New Insights on DNA Recognition by ets Proteins from the Crystal Structure of the PU.1 ETS Domain-DNA Complex*

(Received for publication, May 24, 1996, and in revised form, July 6, 1996)

Frédéric Pio‡‡, Ramadurgam Kodandapani‡‡, Chao-Zhou Ni‡, William Shepard‡‡, Michael Klemsz, Scott R. McKercher‡, Richard A. Maki‡‡‡, and Kathryn R. Ely†‡‡

From the ‡‡La Jolla Cancer Research Center, The Burnham Institute, La Jolla, California 92037, §LURE, Université Paris-Sud, Bât. 209, 91405 Orsay Cedex, France, and ‡‡Neurocrine Biosciences, San Diego, California 92121

Transcription factors belonging to the ets family regulate gene expression and share a conserved ets DNA-binding domain that binds to the core sequence 5′-(C/A)GGA(A/T)-3′. The domain is similar to α+β (“winged”) helix-turn-helix DNA-binding proteins. The crystal structure of the PU.1 ETS domain complexed to a 16-base pair oligonucleotide reveals a pattern for DNA recognition from a novel loop-helix-loop architecture (Kodandapani, R., Pio, F., Ni, C.-Z., Piccialli, G., Klemsz, M., McKercher, S., Maki, R. A., and Ely, K. R. (1996) Nature 380, 456–460). Correlation of this model with mutational analyses and chemical shift data on other ets proteins confirms this complex as a paradigm for ets DNA recognition. The second helix in the helix-turn-helix motif lies deep in the major groove with specific contacts with bases in both strands in the core sequence made by conserved residues in α3. On either side of this helix, two loops contact the phosphate backbone. The DNA is bent (8°) but uniformly curved without distinct kinks. ETS domains bind DNA as a monomer yet make extensive DNA contacts over 30 Å. DNA bending likely results from phosphate neutralization of the phosphate backbone in the minor groove by both loops in the loop-helix-loop motif. Contacts from these loops stabilize DNA bending and may mediate specific base interactions by inducing a bend toward the protein.

Transcription factors bind to target DNA sequences to regulate metabolic functions such as growth and differentiation. Typically, the molecular scaffold for DNA recognition is conserved within a given family of DNA-binding proteins. In some cases the similarity of these scaffolds suggests an evolutionary relationship between different families or comparison of scaffolds reveals a structural similarity that was obscured by sequence comparisons alone.

A recently discovered family of regulatory proteins, the ets gene family, includes more than 45 members in a variety of organisms from Drosophila to humans (1, 2). These molecules play a role in normal development and have been implicated in malignant processes such as erythroid leukemia and Ewing’s sarcoma. The DNA-binding domain of ets proteins is a conserved region (ETS domain) that is about 85 residues in length. Although ets proteins share a homologous sequence in the ETS domain, they differ in length and in the relative position of this domain. In some molecules, the ETS domain is found at the carboxyl terminus (e.g. PU.1 (3); ets-1 (4); ets-2 (5)), while in others the domain is located in the middle of the sequence (erg (6)), or in the amino-terminal region (elk-1 (7)). Flanking regions are thought to form other functional domains that influence protein-protein recognition or inhibitory domains that mask the DNA-binding site (8, 9). In ets-1, an α-helix that is located in an inhibitory domain immediately NH2-terminal to the ETS domain unfolds on DNA-binding (10). Regardless of the position of the ETS domain within the intact ets proteins, there is strong sequence homology in this conserved region.

We have determined the crystal structure of the ETS domain of the PU.1 transcription factor complexed to DNA (11). The domain is similar to α + β helix-turn-helix (HTH) DNA-binding proteins and contacts a 10-base pair region of duplex DNA that is bent (8°) but uniformly curved without distinct kinks. The PU.1 domain assumes a tight globular structure with three α-helices and a four-stranded antiparallel β-sheet enclosing a hydrophobic core. The topology of the domain is similar to the structures of other ets family proteins fli-1 (12), murine ets-1 (13), and human ets-1 (14) determined in solution by NMR. The common molecular scaffold is similar to DNA-binding proteins such as CAP (15) and resembles “winged”-HTH proteins including HNF-3γ (16). ETS domains bind as a monomer to the core sequence 5′-(C/A)GGA(A/T)-3′. The PU.1 domain contacts DNA from three sites: the recognition helix (α3) interacts with the GGAA core sequence in the major groove, while contacts with the phosphate backbone on either side of this site are made in the minor groove by two loops. Therefore, the PU.1 ETS domain binds DNA by a loop-helix-loop motif. One loop is formed between β-strands 3 and 4 (a “wing”) and the other is a loop in the position of the turn in the HTH motif (α2-turn-α3). The protein-DNA contacts stabilize a uniform bending of the duplex DNA that is due to phosphate neutralization by the PU.1 domain. Surprisingly, the protein-DNA interactions reported in the NMR structure of a human ets-1-DNA complex (14) differed dramatically from this pattern, involving different contacts and significant DNA deformation. Because of this discrepancy, we chose to test the validity of the PU.1-DNA complex as a model for other ets proteins. As reported here, when the results of mutational analyses on a number of ets proteins are correlated with the structure of the PU.1-DNA complex and with chemical shift data measured with the fli-1 (12) and murine ets-1 (13) mole-

*This work was supported in part by USAMRDC Department of the Army Grants DMD17–944–4439 (to K. R. E.), National Institute of Allergy and Infectious Disease Grant AI20194 (to R. A. M.), and National Cancer Institute Grant CA63489 (to K. R. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Contributed equally to the results presented in this study.

§ Present address: Dept. of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, IN 46202.

‡‡ To whom correspondence should be addressed: The Burnham Institute, 10901 N. Torrey Pines Rd., La Jolla, CA 92037. Tel.: 619-646-3135; Fax: 619-646-3196.

1 M. Klemsz and R. A. Maki, unpublished results.
2 The abbreviation used is: HTH, helix-turn-helix.
cules, the loop-helix-loop scaffold is confirmed as a general model for DNA recognition by ets proteins. This pattern defines a new class of HTH DNA-binding proteins. The molecular pattern of DNA recognition by ets proteins is compared to other HTH proteins for which crystal structures of the protein-DNA complexes are available.

**EXPERIMENTAL PROCEDURES**

**PU.1 DNA Complex**—A recombinant fragment encompassing residues 160–272 from the murine ets protein PU.1 was crystallized in complex with a 16-base pair oligonucleotide representing a consensus PU.1 DNA-binding site (3) as described previously (17). The complex crystallized in space group C2 with $a = 89.1$, $b = 101.9$, $c = 55.6$ Å, and $\beta = 111.2^\circ$. There are two complexes in the asymmetric unit. The length of the oligonucleotide was critical for crystallization and the oligonucleotide used to form the complex permitted end-to-end stacking of the DNA in the crystal lattice with the formation of pseudo-base pairing by the overhanging A and T bases.

Crystallographic Analyses—The initial structure analysis of the complex solved by the MIRAS method was reported (17). For this first phase of the study, a native data set and four heavy atom data sets were collected using a Rigaku RU200 rotating anode x-ray source and two San Diego Multiwire Systems area detectors. The initial data sets were collected from flash frozen crystals at 2.3-Å resolution. To refine the structure further, another native data set extending to 2.1 Å was collected at the LURE synchrotron source in Orsay, France. Diffraction data were collected at station D41 interfaced with the Mark III multiwire proportional area detector. Data sets were processed using MOSFLM (18) and ROTAVATA, AGROVATA, and TRUNCATE in the CCP4 package (19). In the present study, this native data set was scaled to the data collected in the home laboratory by Wilson scaling and the synchrotron data were incorporated into the refinement. The programs PHASES (20), FRODO (21), and X-PLOR (22) were used for structure solution, model building, and refinement. The current $R$-factor is 22.5 for 6 to 2.1 Å data (22,022 reflections). The average overall $B$-factor for 2929 non-hydrogen atoms (1486 protein atoms + 1300 DNA atoms + 143 solvent oxygens) is 31.6 Å$^2$. The refinement statistics are presented in Table I. There were 11 disordered residues at the amino terminus of the domain and 14 disordered residues at the carboxyl terminus of the recombinant fragment that were excluded from the model. These residues were not ordered even when the resolution was extended to 2.1 Å. For all residues representing the complete ETS domain (residues 171–258), the electron density was clear and permitted unambiguous fitting of both backbone and side chain atoms. More solvent atoms have been added to the model. Only minimal changes in the configuration of some side chains were evident in the high resolution map. The stereoechemistry of all main chain torsion angles in the domain fall within energetically favorable limits (Fig. 1) indicating that no segment of the domain is denatured or randomly configured. The DNA was clearly defined even in the first MIRAS map.

Analyses of DNA Helical Parameters—To analyze the stereochemical basis for the uniform bending observed in the oligonucleotide bound in complex to PU.1, the DNA superstructure was measured (23, 24) and four parameters were calculated that describe the conformation of the DNA bases and the phosphate backbone. The values were calculated (excluding the 5’ A overhang) to analyze helical parameters along the length of the oligonucleotide and to compare these with standard B-DNA parameters. The geometry of dinucleotide steps was analyzed for three rotational angles defining twist, tilt, or roll and for one translational distance, i.e. rise. The values for these parameters are presented in Table II.

**Hydrophobic Core**—The importance of the hydrophobic core was verified by site-directed mutagenesis of the PU.1 domain (11). Of the 14 strictly conserved residues in the domain, seven are found in the hydrophobic core. Single substitution of glycine for five of these residues in PU.1 (Fig. 3) resulted in loss of DNA binding. Two of these core residues also contact the DNA phosphate backbone. The peptide amide nitrogen of Leu$^{374}$ interacts with O2P of C-22 and the side chain NE-1 of Trp$^{315}$ forms a hydrogen bond with O1P from T-23. Mutation of tryp-

---

**Table I**

| Crystallographic refinement statistics |  |
|---------------------------------------|--|
| $R_{	ext{sym}}$ (%)                   | 3.3 |
| Resolution range (Å)                  | 6–2.1 |
| Average $B$ (Å$^2$)                   | 31.65 |
| Crystallographic $R$-factor (%)       | 22.5 |
| $R_{	ext{free}}$ (%)                  | 28.7 |
| Number of reflections used            | 22022 F >2sr(F) |
| Number of protein atoms               | 1486 |
| Number of DNA atoms                   | 1300 |
| Number of solvent atoms               | 143 |

Root mean square deviation from ideal (r.m.s. Target):

- Bond distance (Å): 0.012 (0.06)
- Bond angles (degrees): 1.629 (10)
- Dihedral angles (degrees): 1.575 (20)
tophan 215 to arginine results in loss of DNA binding in ets-1 (28, 29; see Table III). Substitutions in the hydrophobic core affect DNA binding probably because the changes disrupt the tight globular structure of the domain. Residues 174 and 215 are doubly critical for DNA binding since they represent both important structural residues in the domain core and actual DNA contact residues. In summary, residues in the hydrophobic core are critical for the formation of the overall scaffold for ets recognition.

Molecular Scaffold of ETS Domains—To evaluate the conservation of this scaffold within the ets family, the α-carbon backbones of PU.1 (11) and fli-1 (12) domains were superimposed utilizing both sequence homology and secondary structure similarities. For this purpose, a single model from the ensemble of structures deposited in the data bank was used for the NMR-derived fli-1 structure. This scaffold provides the

| Base pair | Helical twist (°) | Roll (°) | Rise (Å) | Slide (Å) | Propeller twist (°) | Buckle (°) |
|-----------|------------------|---------|----------|----------|---------------------|-----------|
| 1         | 36.09            | -0.06   | 3.18     | 0.13     | -18.22              | 11.26     |
| 2         | 39.67            | -0.48   | 3.33     | 0.06     | -16.18              | 13.88     |
| 3         | 36.09            | 0.76    | 3.30     | -0.71    | -14.38              | 1.39      |
| 4         | 36.09            | 0.78    | 3.30     | -1.00    | -17.25              | 3.38      |
| 5         | 34.02            | 3.96    | 3.47     | 0.59     | 7.77                | 5.90      |
| 6         | 35.02            | 6.75    | 3.21     | -0.37    | -14.98              | 7.43      |
| 7         | 36.02            | 9.00    | 3.11     | 0.16     | -21.87              | 10.59     |
| 8         | 39.21            | 3.24    | 3.35     | -0.61    | 19.29               | 5.58      |
| 9         | 24.41            | 0.84    | 3.38     | -0.94    | -13.13              | -10.09    |
| 10        | 37.07            | 3.75    | 3.29     | 0.98     | -12.30              | -7.49     |
| 11        | 32.70            | 9.93    | 3.30     | -0.14    | -6.40               | 3.66      |
| 12        | 33.29            | 4.99    | 3.23     | -0.02    | -10.53              | -8.09     |
| 13        | 33.29            | 4.99    | 3.23     | -0.02    | -9.98               | -2.26     |

FIG. 2. Sequence alignment of the DNA-binding domain of 33 members of the ets family. The amino acid sequence of PU.1 is listed at the top of the figure and residues that are strictly conserved in the family are enclosed in boxes. The sequences were obtained from the SWISSPROT database and original citations for these sequences are given in the database. Secondary structural features of the PU.1 ETS domain are indicated above the alignment. Directly under the PU.1 sequence, the residues that contact DNA are indicated: B, base interaction; P, phosphate backbone interaction; W, water-mediated interaction. Residues found in the hydrophobic core in PU.1 and expected to be located in the hydrophobic interior of all ets proteins are shaded. In some cases, the sequences for ets proteins for two or several species are identical, and therefore only one sequence has been listed to avoid duplication.

DNA Recognition by ets Proteins: A Crystallographic Study

DNA helical parameters of the 16-base pair oligonucleotide bound to PU.1

DNA structural parameters were refined in X-PLOR (22) and then analyzed using the programs developed by Babcock and Olson (24). For comparison, typical twist angles for B-DNA are 34.3°, roll angles are 0°, and rise values are 3.38 Å.
framework for the three structural features arranged in a loop-helix-loop pattern that mediate precise DNA binding by the PU.1 domain. In order to delineate the loop-helix-loop motif in other ets domains and to predict whether this motif is the paradigm for ets recognition, we also superimposed the α-carbon skeleton of the fli-1 domain onto the PU.1 backbone bound to the DNA (Fig. 4). Since this is one of an ensemble of structures from the NMR study, detailed comparisons are not possible. However, general comparisons are useful to establish overall structural similarities between the two related molecules. Although the structure of the fli-1-DNA complex was not determined, it should be noted that the published structure of the fli-1 domain (12) reflects a bound conformation since the NMR experiments were conducted on a 98-residue protein frag-
Contact in PU.1. Mutational substitutions for Lys219 in PU.1 backbone in PU.1 and are in a position to make the same interactions with both adjacent minor grooves (30).

**DNA Conformation in the PU.1 ETS Domain-DNA Complex**—The PU.1 ETS domain contacts DNA over a 10-base pair area. The DNA is bent by 8° in the complex but does not deviate significantly from B-form DNA (Table II). As can be seen in Fig. 4, the DNA is uniformly curved over the length of the 16-base pair fragment. There is an average helical twist of 33°, with 10.8 base pairs per turn and an average rise per base pair of 3.2 Å. The minor groove is slightly enlarged (2–3 Å from the mean) in the GGAA region at the midpoint of the oligonucleotide.

As shown in Fig. 4, there is close similarity in the overall scaffold of the ETS domains but several other features of the superposition are worth noting. First, the positions of the four conserved residues that contact DNA are very similar in PU.1 and fli-1. In PU.1, two conserved arginines, 232 and 235, make hydrogen bonds with the bases GGA of the PU core sequence. Arg<sup>235</sup>(NH-2) forms a hydrogen bond with G-8(O-6) while Arg<sup>232</sup>(NH-1) makes hydrogen bonds with two bases G-9(O-6) and A-10(N-6) on one strand and a water-mediated contact with T-23(O-4) on the opposite strand. These arginines are strictly conserved in all members of the ets family and the GGA sequence is the consensus DNA sequence recognized by the ets proteins. Therefore, these interactions are expected to be reproduced in all ets protein-DNA complexes. When the fli-1 domain is superimposed on PU.1, the side chains of conserved arginines 232 and 235 in the recognition helix are within hydrogen-bonding distance of the same bases in the GGAA core sequence in the major groove. Substitution of these residues by any other amino acid, even closely related hydrophilic amino acids results in loss of DNA recognition in PU.1, fli-1, and other ets proteins (see Table III). Conserved lysines, residues 219 in the loop (HTH) and 245 in the wing contact the phosphate backbone in PU.1 and are in a position to make the same contacts in fli-1. Mutational substitutions for Lys<sup>219</sup> in PU.1 (11) and the equivalents of Lys<sup>219</sup> and Lys<sup>222</sup> (see Table III) in fli-1 (12) or ets-1 (12) disrupt DNA binding, presumably due to the loss of the phosphate backbone interactions. In fli-1, the equivalents of Lys<sup>222</sup> and Met<sup>229</sup> in PU.1 (from the HTH loop) and residues 248/249 (from the wing loop) were identified within 4 Å of DNA by intermolecular NOEs (12). Chemical mapping experiments with the murine ets-1 molecule suggested a similar pattern with a major groove contact zone and interactions with both adjacent minor grooves (30).

**DNA Recognition by ets Proteins: A Crystallographic Study**

![Figure 4](http://www.jbc.org/)

**Comparison of PU.1 and fli-1 ETS domains.** In this stereo image, the α-carbon backbone of the fli-1 ETS domain (thin line; residues 276–373), determined in solution by NMR (12) was superimposed on the PU.1 backbone (bold line) using constraints to match similar structural features. The DNA shown in the figure is the oligonucleotide bound to the PU.1 domain.

![Figure 5](http://www.jbc.org/)

**Sequence of the oligonucleotide bound to the PU.1 protein in the crystal structure.** The GGAA recognition core sequence as well as the bases on the complementary strand are enclosed in a box. The PU.1 domain makes contacts with bases on both strands within this core. The dots designate seven phosphates that are neutralized by interactions with basic residues. With the exception of the phosphate at base 14, all of these phosphates lie on one face of the DNA helix.

The DNA bending that is stabilized by the PU.1 domain may serve as an illustration of the hypothesis of DNA bending by phosphate neutralization. It has been demonstrated, by the introduction of neutral methylphosphonate analogues in DNA fragments bearing polyadenylate tracts (32) that bending of the DNA occurs when the phosphate charges are neutralized on one face of the DNA helix, due to repulsion of the remaining anionic phosphates. It was proposed (32) that bending of proteins with cationic surfaces to DNA could also cause the DNA double helix to “spontaneously relax” toward the surface where cationic amino acids neutralized phosphate anions through formation of salt bridges. The PU.1 ETS domain makes neutralizing contacts with phosphate groups on one face of the DNA helix, involving consecutive phosphates on either side of the major groove. The sites of phosphate neutralization are shown on the DNA sequence in Fig. 5. On the GGAA strand, neutralizing contacts with the phosphate backbone 5’ to the core sequence are made by Lys<sup>208</sup> and Lys<sup>245</sup> from the wing. On the complementary strand, the phosphate contacts are 5’ to the core sequence as well as with the phosphate backbone within the core: Arg<sup>173</sup>, Lys<sup>219</sup>, and Lys<sup>223</sup> from the HTH loop and Lys<sup>229</sup> from helix α3. As predicted by the neutralization exper-
DNA Recognition by ets Proteins: A Crystallographic Study

Actions in the minor groove involve conserved residues, Lys219 of the HTH motif in PU.1 with classic HTH proteins (43, 44) or HNF-3 in the complexes of oligonucleotides with paired homeodomains of the DNA double helix (34). The moderate DNA bending seen in the complex, since only one lysine and one arginine form phosphate-side chain salt bridges. The arginine is the equivalent of Arg235 in PU.1 that forms a hydrogen bond with base G-8 in the GGA core.

Target Specificity—The superimposed models in Fig. 4 suggest that a loop-helix-loop scaffold that brings together conserved amino acids and conserved DNA bases is a general mode of DNA recognition by ets proteins. Yet, ets transcription factors bind to the GGA(A/T) core motif in the context of specific promoters. To begin to identify residues that influence target specificity, it is necessary to look for mutations of non-conserved residues that affect DNA binding. Of the 14 absolutely conserved residues in the domain, seven contact DNA in the PU.1-DNA complex. These contacts would be expected to be maintained for all ets-DNA complexes. In studies of a number of members of the ets family, mutations have been reported that affect DNA binding. These mutations, summarized in Table III, can now be correlated with the atomic model of the PU.1-DNA complex. Some of these residues are conserved residues, but others are unique to a particular molecule.
It should be emphasized that PU.1 contacts both strands at the GGAA core. Interactions are made by conserved residues as well as residues where sequence variability exists in the ets family. Therefore, ets recognition requires specific base contacts with the GGAA sequence and the bases on the complementary strand. For example, it has been shown that a single residue converts DNA recognition of ets proteins from GGAA to GGAT. When a lysine in chicken ets-1 (equivalent to residue 229 in PU.1) is altered to threonine found in this position in Elf-1 and E74, the resultant protein exhibits a restricted selectivity for GGAA like the Elf-1/E74 proteins and the reverse mutation causes the converse change in DNA recognition (38). In the PU.1 complex, Lys229 is located in the recognition helix and makes a water-mediated contact to base C-25 on the anti-sense strand at the GGAA core. Twelve well-defined water molecules are hydrogen-bonded to the bases and also form a hydrogen-bonded network between the two strands. This water network may contribute to the stability of the duplex and consequently influence specific DNA recognition.

Since the side chain of lysine is long, it is possible that the contact of a shorter residue such as threonine would not bind to this water network and could contact a different base, i.e. T-23. The water network itself could also change. Or, the interchange of lysine→threonine could permit DNA contact reflecting the stereochemical difference in size of adenine versus thymine bases.

**HTH Motif**—All of the direct contacts with specific bases in the PU.1-DNA complex are made by residues in the α3 recognition helix. Two non-conserved residues, Thr226 and Gln228, at the amino-terminal end of this helix, make water-mediated contacts with bases C-25 and C-26, respectively, that are base paired to guanines 8 and 9 in the core GGAA sequence. Both of these residues are unique to PU.1/SpiB in the ets family, so these may represent PU.1-specific contacts.

Tyr227, which is strictly conserved in the ets family, is located in the hydrophobic interior of the protein. While the phenyl ring of this tyrosine is buried, the hydroxyl group is exposed and lies within 3.6 Å of G-6(O1P). This residue was not included in our list of DNA contacts using a conservative cut-off.
of 3.2 Å for hydrogen bonds/ionic interactions. Although this interaction may not occur in PU.1, with a simple side chain rotation, a hydrogen bond is possible with the phosphate backbone. This may be an example of a contact made by a conserved residue that influences DNA recognition by selected family members. Substitution of cysteine for this tyrosine abolishes DNA binding in ets-1 (28).

In Fig. 6a, the sequence of the HTH motif of PU.1 is compared with the sequence of “classic” bacterial HTH proteins and other winged-HTH proteins. The glycine required in the turn between helices in HTH proteins (39) is also conserved in this position in ETS domains, although the α2 helix is one turn longer than the helix in HTH proteins. In PU.1, the glycine lies in the last turn of this helix. This glycine and other hydrophobic residues in α2 and α3 stabilize the arrangement of these two helices in HTH proteins. Even this pattern of conserved hydrophobic residues is seen in ets proteins. In other winged-HTH proteins, HNF-3γ (16) or heat shock factor (40), the sequence similarities are not as apparent. These two proteins have prolines in the equivalent position of the conserved glycine and the presence of this proline may influence the configuration in the “turn.” On the other hand, ets proteins may exhibit a helical arrangement that is structurally closer to that in “classic” HTH proteins. When HTH elements of PU.1 and HTH molecules such as λ (41) or 434 cro (42) repressors are superimposed, the glycine is in a structurally equivalent position (not shown). Moreover, the overall pattern of docking of the recognition helix in the major groove is quite similar when 434 cro repressor (42), CAP (15), and PU.1 are compared bound to DNA (Fig. 7). The major difference is the fact that the recognition helix in PU.1 docks deep in the major groove with contacts to the bases involving residues along the entire length of the helix, while DNA contacts in CAP and other classic HTH proteins are made from residues at the amino-terminal portion of the helix.

None of the related proteins in the HTH superfamily actually contact DNA by residues in the HTH turn (43, 44). This novel DNA contact may be possible in PU.1, as well as other ets proteins, because the connecting segment between helices is more of a loop than a turn. The corresponding HTH motifs of heat shock factor (40) and CAP (15) are compared to PU.1 in Fig. 6b. But it is not simply the length of the “turn” or “loop” in the HTH motif that accounts for this DNA contact in PU.1, since other eukaryotic HTH proteins contain even longer connecting segments (43, 44) and yet do not contact DNA by this structural feature, for example HNF-3γ (16). Thus the contacts made by this loop in PU.1 illustrate a new DNA contact that, to date, is unique to the ets proteins as the newest members of the HTH superfamily.

Loops and Minor Groove Contacts—Since the sequences in the HTH loop as well as the loop (wing) between strands β3 and β4 are not strictly conserved among members of the ets family, these residues may be important sites for specific recognition by individual members of the family. In the PU.1-DNA complex, these two loops contact the minor groove through interactions with the phosphate backbone closest to the major groove. It is also interesting to note that the length of both of the contact loops differs among members of the family, with the PU.1 loop containing an “extra” glycine at residue 220 and lacking a glycine after residue 247. Other residues in these loops may also provide specific contacts to bases in other ets proteins. For example, the change of arginine→aspartic acid (equivalent to 244 in PU.1) affects DNA binding in Elk-1 (45).

Since ets proteins bind DNA as monomers, it could be expected that there would be extensive contacts to stabilize the interaction. HNF-3γ also binds DNA as a monomer (16). In the HNF-3γ complex, three regions were involved in DNA recognition: the recognition helix and two wings. The location of the first wing between the last two strands in the β-sheet corresponds topologically to the wing in PU.1, but contacts from the second wing emanate from a loop at the COOH terminus of the domain. The structural equivalent of this second loop is absent in PU.1. In CAP, the major DNA contacts are made from the recognition helix. This protein binds DNA as a dimer. The surface area on CAP that is buried on DNA binding is 1187 Å². Similarly, the surface area buried when 434 cro repressor binds DNA is 1306 Å². But with the formation of the DNA complex with the PU.1 ETS domain, 1701 Å² surface area is buried. The significantly greater surface area of the PU.1 domain covered reflects the extensive protein-DNA contact region extending for more than 30 Å (11).

The PU.1-DNA model suggests that residues from the two loops contribute the critical interactions for recognition of bases other than the conserved GGAA core when the core is embedded in specific promoter sequences. The loops approach segments of the DNA that are adjacent to the conserved core sequence and therefore these interfaces are stereochemically suitable to permit sequence-specific interactions by a given family member while maintaining the consensus interactions at GGAA/AT). Moreover, the contacts from these loops may mediate specific base interactions by stabilizing a bend toward the protein. Future extensive mutational studies of amino acids that contact DNA are needed to identify these residues. Ultimately, crystal structures of other ets proteins complexed to DNA can be compared to distinguish unique DNA contacts.

Acknowledgments—We thank Dr. Roger Fourme and the staff at the LURE synchrotron for consultation and support for this project. In addition, we are grateful to John Knight and Rick Mitchell for synthesis and purification of oligonucleotides, and to Kelly Riddle-Hilde for preparing the manuscript for publication.

REFERENCES

1. Moreau-Gachelin, F. (1994) Biochim. Biophys. Acta 1198, 149–163
2. Wasylyk, B., Hahn, S. L., and Giovane, A. (1993) Eur. J. Biochem. 211, 7–18
3. Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C., and Maki, R. A. (1988) Cell 61, 113–124
4. Watson, D. K., McWilliams-Smith, M. J., Munn, M. F., Duesberg, P. H., O’Brien, S. J., and Papas, T. S. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 7294–7298
5. Reddy, E. S., and Rao, V. N. (1988) Oncogene Res. 3, 239–246
6. Reddy, E. S., Rao, V. N., and Papas, T. S. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6131–6135
7. Rao, V. N., Huebner, K., Jaseb, M., ar-Rushdi, A., Croce, C. M., and Reddy, E. S. (1989) Science 244, 66–70
8. Wasylyk, C., Kercheva, J.-P., and Wasylyk, B. (1992) Genes Dev. 4, 965–974
9. Lelong, P., Krust, N., Pramptonn, W. K. (1994) EMBO J. 13, 641–652
10. Petersen, J. M., Skalicky, J. J., Donaldson, L. W., McIntosh, L. P., Alber, T., and Graves, B. J. (1995) Science 269, 1866–1869
11. Kodandapani, R., Pio, F., Ni, C.-Z., Piccialli, L., Klemsz, M., McKercher, S., Maki, R. A., and Ely, K. R. (1996) Nature 380, 456–460
12. Liang, H., Mao, X., Oleiniczak, E. T., Nettesheim, D. G., Yu, L., Meadows, R. P., Thompson, C. B., and Fesik, S. W. (1994) Nature Struct. Biol. 1, 871–876
13. Donaldson, L. W., Petersen, J. M., Graves, B. B., and McIntosh, L. P. (1996) EMBO J. 15, 125–134
14. Werner, M. H., Clore, G. M., Fisher, C. L., Fisher, J. R., Trinh, L., Shiloh, J., and Gronenborn, A. (1997) Nature Struct. Biol. 4, 761–772
15. Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991) Science 253, 1001–1007
16. Clark, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993) Nature 364, 412–420
17. Pio, F., Ni, C.-Z., Mitchell, R. S., Knight, J., McKercher, S., Klemenz, M., Lombardo, A., Maki, R. A., and Ely, K. R. (1995) J. Biol. Chem. 270, 24258–24263
18. Leslie, A. (1994) MOSFLM User Guide, Mosflm version 5.20, MRC Laboratory of Molecular Biology, Cambridge, England
19. Collaborative Computational Project, Number 4 (1994) Acta Crystallogr. B50, 385–402
20. Furey, W., and Swaminathan, S. (1990) Am. Cryst. Asso. Meeting 18, 73
21. Jones, T. A. (1986) Methods Enzymol. 115, 157–171
22. Brünger, A. T. (1992) X-PLOR Manual, Version 3.1, Yale University, New Haven, CT
23. Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. T., and Berman, H. M. (1996) Acta Crystallogr. D52, 57–64
24. Balbach, J. (1995) J. Mol. Biol. 237, 98–124
25. Holm, L., and Sander, C. (1994) Nucleic Acids Res. 22, 3600–3609
26. Holm, L., and Sander, C. (1995) J. Mol. Biol. 253, 123–138
27. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E. Jr., Brice, M. D.,
DNA Recognition by ets Proteins: A Crystallographic Study

23337

Rodgers, J. R., Kennard, O., Shimomouchi, T., and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542

28. Mavrothalassitis, G., Fisher, R. J., Smyth, F., Watson, D. K., and Papas, T. S. (1994) Oncogene 9, 425–435

29. Wang, C.-Y., Petryniak, B., Ho, I.-C., Thompson, C. B., and Leiden, J. M. (1992) J. Exp. Med. 175, 1391–1399

30. Nye, J. A., Petersen, J. M., Gunther, C. V., Jensen, M. D., and Graves, B. J. (1992) Genes Dev. 6, 975–990

31. Drew, H. R., and Dickerson, R. E. (1981) J. Mol. Biol. 151, 535–556

32. Strauss, J. K., and Maher, L. J., III (1994) Science 266, 1829–1834

33. Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993) Nature 365, 520–527

34. Xu, W., Rould, M. A., Jun, S., Desplan, C., and Pabo, C. O. (1995) Cell 80, 639–650

35. Wilson, D. S., Guenther, B., Desplan, C., and Kuriyan, J. (1995) Cell 82, 709–719

36. Shore, P., Bisset, L., Lakey, J., Walthe, J. P., Virden, R., and Sharrocks, A. D. (1995) J. Biol. Chem. 270, 5805–5811

37. Bosselut, R., Levin, J., Adjadi, E., and Ghysdael, J. (1993) Nucl. Acids. Res. 21, 5184–5191

38. Brennan, R. G., and Matthews, B. W. (1989) J. Biol. Chem. 264, 1903–1906

39. Harrison, C. J., Bohn, A. A., and Nelson, H. C. M. (1993) Science 263, 224–227

40. Beamer, L. J., and Pabo, C. O. (1992) J. Mol. Biol. 227, 177–196

41. Mondragon, A., and Harrison, S. C. (1991) J. Mol. Biol. 219, 321–334

42. Pabo, C. O., and Sauer, R. T. (1992) Annu. Rev. Biochem. 61, 1053–1095

43. Brennan, R. G. (1992) Curr. Opin. Struct. Biol. 2, 100–108

44. Janknecht, R., Zinck, R., Ernst, W. H., and Nordheim, A. (1994) Oncogene 9, 1273–1278

45. Ramachandran, G. N., and Sasisekharan, V. (1968) Adv. Protein Chem. 23, 263–437

46. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Cryst. 26, 203–209

47. Soudant, N., Albaghi, O., Dhordain, P., Fleurins, A., Stéhélin, D., and Leprince, D. (1994) Nucl. Acids Res. 22, 3871–3879
New Insights on DNA Recognition by ets Proteins from the Crystal Structure of the PU.1 ETS Domain-DNA Complex
Frédéric Pio, Ramadurgam Kodandapani, Chao-Zhou Ni, William Shepard, Michael Klemsz, Scott R. McKercher, Richard A. Maki and Kathryn R. Ely

J. Biol. Chem. 1996, 271.23329-23337.
doi: 10.1074/jbc.271.38.23329

Access the most updated version of this article at http://www.jbc.org/content/271/38/23329

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 46 references, 13 of which can be accessed free at http://www.jbc.org/content/271/38/23329.full.html#ref-list-1
The DNA junction-resolving enzyme endonuclease VII of bacteriophage T4 contains a zinc-binding region toward the N-terminal end of the primary sequence. In the center of this 39-amino acid section (between residues 38 and 44) lies the sequence HLDHDHE, termed the His-acid cluster. Closely related sequences are found in three other proteins that have similar zinc-binding motifs. We have analyzed the function of these residues by a site-directed mutagenesis approach, modifying single amino acids and studying the properties of the resulting N-terminal protein A fusions. No sequence changes within the His-acid cluster led to a change in zinc content of the protein, indicating that these residues are not involved in the coordination of zinc. We found that the N-terminal aspartate residue (Asp-40) and the two histidine residues (His-41 and His-43) within the cluster are essential for junction-cleavage activity of the proteins. However, all sequence variations within this region generate proteins that retain their ability to bind to four-way DNA junctions (with minor changes in binding affinity in some cases) and to distort their global structure in the same manner as active enzymes. We conclude that the process of cleavage can be uncoupled from those of binding to and distortion of the junction. It is probable that some amino acid side chains of the His-acid cluster participate in the phosphodiester cleavage mechanism of endonuclease VII. The essential aspartate residue might be required for coordination of catalytic metal ions.

The DNA junction-resolving enzymes are a class of nucleases that recognize the structure of branched DNA molecules. Such enzymes are important in DNA recombination and repair for the processing of four-way DNA junctions created as intermediates (1–10). Junction-selective nucleases have been isolated from bacteriophage-infected eubacteria (11, 12), Escherichia coli (13–16), yeast (17, 18), and mammalian cells (19, 20) and their viruses (21) and are very probably ubiquitous cellular enzymes.

These proteins are fundamentally structure-selective. For example, the complexes formed between T7 endonuclease I, T4 endonuclease VII, or yeast CCE1, and four-way DNA junction are not displaced by a 100-fold excess of duplex DNA of the same sequence (22–24). Although RuvC of E. coli and yeast CCE1 exhibit significant sequence specificity, this is manifested at the level of the cleavage reaction, and these enzymes bind to four-way junctions of any sequence (24, 25). The existence of mutants of T7 endonuclease I and T4 endonuclease VII that bind normally to DNA junctions but are defective in cleavage suggests that binding and catalysis are separable events. While the binding of resolving enzymes is selective for the structure of DNA junctions, the act of binding in general also distorts the global configuration of helical arms (22, 26, 27).

Endonuclease VII of T4 is required during late infection in order to resolve DNA branch points prior to packaging of the DNA into phage heads (11, 28). Examination of the primary sequence of endonuclease VII (29) suggests the existence of three sections that might form modules within the overall protein structure. There is a region at the C terminus that is 48% identical to a sequence within the pyrimidine dimer glycosylase and nuclease T4 endonuclease V. The structure of this repair enzyme is known (30), and the region of similarity comprises a helix and an extended section. Replacement of the region of endonuclease VII with the corresponding sequence of endonuclease V resulted in a chimeric protein that retained its specificity for the precise cleavage of four-way DNA junctions (23). In the center of the endonuclease VII sequence is a section with some similarity to a region of the functionally related resolving enzyme T7 endonuclease I. We have previously found that in selection of non-functional mutants of T7 endonuclease I, all such mutants map within this region of the protein (22), suggesting that it may comprise a significant part of the active site for DNA cleavage. We have shown that a mutation within the corresponding part of the primary sequence of T4 endonuclease VII results in a catalytically inactive protein (27).

The N-terminal section of endonuclease VII contains a 40-amino acid region bounded by two Cys-X-X-Cys motifs that binds an atom of zinc (23). This region is 42% identical to a section found in a protein (gp59) of unknown function encoded by mycobacteriophage L5 (31). In addition, an open reading frame identified downstream of the secF gene of E. coli and Salmonella typhimurium (32) encodes a 109-amino acid protein of unknown function that includes a region that is 43% identical with this section, which is also bounded by Cys-X-X-Cys motifs. The four sequences are collected together in Fig. 1. The comparison reveals a number of conserved features, including arginine (position 28 in endonuclease VII), asparagine (position 31), and glycine (position 51). In addition, there is a conserved four-residue sequence Asp-His-Asp-His beginning with aspartate 40 in endonuclease VII. Indeed, this cluster of histidine and acidic residues can be extended in endonuclease VII, beginning with histidine 38, to read HLDHDHE. Acidic residues are frequently involved in the catalytic sites of nucleases (33–36), where they coordinate metal ions that participate in the chemistry of phosphodiester bond cleavage, and we were curious to learn whether any or all of these residues might be involved in catalysis. We have therefore made point mutants of endonuclease VII in which amino acids within the cluster of
The Histidine-Aspartate Cluster of T4 Endonuclease VII

Figure 1. The potential domain structure of T4 endonuclease VII. A schematic map of the regions of endonuclease VII, showing the locations of amino acid residues changed in this study. The N-terminal section of the protein binds a zinc ion (23) that is coordinated by four cysteine residues, and the sequence of this region is shown. The His-acid cluster at the center of the section is highlighted in larger type. The C-terminal section is similar to a region of T4 endonuclease V. The central section has some similarity to a region of T7 endonuclease I, and the catalytically inactivating Es6A mutation (27) lies in this region. B, an alignment of sequences of probable zinc-binding domains in T4 endonuclease VII, gp59 of mycobacteriophage L5 (31), and open reading frames identified in the secD locus of E. coli and S. typhimurium (32). The sequences are aligned by their CXXC sequences. Note the presence of the His-acid cluster within each of these sequences.

Materials and Methods

Synthesis of Oligonucleotides—Oligonucleotides were synthesized using β-cyanoethyl phosphoramidite chemistry (44, 45) implemented on a 394 DNA/RNA synthesizer (Applied Biosystems). Fully deprotected oligonucleotides were purified by gel electrophoresis in 10% polyacrylamide containing 7 M urea, the bands were excised, and DNA was eluted and recovered by ethanol precipitation.

Construction of Four-way DNA Junctions—A four-way DNA junction with four arms of 25 bp1 was generated by the hybridization of four oligonucleotides of 50 nucleotides each, one of which was radioactively 5'-32P-labeled. The oligonucleotides were based upon the sequence of junction 3 studied by Duckett et al. (39); thus, the b strand had the sequence 5' CCGAGGAGCTCCGTCTCGTTGCAGAGAAGG 3', which was used for changing the histidine 38 to a glutamine was 5' CACACGTCCAAGCTAATCAACTTGACCAC 3'.

For the comparative gel electrophoretic analysis of the global structure of the complex between junction and endonuclease mutants, six junctions comprising two long arms of 40 bp and two short arms of 15 bp were each generated by hybridization of appropriate oligonucleotides based on the sequence of junction 3 as described above. As an example, the BH junction, where the B and H arms are long) was obtained using the following four oligonucleotides: b strand, 5' CCGAGGAGCTCCGTCTCGTTGCAGAGAAGG 3'; c strand, 5' CCGAGGAGCTCCGTCTCGTTGCAGAGAAGG 3'; d strand, 5' GTCCGAACGTCCAAGCTAATCAACTTGACCAC 3'; e strand, 5' GTCCGAACGTCCAAGCTAATCAACTTGACCAC 3'; f strand, 5' GTCCGAACGTCCAAGCTAATCAACTTGACCAC 3'.

In each case, stoichiometric quantities of three unlabeled and one radioactively 5'-32P-labeled strands were annealed by incubation in 50 mM Tris-HCl (pH 7.6), 10 mM MgCl2, 5 mM dithiothreitol, 0.1 mM spermidine, and 0.1 mM EDTA for 3 min at 85 °C, followed by slow cooling. Assembled junctions were purified by gel electrophoresis in 5% polyacrylamide (acrylamide/bisacrylamide, 20:1) gels and stained with Coomassie Brilliant Blue R-250. The DNA was excised and DNA was recovered by electroelution. DNA concentrations were measured spectrophotometrically at 260 nm, using an extinction coefficient of ε = 6.5 × 10^5 M^-1 cm^-1 bp^-1.

Enzymes—Oligonucleotides were radioactively labeled at their 5' termini using [γ-32P]ATP and T4 polynucleotide kinase (Amersham). Ligation of DNA was performed using T4 DNA ligase (Amersham) under standard conditions (46).

Cloning of Endonuclease VII to Encode an Oligohistidine Fusion Protein—A fragment containing part of the synthetic gene encoding endonuclease VII was obtained by digestion of pAT153-SEVII (23) by HindIII and BamHI. The sequence lost from the 5' end of the gene was restored by hybridization of two oligonucleotides, giving a DNA fragment containing an NcoI site at the 5' end and a HindIII site at the 3' end; this also added the coding sequence for 10 histidine residues and a site for proteolytic cleavage by enterokinase. This was ligated into pET-19b (previously digested by NeoI and BamHI) to generate the plasmid pET-SEVII (8), which was used to transform the E. coli strain HMS174(DE3) pLYsS.

Mutagenesis by Polymerase Chain Reaction—Mutagenesis of the synthetic endonuclease VII gene was performed by polymerase chain reaction (PCR) using one primer that differed in sequence from the synthetic gene sequence by one or two nucleotides (therefore changing the appropriate codon) and one non-mutated primer. In order to maximize the stability of the hybrid, at least seven nucleotides were placed between the mutagenic mismatch(es) and the ends of the oligonucleotide primers. In some cases, the mutated primer was extended to the nearest restriction site to facilitate the cloning. For example, the sequence of the primer used for changing the histidine 38 to a glutamine was 5' CGAACGCATCCGAACGTCCAAGCTAATCAACTTGACCAC 3', where the NcoI cloning site and mutagenic adenine base are underlined. The other primer corresponded to a sequence downstream to the targeted area, 5' TGACGTCCAAGCTAATCAACTTGACCAC 3'.

PCR reactions were performed as described by Landt et al. (47) using 1 unit of Taq DNA polymerase, 1 ng of plasmid DNA, 100 pmol of each primer, 50 μM dNTPs in 50 mM Tris-HCl (pH 9), 1 mM MgCl2, 0.1% Triton X-100 for 30 cycles (1 min at 93 °C, 1 min at 45 °C, and 45 s at 72 °C). The PCR products were digested with the restriction enzymes AatII and BIII, and the fragment was purified by gel electrophoresis and used to replace the wild-type sequence in pK19SEVII (23).

DNA Sequencing—The base sequences of wild-type and mutated genes were obtained by primer extension-dideoxy sequencing (48).

Preparation of Oligohistidine-Endonuclease VII Fusion Protein from E. coli—Endonuclease VII as an oligohistidine fusion protein was prepared from 1 liter of E. coli strain HMS174 (DE3) pLYsS transformed with pK19SEVII. The cells were grown to an A660 of 0.6 and then induced with IPTG to a final concentration of 1 mM for 2 h. Cells were harvested, resuspended in 20 ml of 20 mM Tris-HCl (pH 7.5), 0.5 mM NaCl, 1 mM imidazole, and lysed by sonication. Unbroken cells and cell debris were removed by centrifugation (40,000 × g for 15 min). The protein was purified by affinity chromatography using a Fractogel EMD chelate column previously charged with nickel chloride. The protein was eluted using a gradient of imidazole from 0 to 500 mM in 20 mM Tris-HCl (pH 7.5), 0.5 mM NaCl. The protein-containing fractions were pooled and dialyzed for 2 h at 4 °C against 50 mM Tris-HCl (pH 7.4), 1 mM DTT, 50% glycerol and were stored at −20 °C.

Preparation of Protein A-Endonuclease VII Fusion Proteins from E. coli—E. coli strain JM101 transformed with the appropriate protein A-fusion plasmids based on the pK19 system (37) was grown in A660 of 0.6 and induced with 0.5 mM IPTG for 2 h. Cells were harvested by centrifugation and resuspended in 20 ml of 20 mM MES (pH 6). Cells were lysed by sonication as described above. Two ammonium sulfate precipitation steps were performed. Ammonium sulfate was added to 40% saturation and the precipitate was discarded. Further ammonium sulfate was added to the supernatant to a final concentration of 65% saturation. The pellet was redissolved in 5 ml of a solution of 20 mM MES (pH 6), 1 mM DTT and was applied to an S-Sepharose ion-exchange column. A gradient of NaCl in the same buffer was applied to the column. The peak fractions containing the protein were dialyzed against 20 mM Tris-HCl (pH 7.4), 1 mM dithiothreitol, 50% glycerol for 2 h at 4 °C. The concentration of the protein was determined by the Bradford method, calibrated against a previous amino acid analysis for endonuclease VII-H41T (23). The purity of the proteins obtained was verified by polyacrylamide gel electrophoresis in the presence of SDS.

Preparation of Endonuclease VII H38T Without Fusion Polypeptide—Protein A-endoronuclease VII H38T was prepared as described above, but after elution from the S-Sepharose column the pooled fractions were not...
dialyzed in the glycerol-containing solution. The buffer was changed by centrifugation in a Centriplus concentrator (Amicon) to 20 mM sodium phosphate (pH 7.4), and the sample was concentrated to 5 ml. The protein was digested overnight with a ratio of 1:500 (w/v) of protease factor Xa. The digested protein was reapplied to the S-Sepharose column. The released protein A was not retained by the resin, and non-fusion endonuclease VII H38T was eluted using a gradient of NaCl in 20 mM MES (pH 6), 1 mM DTT. The protein-containing fractions were dialyzed in 20 mM Tris-Cl (pH 7.4), 1 mM DTT, 50% glycerol for 2 h at 4°C.

Cleavage of Four-way DNA Junctions—Reactions were performed on ice in 10 μl of 112 mM four-way DNA junction 3 individually 5′-32P-labeled on either the b, h, r, or x strand and endonuclease VII or mutant protein in 50 mM Tris-Cl (pH 7.4), 50 mM NaCl, 1 mM EDTA. The samples were loaded on a 10% polyacrylamide denaturing gel (acylamide/bisacylamide, 29:1). After electrophoresis the gels were dried onto Whatman 3MM paper and subjected to autoradiography at ~70 °C using Fuji RX x-ray film with Ilford fast tungstate intensifier screens.

Gel Electrophoretic Retardation Analysis of Protein A-Endonuclease VII Mutants and Measurement of the Apparent Binding Constants—Varying amounts of each mutant and wild-type protein were incubated with 24.2 ± 5 ′32P-labeled junction for 10 min at room temperature in 10 μl of binding buffer (50 mM Tris-Cl (pH 7.4), 100 mM NaCl, 1 mM dithiothreitol, and either 1 mM EDTA or 200 μM MgCl2) for 20 min. The reactions were terminated by addition of 10 μl of formamide, 50 mM EDTA. The samples were loaded on a 10% polyacrylamide denaturing gel (acylamide/bisacylamide, 29:1) in the presence of either TBE or TBM. Electrophoresis was performed on an Applied Biosystems 420H analyzer.

RESULTS

Construction and Expression of Endonuclease VII Mutants—We have constructed a series of point mutants within the His-acid cluster of T4 endonuclease VII by means of site-directed mutagenesis of the synthetic gene described previously (23). A section of the gene was replicated by means of the PCR that included one mutagenic primer. The amplified fragment was cloned as translational fusions with protein A in the plasmid pK19PRA (37) and transformed in E. coli JM101. Expression was under the control of the lac promoter and was induced by the addition of IPTG. Following ammonium sulfate precipitation, the protein A fusion polypeptides were purified by ion exchange chromatography. The endonuclease VII variants could be released from the protein A fusion by digestion with Factor Xa protease.

The wild-type sequence protein was studied as an N-terminal oligohistidine fusion, while all the mutant proteins were analyzed as N-terminal protein A fusions. Dissociation constants were calculated by measuring the extent of binding to DNA junctions of the proteins as a function of their concentration, fitting the data as described under "Materials and Methods." These data were all measured in the absence of added magnesium ions. For protein A-endonuclease VII H38T the errors are the standard error obtained from three independent experiments; for the other proteins the errors are derived from the fit of individual data points. Zinc stoichiometries were measured using a colorimetric assay. The errors are the random error on the data points in the absorption plateau region, and the full experimental error is probably larger than this.

| Mutation | Activity | Kd (nM) | Zinc content |
|----------|----------|---------|--------------|
| Wild type | Active | 38 ± 16 | 1.1 ± 0.08 |
| H38T | Active, but thermosensitive | 45 ± 2 |
| H38Q | Active, but thermosensitive | 25 ± 16 |
| H38S | Active, but thermosensitive | 5.3 ± 5 |
| D40N | Inactive | 20 ± 8 |
| D40A | Inactive | 20 ± 19 |
| H41T | Inactive | 96 ± 47 |
| D42N | Active | 25 ± 20 |
| D42A | Active | 28 ± 15 |
| H43T | Very low activity | 6 ± 2 |

The Histidine-Aspartate Cluster of T4 Endonuclease VII

Table I

Mutations introduced into the His-acid region of endonuclease VII and their properties

The histidine-aspartate cluster of T4 endonuclease VII (23) was engineered by site-directed mutagenesis of the synthetic gene described previously (23). A section of the gene was replicated by means of the PCR that included one mutagenic primer. The amplified fragment was cloned as translational fusions with protein A in the plasmid pK19PRA (37) and transformed in E. coli JM101. Expression was under the control of the lac promoter and was induced by the addition of IPTG. Following ammonium sulfate precipitation, the protein A fusion polypeptides were purified by ion exchange chromatography. The endonuclease VII variants could be released from the protein A fusion by digestion with Factor Xa protease.

The wild-type sequence protein was studied as an N-terminal oligohistidine fusion, while all the mutant proteins were analyzed as N-terminal protein A fusions. Dissociation constants were calculated by measuring the extent of binding to DNA junctions of the proteins as a function of their concentration, fitting the data as described under "Materials and Methods." These data were all measured in the absence of added magnesium ions. For protein A-endonuclease VII H38T the errors are the standard error obtained from three independent experiments; for the other proteins the errors are derived from the fit of individual data points. Zinc stoichiometries were measured using a colorimetric assay. The errors are the random error on the data points in the absorption plateau region, and the full experimental error is probably larger than this.

Table I

Mutations introduced into the His-acid region of endonuclease VII and their properties

The wild-type sequence protein was studied as an N-terminal oligohistidine fusion, while all the mutant proteins were analyzed as N-terminal protein A fusions. Dissociation constants were calculated by measuring the extent of binding to DNA junctions of the proteins as a function of their concentration, fitting the data as described under "Materials and Methods." These data were all measured in the absence of added magnesium ions. For protein A-endonuclease VII H38T the errors are the standard error obtained from three independent experiments; for the other proteins the errors are derived from the fit of individual data points. Zinc stoichiometries were measured using a colorimetric assay. The errors are the random error on the data points in the absorption plateau region, and the full experimental error is probably larger than this.

Table I

Mutations introduced into the His-acid region of endonuclease VII and their properties
Wild-type sequence endonuclease VII was also expressed as a fusion with an N-terminal oligohistidine sequence by transferring the gene into the plasmid pET-19b. The protein was purified by affinity chromatography on a column to which nickel ions were chelated.

The purity of the proteins was analyzed by polyacrylamide gel electrophoresis in a buffer containing SDS, and the preparations were generally found to contain a single polypeptide migrating at the position expected for the calculated mass (Fig. 2).

The single-amino acid changes introduced into endonuclease VII are summarized in Table I. In general we have altered histidine residues to threonine, glutamine, or serine, and aspartate residues to asparagine or alanine.

Mutants with Altered Sequences in the His-Acid Cluster Have Normal Zinc Content—The His-acid cluster is centrally located within the zinc-binding region of endonuclease VII. While previous studies have strongly implicated the four cysteine residues in the coordination of the zinc ion (23), we could not exclude some role for other amino acids, particularly the histidine residues. We therefore measured the zinc content of mutants representative of each position (as N-terminal protein A fusions) using a colorimetric assay (38). The results are summarized in Table I, where it can be seen that each of the mutant proteins analyzed contains 1 mol of zinc/mol of protein within the probable experimental error. Thus, the zinc content of the protein has not been altered by mutation of any of these residues from the wild-type sequence, strongly indicating the lack of a role in zinc coordination for these amino acids.

Activity of Mutant Proteins in the Cleavage of Four-way DNA Junctions—Wild-type sequence endonuclease VII cleaves four-way DNA junctions with considerable selectivity. It cleaves junctions with the central sequence of junction 3 on two diametrically related strands (called b and r), three bases from the point of strand exchange. The activities of the mutant proteins (as N-terminal protein A fusions) were examined under stand-
and conditions, using a four-way DNA junction with the central sequence of junction 3 of Duckett et al. (39) (Fig. 3), and the results are summarized in Table I. It is clear that aspartate 40 and histidine 41 are essential to activity, because all mutations of these residues result in total loss of detectable activity (note that we have used a 17-fold higher concentration of the inactive mutant proteins (1 μM) compared with the active ones (60 nM)). Histidine 43 is also important, because alteration to threonine leads to an almost total loss of activity. By contrast, aspartate 42 can be replaced by asparagine or alanine without detectable loss of activity. Histidine 38 can be changed to threonine with retention of activity, but H38T has an interesting thermal dependence of junction retention of activity, but H38T has an interesting thermal dependence of junction 3.

**Fig. 4.** Binding of endonuclease VII and derived mutant proteins to a four-way DNA junction. 24.2 nM of radioactively 5'-32P-labeled (h strand) junction 3 was incubated with increasing concentrations of endonuclease VII-derived mutant proteins as N-terminal protein A fusions for 10 min at room temperature. Free junction and DNA-protein complexes were separated by electrophoresis in polyacrylamide and visualized by autoradiography. The protein-junction complex migrates as a retarded species relative to the free junction (both species indicated on left in A). A, binding of protein A-endonuclease VII H38T to junction 3. Tracks 1–16, protein concentrations (calculated for dimeric species) of 1.2, 2.4, 4.8, 7.3, 12.1, 17, 19.4, 24.2, 30.3, 36.4, 42.4, 48.5, 54.5, 60.6, 66.6, and 72.7 nM, respectively. B, binding of protein A-endonuclease VII H41T and H43T to junction 3. Tracks 1–8, protein A-endonuclease VII H41T concentrations (calculated for a dimeric species) of 1.2, 3.2, 5.6, 10, 17.8, 31.7, 56.2, and 100 nM, respectively; tracks 9–16, protein A-endonuclease VII H43T concentrations (calculated for dimeric species) of 1.2, 3.2, 5.6, 10, 17.8, 31.7, 56.2, and 100 nM, respectively; tracks 19–25, protein A-endonuclease VII H43T concentrations (calculated for dimeric species) of 1.2, 3.2, 5.6, 10, 17.8, 31.7, 56.2, and 100 nM, respectively. C, binding of protein A-endonuclease VII D40N and D42N to junction 3. Tracks 1–8, protein A-endonuclease VII D40N concentrations (calculated for dimeric species) of 1.2, 3.2, 5.6, 10, 17.8, 31.7, 56.2, and 100 nM, respectively. Tracks 9–16, protein A-endonuclease VII D42N concentrations (calculated for dimeric species) of 1.2, 3.2, 5.6, 10, 17.8, 31.7, 56.2, and 100 nM, respectively.

**Fig. 5.** Binding isotherms for endonuclease VII and derived mutant proteins binding to a four-way DNA junction. Extent of protein binding to four-way junctions as a function of total protein concentration was estimated by gel electrophoresis (see the legend to Fig. 4). The fraction of DNA junction bound to protein was calculated for each protein concentration and plotted against the protein molarity (calculated for a dimeric species) on a logarithmic scale. The data were fitted to a model for the binding process (see under "Materials and Methods") from which the binding affinities were calculated. The points plotted are experimental data, and the lines are simulations derived using the association constants derived from the fits. A, binding of protein A-endonuclease VII H38T. Three independently measured sets of data are plotted, differentiated by the use of three different plotting symbols. The line was calculated for a $K_a = 3.29 \times 10^7$ M$^{-1}$, B, binding of different protein A-endonuclease VII histidine mutants. Data shown for the mutant sequences H38T (●), H38Q (○), H38S (□), H41T (△), and H43T (▲). The lines were calculated for $K_a = 2.0 \times 10^7$ M$^{-1}$, 3.29 $\times 10^7$ M$^{-1}$, and 7.39 $\times 10^6$ M$^{-1}$. C, binding of protein A-endonuclease VII aspartate 40 mutants. Data are shown for the mutant sequences D40N (●) and D40A (○). The line was calculated for a $K_a = 5.4 \times 10^7$ M$^{-1}$. D, binding of protein A-endonuclease VII aspartate 42 mutants. Data are shown for the mutant sequences D42N (●) and D42A (○). The lines were calculated for $K_a = 7.0 \times 10^7$ M$^{-1}$ and 2.35 $\times 10^8$ M$^{-1}$.

**Binding to Four-way DNA Junctions—Endonuclease VII of wild-type sequence binds selectively to four-way DNA junctions.** This is also true of a non-catalytic mutant protein endonuclease VII E86A; this protein (as either N-terminal fusions or non-fusion) binds to four-way junctions in the presence or absence of magnesium ions and is not displaced by a 1000-fold excess of duplex competitor of the same sequence (27).

We examined the binding of the proteins that were mutated in the His-acid cluster to radioactively labeled four-way DNA junctions. All were found to bind DNA junctions. Binding titrations were carried out using gel electrophoretic retardation in the presence of 1 mM EDTA (Fig. 4). The titrations are well behaved for all the proteins, giving increasing fractions of a single retarded species as the protein concentration is raised. At protein concentrations higher than 100 nM some super-retarded species could be found in some cases, and thus such data were not used in the calculation of binding affinities.

The ratios of bound and free junction were quantified by phosphorimaging, from which apparent dissociation constants (KD) were calculated assuming binding of a dimeric species (Fig. 5). Most of the proteins bound with affinities that were close to that of the wild-type sequence ($K_D = 20\text{–}40$ nM). Protein A-endonuclease VII H38S and H41T had higher affinity ($K_D = 5$ nM), while H41T had lower affinity ($K_D = 96$ nM). These results indicate that the loss in activity of the proteins with sequence alterations at Asp-40, His-41, and...
The Histidine-Aspartate Cluster of T4 Endonuclease VII

The binding affinities of inactive mutant proteins could also be measured in the presence of magnesium ions. We found that protein A-endonuclease VII H41T bound around 2-fold more tightly in the presence of 200 μM magnesium ions. The binding affinity of protein A-endonuclease VII D40N was increased by a factor of 1.3 under the same conditions.

Distortion of the Global Structure of Junctions on Binding Endonuclease VII Variants—On binding to junctions, endonuclease VII induces a change in the global configuration of arms, demonstrated by comparative gel electrophoresis studies (27). In this method a four-way junction with arms of 40 bp each in length is subjected to shortening of two arms by restriction cleavage in the six possible combinations (39–41). The electrophoretic mobility in polyacrylamide of these six two-long, two-short arm species are compared and analyzed on the basis of the expected relationship (42) between electrophoretic mobility and the angle included between the two long arms. We used this method originally to analyze the structure of the free junction under different conditions (39), but it has more recently been applied to junction-protein complexes (22, 27, 43).

Both endonuclease VII H38T and E86A induce a change in the global folding of DNA four-way junctions (27). The same structure is generated by the inactive mutant endonuclease VII E86A in either the presence or absence of added magnesium ions and is different from that of the free DNA junction under either set of conditions. It is clear that the four-way junction is extensively manipulated by endonuclease VII, and we asked whether the mutant proteins retained the ability to induce the same structural alteration.

The binding of protein A-endonuclease VII to junction 3 generates a pattern of electrophoretic mobilities described by intermediate-slow-intermediate-intermediate-fast-intermediate-slow pattern indicative of the stacked X structure with B on X (39). This structure of this stacking isomer gives rise to the three pairs of long-short arm species in which the included angles between the long arms are acute (BH and RX are slow species), obtuse (BR and HX are intermediate species), or linear (BX and RR are fast species). This interpretation is summarized in the schematic of the stacked X structure on the right. The pattern of mobilities for the long-short species of the junction-protein complex is clearly different from that of the free DNA and is the same as that found in the presence of EDTA (compare with A). The electrophoretic pattern (and thus the global structure of the DNA junction in the presence of this mutant) is unchanged by the presence or absence of magnesium ions, just as was found previously for the complex with protein A-endonuclease VII E86A. Track 1, BH species; track 2, BR species; track 3, BX species; track 4, HR species; track 5, HX species; track 6, RX species.

His-43 is not due to impairment in substrate binding, since the mutant proteins D40N, D40A, H38T, and H38Q bind normally, and H43T has a 3-fold higher affinity than the enzyme of wild-type sequence.

Further studies have shown that the binding properties of the various endonuclease VII variants are similar to those of wild-type enzyme. The variants of endonuclease VII appear to impose the same global structure on the four-way DNA junction. All of the variants of endonuclease VII appear to impose the same global structure on the four-way DNA junction. All of the variants of endonuclease VII appear to impose the same global structure on the four-way DNA junction. All of the variants of endonuclease VII appear to impose the same global structure on the four-way DNA junction.
in terms of a principal binding across arms H and X (in which the cleavages are introduced by the active enzyme in the presence of magnesium ions) and a rotation of arms B and R toward arms H and X, respectively, together with a movement out of the plane (27).

Fig. 6A compares the electrophoretic patterns of the complexes of junction 3 with protein A fusions of the three histidine-to-threonine mutants of endonuclease VII, in the presence of 1 mM EDTA (TBE buffer) to prevent cleavage by active enzyme. The patterns of mobilities of the six long-short species are identical for all three proteins, indicating that all three induce the same global conformation of arms on the four-way DNA junction. The experiment was repeated for protein A endonuclease VII H41T in the presence of 200 μM magnesium ions (TBM buffer), conditions where the free DNA junction folds into the stacked X structure. Since this mutant is completely inactive as a nuclease, the experiment can be carried out in the presence of magnesium ions without inducing cleavage of the DNA. Despite the change of conditions, the complex clearly has the same global structure, resulting in an unchanged pattern of electrophoretic mobilities (Fig. 6B).

The aspartate mutants also generated the same structure in the four-way junction. Fig. 6C shows the comparative gel electrophoretic analysis of the complexes of the six long-short variants of junction 3 with endonuclease VII D40N and D42N in the presence of 1 mM EDTA (TBE buffer). Once again the pattern is unchanged from those of all the endonuclease VII variants.

These results suggest that the binding process is very similar for all of the mutants of the His-acid cluster studied. This is further evidence that the lack of catalytic activity in some mutant proteins was not due to impairment of binding.

DISCUSSION

The mutation analysis confirms that amino acids contained within the conserved His-acid cluster are important in the function of T4 endonuclease VII. In particular, one aspartate (Asp-40) and two histidine residues (His-41 and His-43) are required for cleavage of DNA junctions. Since mutation of these residues leads to only small changes in binding affinity for DNA junctions, it is likely that they are involved (directly or indirectly) in the catalyzed phosphodiester bond hydrolysis. The zinc content of none of the mutants was significantly altered from 1 mol/mol of protein, and thus a role in the coordination of zinc is not likely. We have previously shown (27) that another acidic residue, glutamate 86, is essential for catalytic activity. This is located in the second region of the primary sequence of the protein, which exhibits some similarity to endonuclease I of T7. It is therefore possible that the active site of the enzyme comprises amino acid side chains from both of these regions of the polypeptide. We suspect that acidic side chains may be involved in the coordination of a magnesium ion required for the cleavage reaction. This is a common feature of nucleases (33–35), including the E. coli junction-resolving enzyme RuvC (36), where the catalytic center contains three aspartate residues and one glutamate residue.

All the endonuclease VII sequence variants examined retained selective binding to DNA junctions. This is a further indication of the divisibility of binding and catalysis in this enzyme. Small changes in binding affinities were measured over the set of mutant proteins, but this difference was only about 20-fold between the tightest (endonuclease VII H38S) and weakest binding (endonuclease VII H41T) protein, corresponding to an overall difference in binding free energy of 1.7 kcal mol−1. In most cases the differences are much smaller than this. This suggests that sequence changes in the His-acid cluster tend to affect interactions primarily with the transition state rather than the ground state DNA structure. It is interesting to note that the mycobacteriophage L5 gp59 protein has a serine at the position corresponding to histidine 38 in endonuclease VII. This protein therefore has the same sequence at this position as the tightest binding endonuclease VII variant.

An interesting feature of the binding of endonuclease VII to four-way junctions is the change in the global structure of the DNA (27). Distortion of DNA structure upon binding of junction-selective proteins appears to be rather general, having also been observed for T7 endonuclease I (22), E. coli RuvA (43) and RuvC (26), and yeast CCE1.2 The distortion imposed on the global configuration of helical arms in the junction by endonuclease VII is quite significant, and we have proposed a model of the structure of the bound junction involving an unstacking of the arms at the point of strand exchange (27). All of the sequence variants studied here appear to induce the same change in junction structure, whether active or inactive, despite small differences in binding affinity. This further supports the contention that the basic binding processes are unaltered by any of the sequence changes in the His-acid cluster. This is also indicated by the fact that mutations within the region result in either proteins that cleave at exactly the same positions as the wild-type enzyme or fail to cleave at all. None causes an alteration in the cleavage pattern that might be expected if the manner of substrate binding had been changed.

In summary, we have found that a number of sequence changes in the His-acid cluster of endonuclease VII lead to reduced activity of the enzyme. In particular aspartate 40 and histidines 41 and 43 appear to be required for cleavage of DNA junctions. However, none of these mutations appears to affect binding to DNA junctions (beyond relatively small changes in affinity) or the distortion of DNA structure. It is therefore quite likely that the His-acid cluster will prove to be important in generating the active site of T4 endonuclease VII.

Acknowledgments—We thank our colleagues Drs. Richard Pöhler and Malcolm White for many valuable discussions and the Cancer Research Campaign for generous financial support.

REFERENCES
1. Holliday, R. (1964) Genet. Res. 5, 282–304
2. Broker, T. R., and Lehman, I. R. (1971) J. Mol. Biol. 60, 131–149
3. Orr-Weaver, T. L., Szostak, J. W., and Rothstein, R. J. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 6534–6538
4. Potter, H., and Dressler, D. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 3000–3004
5. Potter, H., and Dressler, D. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 3696–3702
6. Kita, P. A., and Nash, H. A. (1987) Nature 329, 346–348
7. Nunes-Duby, S. E., Matsumoto, L., and Landy, A. (1987) Cell 50, 779–788
8. Hoess, R., Wierzchicki, A., and Abremski, K. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 6840–6844
9. Jayaram, M., Crain, K. L., Parsons, R. L., and Harshay, R. M. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 7902–7906
10. McCulloch, K., Coggins, L. W., Collins, S. D., and Sherratt, D. J. (1994) EMBO J. 13, 1844–1855
11. Kemper, B., and Garabett, M. (1981) Eur. J. Biochem. 115, 123–131
12. de Massy, B., Studier, F. W., Dorgai, L., Appelbaum, F., and Weisberg, R. A. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 715–726
13. Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A., and Shinagawa, H. (1991) EMBO J. 10, 4381–4389
14. Connolly, B., Parsons, C. A., Benson, F. E., Duderla, H. J., Sharpley, G. J., Lloyd, R. G., and West, S. C. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 6063–6067
15. Sharples, G. J., and Lloyd, R. G. (1991) J. Bacteriol. 173, 7711–7715
16. Sharples, G. J., Chan, S. N., Mahal, A. A., Whitty, M. C., and Lloyd, R. G. (1994) EMBO J. 13, 6133–6142
17. West, S. C., and Korner, A. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 6455–6459
18. Symington, L., and Kolodner, R. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 7247–7251
19. Elborough, K. M., and West, C. S. (1990) EMBO J. 9, 2931–2936
20. Hyde, D., Davies, A. A., Benson, F. E., and West, S. C. (1994) J. Biol. Chem. 269, 5202–5209
21. Stuart, D., Ellison, K., Graham, K., and McFadden, G. (1992) J. Virol. 66, 1551–1561
2. M. F. White and D. M. J. Lilley, unpublished data.
The Histidine-Aspartate Cluster of T4 Endonuclease VII 33155

22. Duckett, D. R., Giraud-Panis, M.-J. E., and Lilley, D. M. J. (1995) *J. Mol. Biol.* 246, 95–107
23. Giraud-Panis, M.-J. E., Duckett, D. R., and Lilley, D. M. J. (1995) *J. Mol. Biol.* 252, 586–610
24. White, M. F., and Lilley, D. M. J. (1996) *J. Mol. Biol.* 257, 330–341
25. Bennett, R. J., Dunderdale, H. J., and West, S. C. (1993) *Cell* 74, 1021–1031
26. Bennett, R. J., and West, S. C. (1995) *J. Mol. Biol.* 252, 213–226
27. Pohler, J. R. G., Giraud-Panis, M.-J. E., and Lilley, D. M. J. (1996) *J. Mol. Biol.* 260, 678–696
28. Kemper, B., and Janz, E. (1976) *J. Virol.* 18, 992–999
29. Tomaszewski, J. (1988) *Endonuklease VII des bakteriophagen T4: sequenzierung und überexpression.* Ph.D. thesis, Universität Bochum
30. Morikawa, K., Ariyoshi, M., Vasylyev, D. G., Matsumoto, O., Katayanagi, K., and Ohtsuka, E. (1985) *Nature* 313, 154–156
31. Cooper, J. P., and Hagerman, P. J. (1987) *J. Mol. Biol.* 198, 711–719
32. Lumpkin, O. J., and Zimm, B. H. (1982) *Biopolymers* 21, 2315–2316
33. Beaucage, S. L., and Caruthers, M. H. (1981) *Tetrahedron Lett.* 22, 1859–1862
34. Sinha, N. D., Biernat, J., McManus, J., and Koster, H. (1984) *Nucleic Acids Res.* 12, 4539–4557
35. Landt, O., Grunert, H.-P., and Hahn, U. (1990) *Gene (Amst.)* 96, 125–128
36. Saito, A., Iwasaki, H., Ariyoshi, M., Morikawa, K., and Shinagawa, H. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 7470–7474
37. Zueco, J., and Boyd, A. (1992) *Anal. Biochem.* 207, 348–349
38. Giedroc, D. P., Keating, K. M., Williams, K. R., Konigsberg, W. H., and Coleman, J. E. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 8452–8456
39. Duckett, D. R., Murchie, A. I. H., Diekmann, S., von Kitzing, E., Kemper, B., and Lilley, D. M. J. (1988) *Cell* 55, 79–89
40. Gough, G. W., and Lilley, D. M. J. (1985) *Nature* 313, 154–156
41. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Vol. 1, p. 168, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
42. Hunt, J. B., Neece, S. H., Schachman, H. K., and Ginsberg, A. (1984) *J. Biol. Chem.* 259, 14793–14803