidin and the N-3 of cytidine. In this instance, however, the steric effect of the proton at the 7-position of the purine analogue could force the exo-amino group out of the plane of the aromatic ring system, reducing thus the electron delocalization and weakening the H-bond at this site. As the mismatch analogues lie immediately adjacent to the site of cleavage by HindII, it is perhaps not surprizing that the rate of hydrolysis of the phosphodiester moiety in these duplexes is affected.

Our hypothesis (2) that TaqI requires only one strand of a duplex for site-recognition could also be substantiated by the current results. All the duplexes R/C were cleaved with approximately equal efficiencies (Table 1). The duplexes G/T and A/C, in which at least one strand carries the sequence TCGA, were

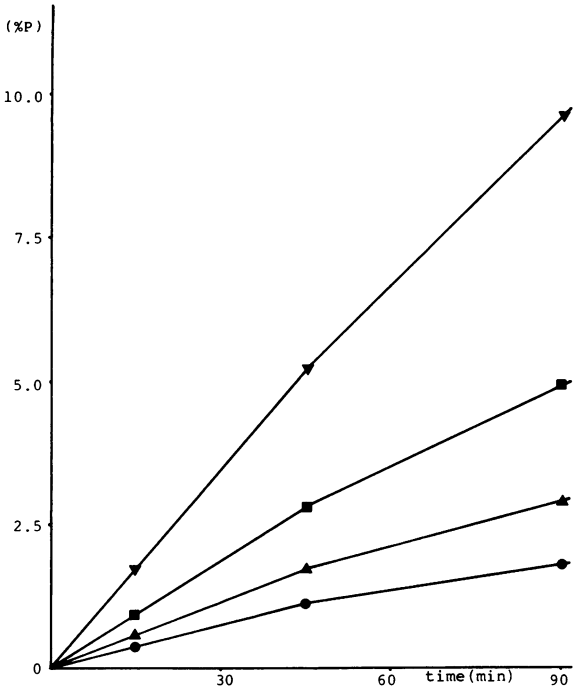

Figure 5: Room temperature TaqI digests of A/C (●), Tu/C (▲), N/C(■) and DN/C(▼).

also cleaved with similar rates. By contrast, the duplex A/T was not cleaved, as neither strand contains a TaqI restriction site. The I/T duplex, which lacks an exo-amino group in the minor groove of the restriction site was cleaved, but the rate of digest was substantially lowered. The DI/T duplex, which contains an additional major groove modification at the same position, was digested extremely slowly (Table 1). At 37°C, the rates of digest of all the duplexes R/C were closely related. At room temperature the mismatched duplexes were cleaved in the order DN/C N/C Tu/C A/C (Fig.5), which reflects the following base-pair structures: no H-bond + major groove destabilization, no H-bond, one H-bond + major groove destabilization, one H-bond, respectively. These results further support the theory that TaqI requires melted DNA for recognition of its restriction sites.

CONCLUSIONS.

The evidence emerging from recent studies with restriction enzymes (2-6), as well as some other well-characterized proteins such as the CRO repressor (for review see 17), is beginning to reveal a number of common structural features, which are required in the interactions of proteins with DNA. Major groove determinants seem to be implicated in a vast majority of these events, the N-7 position of purines being of particular importance. To date, however, the interactions of a specific DNA sequence were carried out with one type of protein only. The sequences of the oligodeoxyribonucleotides used in this study comprise the BamHI-PstI fragment of the filamentous bacteriophage M13mp9, and can thus be used, following their incorporation into the M13 replicative form, to investigate their interactions with *E. coli* proteins *in vivo*. Our experiments, concerned with the repair of the above-studied purine analogues from mismatches, are described in the accompanying report.

ACKNOWLEDGEMENTS.

The authors wish to thank Drs. K. Wiebauer, P. Kuenzler and C. Moroni for critical reading of the manuscript.

REFERENCES.

1. Kessler,C.; Neumaier,P.S.; Wolf,W.; (1985) Gene 33, 1-102.
2. Jiricny,J., Martin,D., (1986) Nucl. Acids Res., 14(5), 1943-1949.
3. Bodnar,J.W.; Zempsky,W.; Warder,D.; Bergson,C.; Ward,D.C.; (1983) J.Biol.Chem.,258, 15206-15213.
4. Ono,AY.; Ueda,T.; (1984) Nucl.Acids Res., 12, 8939-8949.
5. Seela,F., Driller,H., (1986) Nucl. Acids Res., 14(5), 2319-2332.
6. Fliess,A., Wolfes,H., Rosenthal,A., Schwellnus,K., Bloecker,H., Frank,R., Pingoud,A., (1986) Nucl. Acids Res., 14(8), 3463-3474.
7. Robins,M.J., Wilson,J.S., (1981) J.Am.Chem.Soc., 103, 932-933.
8. Pike,J.E., Slechta,L., Wiley,P.F., (1964) J.Heterocyclic Chem., 1, 159-164.
9. Gerster,J.F., Carpenter,B., Robins,R.K., Townsend,L.B., (1967) J.Med.Chem., 10,326-331.
10. Barone,A.D., Tang,J.-Y., Caruthers,M.H., (1984) Nucl.Acids Res., 12(10), 4051-4061.
11. Atkinson,T; Smith,M.;(1984) Oligonucleotide synthesis (M.J. Gait ed.), practical approach series, IRL press Oxford.Washington DC, 35-81.

12. Zoller,M.J., Smith,M., (1983) Methods in Enzymology, 100, 468-500.
13. Frank,R., Bloecker,H., (1982) Chemical and Enzymatic Synthesis of Gene Fragments, (Eds. Gassen,H.G., Lang,A.,) 225-246.
14. Patel,D.J.; Kozlowski,S.A.; Marky,L.A.; Rice,J.A.; Broka,C.; Dallas,J.; Itakura,K.; Breslauer, K.J.; (1982) Biochemistry, 21, 437-444.
15. Brown,T.; Kennard,O.; Kneale,G.; Rabinovich,D.; (1985) Nature, 315, 604-606.
16. Hunter,W.N., Brown,T., Anand,N.N., Kennard,O., (1986) Nature 552-555.
17. Pabo,C.O., Sauer,R.T., (1984) Ann.Rev.Biochem., 53,293-321.