CHAPTER III

STRUCTURE DETERMINATION OF THE AUTO-TRANSPHOSPHORYLATION DOMAIN (DOMAIN A) OF ENVZ

III-1. INTRODUCTION

The cytoplasmic region of EnvZ comprises three enzymatic activities: auto-transphosphorylation, OmpR kinase, and phospho-OmpR phosphatase. Most of histidine kinases exhibit similar three biochemical activities as EnvZ. Among these three activities, the first event occurring in response to the osmolarity change is the auto-transphosphorylation by ATP at a conserved histidine residue (His-243).

In the His-Asp phosphorelay system, the signal transduction that responds chemotaxis is catalyzed by a protein distinct from a histidine kinase. Dephosphorylation of phospho-CheY is stimulated by CheZ, not the histidine kinase CheA. In addition, the amino acid sequence around the conserved autophosphorylation site of CheA (P1 domain including His-48) shows poor similarity to EnvZ. Nevertheless, CheA contains the conserved catalytic region among histidine kinases and performs the His-Asp phosphorelay signal transduction. This signal transduction system, therefore, can be interpreted as follows: the phospho transfer itself is a pinpointing and common event between the conserved histidine and aspartic acid residues, subject to the specific rescognition of the cognate protein as another prerequisite event which preferentially involves several residues. CheA study (Zhou *et al.*, 1996), and other phosphotransfer chemistry on small-molecule phosphodonors (Feng *et al.*, 1992; Lukat *et al.*, 1990; McCleary and Stock, 1994; Schroder *et al.*, 1994) have proposed the aspect of this pinpointing phospho-transfer event. The autophosphorylation by ATP provides a phospho-histidine (3-imidazole-phosphate), and subsequent kinase activity transfers the phosphate to a response regulator that involves a phospho-aspartic acid (carboxyphosphate). The query remains how they interact with each other.

EnvZ domain A (residues 223-289) contains the conserved auto-transphosphorylation site, His-243 (Figure III-1). Recent report by Park *et al.* (1998) indicated that the autophosphorylation domain (domain A: residues 223-289) is responsible for the dimer formation of EnvZ and also the binding of EnvZ to the cognate receptor, OmpR. In this chapter, the solution structure of domain A is elucidated by means of NMR. The structure of domain A reveals a four-helix bundle in dimer and each identical subunit contains helix-loop-helix motif. Close to the solvent exposed His-243, helices are somewhat loosened and has a negatively charged acidic patch at the interface of two subunits.

This loosened structure region (242-248) contains well-conserved sequence motif in histidine kinases. The structure observation suggests the importance of inter-subunit surface of domain A that could contribute to the interaction with the other protein such as OmpR.

III-2. MATERIALS AND METHODS

III-2-1. Sample preparation

EnvZ domain A (residues 223-289) was provided by Prof. M. Inouye (UMDNJ, USA). It was expressed and purified as previously described (Egger and Inouye, 1997) except that uniformly ^{15}N- or $^{15}N/^{13}C$-labeled proteins were obtained by using $^{15}NH_4Cl$ or $^{15}NH_4Cl/[^{13}C_6]$-D-glucose as the sole nitrogen and nitrogen/carbon sources in M9 medium, respectively. NMR samples contained 1.0 to 1.5 mM uniformly ^{15}N- or $^{15}N/^{13}C$-labeled, or unlabeled protein in 95 % H_2O/5% ^2H_2O containing 20 mM sodium phosphate, 50 mM KCl, 0.5 mM AEBSF (ICN Pharmaceutical Inc.), 50 μM sodium azide, 5 mM perdeuterated dithiothreitol (DTT), and 5 mM $MgCl_2$. pH was adjusted to 7.0 without correction for deuterium isotope effects.

The mass of the EnvZ domain A was determined by light scattering and gel filtration analyses. They indicated that domain A is dimeric in solution (see Appendix Table A-1.).

III-2-2. NMR spectroscopy

All NMR experiments were performed at 25 °C, unless otherwise noted, on a Varian four-channel UNITY 500 spectrometer, equipped with an actively z gradient shielded triple-resonance probe and a PFG driver. All data were processed with nmrPipe and nmrDraw programs (Delaglio *et al.*, 1995), and analyzed with PIPP and STAPP (Garrett *et al.*, 1991).

On the 10% ^{13}C-labeled and 90% unlabeled sample, 2D ^1H-^{13}C CT-HSQC (Vuister and Bax, 1992) spectra was recorded on a UNITY 600 spectrometer with the following number of complex points and acquisition times: ^{13}C (F1) 160, 27 ms, ^1H (F2) 384, 48 ms with 16 transients. A homonuclear 2D NOESY (Jeener *et al.*, 1979) spectrum was recorded also on a UNITY 600 spectrometer, using following numbers of complex points and acquisition times: ^1H (F1) 512, 85 ms, ^1H (F2) 1024, 171 ms with 64 transients.

Three-dimensional HNCO, (HB)CBCA(CO)NNH (Grzesiek and Bax, 1992), and HNCACB (Wittekind and Mueller, 1993) spectra were recorded with the following numbers of complex points and acquisition times: ^{13}C (F1) 80, 53 ms, ^{15}N (F2) 32, 25 ms, ^1H (F3) 512, 64 ms with 8 transients for HNCO, and 13C (F1) 48, 6 ms, 15N (F2) 32, 25 ms, ^1H (F3) 512, 64 ms with 24 transients for (HB)CBCA(CO)NNH, and ^{13}C (F1) 38, 5 ms, ^{15}N (F2) 28, 22 ms, ^1H (F3) 512, 64 ms with 64

transients for HNCACB. Three-dimensional HCCH-TOCSY (Bax *et al.*, 1990; Kay *et al.*, 1993) spectrum was recorded with the following numbers of complex points and acquisition times: ^{1}H (F1) 128, 37 ms, ^{13}C (F2) 28, 9 ms, ^{1}H (F3) 416, 52 ms with 16 transients. A 3D H(CCO)NH and C(CO)NH (Grzesiek *et al.*, 1993) spectrum was recorded with the following numbers of complex points and acquisition times: ^{13}C (F1) 56, 7 ms, ^{15}N (F2) 32, 25 ms, ^{1}H (F3) 384, 64 ms with 40 transients.

Three-dimensional simultaneous ^{15}N/^{13}C-edited NOESY-HSQC (Pascal *et al.*, 1994) and ^{1}H-^{15}N HMQC-NOESY-HMQC (Ikura *et al.*, 1990) spectra were recorded on a UNITY 600 spectrometer with the following numbers of complex points and acquisition times: ^{1}H (F1) 120, 20 ms, ^{13}C (F2) 28, 9 ms, ^{1}H (F3) 384, 48 ms with 16 transients for simultaneous ^{15}N/^{13}C-edited NOESY-HSQC and ^{15}N (F1) 32, 21 ms, ^{15}N (F2) 32, 21 ms, ^{1}H (F3) 384, 64 ms with 32 transients for ^{1}H-^{15}N HMQC-NOESY-HMQC. A 3D [^{13}C/F_1]-edited [^{13}C/F_3]-filtered HMQC-NOESY (Lee *et al.*, 1994; Zwahlen *et al.*, 1997) spectrum was recorded with the following numbers of complex points and acquisition times: ^{1}H (F1) 96, 19 ms, ^{13}C (F2) 32, 13 ms, ^{1}H (F3) 352, 53 ms with 16 transients.

Three-dimensional HNHB (Archer *et al.*, 1991) and HNHA (Vuister and Bax, 1993) spectra were recorded with the following numbers of complex points and acquisition times: ^{1}H (F1) 54, 10 ms, ^{15}N (F2) 32, 29 ms, ^{1}H (F3) 512, 64 ms with 24 transients for HNHB, and ^{1}H (F1) 64, 16 ms, ^{15}N (F2) 16, 32 ms, ^{1}H (F3) 512, 64 ms with 16 transients for HNHA. Heteronuclear $^{3}J_{NH\beta}$ coupling constants were estimated from a 3D HNHB spectrum and $^{3}J_{NH\alpha}$ from a HNHA spectrum.

Exchanging rate of amide protons was estimated from the difference in the intensities of ^{1}H-^{15}N HSQC cross peaks between the HSQC spectra taken with and without water presaturation (Spera *et al.*, 1991).

III-2-3. Structural constraints

NOEs were assigned from 2D homonuclear and 3D heteronuclear NOESY spectra. Distance restraints between subunits were mainly obtained by a 3D [^{13}C/F_1]-edited [^{13}C/F_3]-filtered HMQC-NOESY recorded on a sample containing a 1:1 mixture of ^{15}N/^{13}C-labeled and unlabeled protein. All

NOEs were grouped into four distance ranges: 1.8-2.9 Å, 1.8-3.5 Å, 1.8-5.0 Å, and 1.8-6.0 Å corresponding to strong, medium, weak and very weak NOEs. Standard pseudo-atom distance corrections (Wüthrich *et al.*, 1983) were incorporated to account for center averaging. An additional 0.5 Å was added to the upper limits for distances involving methyl groups to account for the higher apparent intensity of methyl resonances.

The values of $^{3}J_{NH\alpha}$ give the backbone Φ dihedral angle constraints of domain A were obtained: for residues with $^{3}J_{NH\alpha} < 6$ Hz, Φ was restrained to -50° ± 40°. Ψ angle constraints (-50° ± 50°) were included only for those residues within regular α-helical conformation.

Two distance restraints, $r_{NH\text{-}O}$ (1.8-2.3 Å) and $r_{N\text{-}O}$ (2.3-3.3 Å), was used for each hydrogen bond which is identified by analyzing the $^{1}H/^{2}H$ exchange rates of amide protons.

III-2-4. Structure calculation

Structure calculations were performed using the YASAP protocol (Nilges *et al.*, 1988) with non-crystallographic symmetry restraints (NCS) (Nilges, 1995) within X-PLOR 3.1 (Brünger, 1993). Calculations employed 1697 interproton distance restraints (comprising 632 intraresidue, 480 sequential, 512 short-range, 46 long-range, 27 inter-subunit) supplemented with 100 distance restraints for 50 hydrogen bonds and 168 dihedral angle restraints. Thirty structures with the lowest total energy were selected out of the 60 calculated structures.

III-2-5. Sequence alignment

Sequence alignment for domain A was performed as described in Chapter II. The accession numbers of the SWISS-PROT database for the 11 selected histidine kinases are P02933 (EnvZ), P08982 (EnvZ from *Salmonella typhimurium*), P18392 (RstB), P08336 (CpxA), P39928 (SLN1 from *Saccharomyces cerevisiae*), Q02541 [COPS from *Pseudomonas syringae (pv. tomato)*], P51392 (YC26 from *Porphyra purpurea*), P08400 (PhoR), P06712 (NTRB), Q06067 (ATOS) and P22763 (ArcB). Most of them are *E. coli* proteins except where otherwise noted.

III-3. RESULTS

III-3-1. Resonance assignments

The strategy used here for structure determination of EnvZ domain A (residues 223-289) is similar to that described in Chapter II. Figure III-2 summarizes the assignment of backbone amide ^{1}H and ^{15}N correlations in the ^{1}H-^{15}N HSQC (Kay *et al.*, 1992) spectrum of domain A. The previous biochemical studies (Hidaka *et al.*, 1997; Park *et al.*, 1998) demonstrated that EnvZ domain A forms a stable homodimer with an apparent molecular weight of 19 kDa. The ^{1}H-^{15}N HSQC spectrum of this domain displays about 80 highly dispersed peaks in the backbone NH region, which approximately correspond to the expected number of backbone NH atoms excluding the N-terminal three residues. This observation is consistent with the dimer formation of two identical subunits (each containing 67 backbone NHs and 14 side chain NH_2) related by a two-fold symmetry. The line widths of HSQC peaks also indicate that the overall rotational correlation time of this domain corresponds to that of a 16-18 kDa globular protein.

The HNCACB and (HB)CBCA(CO)NNH pair of experiments produced nearly complete backbone and Cß chemical shift assignments for EnvZ cytoplasmic domain A, excluding only these residues without crosspeaks in ^{1}H-^{15}N HSQC spectrum (Pro-248 and the N-terminal 3 residues). Analysis of ^{13}C-separated spectra produced backbone assignments for all of the remaining residues except Met-223, Ala-224, and Ala-225, suggesting that they have been cleaved off during bacterial expression.

Side-chain resonances were mainly assigned by their correlations to $^{13}C_{\alpha}$-$^{1}H_{\alpha}$ and $^{13}C_{\beta}$-$^{1}H_{\beta}$ crosspeaks in HCCH-TOCSY spectrum. The carbon chemical shifts were obtained also by the ^{13}C-edited NOESY-HMQC spectrum. Backbone and side-chain ^{1}H, ^{13}C, and ^{15}N chemical shifts are reported in Table III-1.

III-3-2. Secondary structure

The secondary structure of domain A are summarized in Figure III-3. In NMR studies of proteins, helices are characterized by the presence of dαN(i, i+3) connectivities and $^{3}J_{NH\alpha}$ coupling

constants that are less than 6 HZ (Wüthrich, 1986). Two helical regions are estimated from the pattern of chemical shift index (CSI) (Venters *et al.*, 1996; Wishart *et al.*, 1992), short range NOEs (i to i+3 and i to i+4, especially), and are further supported by relatively strong sequential amide proton (NH) to amide proton (NH) crosspeaks and small $^3J_{NH\alpha}$ coupling constants. Evidence of two large helices also includes some slowly exchanging backbone amide protons. Within the helix I, residues around His-243 indicate the poor helical character: the backbone NH groups in this region exhibit fast exchanging rates, and no $^3J_{NH\alpha}$ coupling constant data nor long-range NOEs were obtained because of line broadening.

III-3-3. Structure description

A total of 1797 nuclear Overhauser effect (NOE)-derived distance restraints including 27 inter-subunit NOEs and 100 hydrogen bond distance restraints, and 168 dihedral angle restraints were used to determine the protein structure. In order to distinguish between the intra-subunit and inter-subunit NOEs, [$^{13}C/F_1$]-filtered [$^{13}C/F_3$]-edited HMQC-NOESY (Ikura *et al.*, 1992; Lee *et al.*, 1994) was measured on 1:1 mixture of $^{13}C/^{15}N/^{1}H$:$^{12}C/^{15}N/^{1}H$-labeled EnvZ domain A that enabled to observe specifically the NOEs from protons attached to ^{13}C to protons attached to ^{12}C. The NOEs obtained from this experiment proved the close contacts between the ^{13}C labeled subunit and the non-labeled subunit. Consequently, this fragment forms a homodimer in solution. The superposition of 30 simulated annealing structures calculated by Xplor 3.1 (Brünger, 1993) is shown in Figure III-4a. The pairwise root-mean-square (rms) deviation of the NMR-derived structure is 1.10 Å for backbone heavy atoms and 1.90 Å for all heavy atoms only for those residues in α-helices. A summary of structural statistics is given in Table III-2.

The homodimer of EnvZ domain A forms a four-helix bundle with a two-fold symmetry along the helix axis. Each subunit folds into a compact structure consisting of antiparallel helices (helix I, residues 235-255 and helix II, 265-286) connected by a nine-residue long turn (Figure III-4a and 4b). The packing of the two identical subunits is schematically illustrated in Figure III-5. Long amphipathic helices (21 residues in helix I and 22 residues in helix II) are responsible for the elongated shape of the four-helix bundle structure, with approximate overall dimensions of 13 Å x 13 Å x 36 Å (the flexible N-terminal residues 223-234 are omitted). Helix I of subunit A interacts

intimately with helix II of the same subunit as well as helix II of subunit B in an antiparallel manner (interhelical angles are 175 ± 2° and 172 ± 3°, respectively). Consequently, the same helices of subunits A and B are positioned orthogonally in a parallel manner to each other (interhelical angles are 12 ± 4° for helix I-helix I and 5 ± 3° for helix II-helix II, respectively). The hydrophobic core is made of numerous methyl-containing residues (Leu-237, Met-238, Val-241, Leu-245, Ile-252, and Ala-255 in helix I, and Leu-266, Ile-270, Ile-274, Ile-280, Ile-281, and Ile-285 in helix II from each subunit) and a phenylalanine residue (Phe-284) in helix II.

The carbon chemical shifts of proline, especially Cß and Cγ are affected strongly by the conformation (cis or trans) of the peptide bond containing the proline nitrogen (King *et al.*, 1986). Pro-248 has Cα, Cß, Cγ, and Cδ chemical shifts of 65.6, 31.2, 28.3, and 49.8 p.p.m., consistent with a trans peptide bond (Ikura *et al.*, 1991). NOE evidence confirms this interpretation; the α proton of Thr-247 shows strong cross peaks to the δ protons of Pro-248.

III-4. DISCUSSION

Overlooking the entire molecule, three regions stand out as poorly defined by the structure calculations; the N-terminal residues (223-233), the loop between two helices (258-264) and the C-terminal residues (286-289). Detailed analysis of the backbone dihedral angles of these three regions suggests their flexibility. These regions are defined by less than 10 NOE restraints each. Thus the poor definition of these regions in the calculated structures results from the lack of structural constraints.

The structural core of the molecule is created by hydrophobic contacts among the helices, which can be described as amphipathic (Figure III-5). The total surface area buried at the dimer interface is 1198 ± 119 Å^2 per subunit, which corresponds to 21 % of the entire accessible surface area of one subunit (5696 ± 108 Å^2). Ninety percent of the surface area buried at the dimer interface is hydrophobic. Within the four-helix bundle, residues 242-248 of helix I are somewhat poorly defined (rms deviation 1.14 Å for backbone heavy atoms and 1.99 Å for all heavy atoms when calculated with only these residues). The backbone NH groups in this region exhibit fast H/D exchange rates, suggesting that this region is accessible to solvent. In addition, line broadening, no

long-range NOE and no $^{3}J_{NH\alpha}$ coupling constant data suggest that this region undergoes a conformational equilibrium, probably between helical and unfolded states.

Helix I *versus* helix II

Despite of architectural similarity, helices I and II differ from each other with respect to their structural properties. First, helix I contains many amino acid residues that are highly conserved among members of the histidine kinase family (Figure III-6). On the other hand, helix II exhibits significantly lower homology in amino acid sequence. Secondly, helix I contains a segment (residues 242-248) which displays structural flexibility, whereas helix II is well-defined and therefore more rigid throughout. Interestingly, the flexible segment of helix I contains the most conserved residues His-243, Thr-247 and Pro-248.

Among the members of histidine kinases, structures of phosphotransfer domains of CheA (the phosphotransfer domain; residues 1-134) (Zhou *et al.*, 1995) and ArcB (the histidine-related phosphotransfer (HPt) domain; residues 654-777) (Kato *et al.*, 1997), have been determined previously. Although these domains are different from EnvZ domain A in their sequence and function, there are some similarities in structures. Both CheA and ArcB structures are composed of multiple helices involving a four-helix bundle structure. A histidine residue (His-48 of CheA and His-717 of ArcB) which receives a phosphate, is located in the middle of a long helix and largely exposed to the solvent. However, both structures are monomeric and do not have a perfect symmetry like the one found in the homodimeric structure of EnvZ. Notably, histidine residue that receives the phosphate in each structure (autophosphorylation site for EnvZ and CheA, and phosphotransfer site for ArcB) is always situated in an environment comprising multiple helices, which could provide a structural flexibility relative to all β structures.

Molecular surface of domain A

The autophosphorylation site of EnvZ, His-243, is located at the middle of helix I (Figure III-4b). The solvent-exposed side-chain projects away from the four-helix bundle, and is accessible to the catalytic domain of the same protein, and also to the response regulator OmpR. Moreover, His-

243 is situated at the edge of the intra- and inter-subunit surfaces (Figures III-4b and 5). The molecular surface of the four-helix bundle (Figure III-7) can be divided into two symmetric sections originated from the two identical subunits. One is made of helix I of subunit A and helix II of subunit B and *vise versa* (inter-subunit), while the other (i.e. intra-subunit) comprises helices I and II within the same subunit (Figures III-4b and III-5). Interestingly, the inter-subunit surface contains two acidic clusters; one consisting of Asp-244, Asp-273 and Glu-276, and the other Asp-232, Asp-233 and Asp-286. The former cluster is located in close spatial proximity to His-243, while the latter located near the both termini of domain A. In contrast to the inter-subunit surface, the intra-subunit surface encompasses virtually no obvious structural characteristics, except for the presence of Arg-246, the only basic residue near His-243.

In addition to the acidic clusters described above, the inter-subunit surface contains a hydrophobic cluster consisting of Met-238, Ala-239, Leu-236, Phe-284, Ile-280, and Ala-279. These clusters of charged and hydrophobic residues in the vicinity of His-243 are probably involved in the interactions with the catalytic domain of EnvZ (in Chapter II) (Tanaka *et al.*, 1998) and the N-terminal domain of OmpR containing the phosphotransfer site, Asp-55 as will be discussed in Chapter III.

Functional residues and molecular architecture

Numerous natural and site-directed mutagenesis studies on EnvZ have been reported previously (Aiba *et al.*, 1989; Brissette *et al.*, 1991; Forst *et al.*, 1989; Hsing and Silhavy, 1997; Kanamaru *et al.*, 1990; Russo and Silhavy, 1991; Skarphol *et al.*, 1997; Tokishita *et al.*, 1992; Waukau and Forst, 1992). The phenotype of these mutations can be classified into two groups; one with reduced kinase activity while retaining phosphatase activity (kinase-/phosphatase+), and *vice versa* (kinase+/phosphatase-). All these mutations have been conducted on helix I, with exception of those described in the recent study (Hsing *et al.*, 1998), and mapped on the top half of the four-helix bundle in the view shown in Figure III-8. The kinase+/phosphatase- mutations (shown in cyan) are spread over the top half of the four-helix bundle, whereas the kinase-/phosphatase+ mutations (shown in green) are more localized near His-243. It has been noted that no single mutation impaired both kinase and phosphatase activities, indicating any mutation summarized in Figure III-8, does not

completely eliminate the ability to interact with OmpR. Structural consequences of all the mutations studied so far probably differ from one case to another. However, they are located mainly on the inter-subunit surface rather than on the intra-subunit surface, indicating the functional importance of the four-helix bundle formation of homo-dimeric domain A. It is suggested that many of these mutations alter the structural integrity and stability of the four-helix bundle, and amplify the lability around His-243, thereby modulating the kinase or phosphatase activity.

In summary, the present study establishes the structural basis of the dimer formation of EnvZ. The four-helix bundle architecture formed by the homodimeric domain (residues 223-289) reveals a vital platform for autophosphorylation, phosphotransfer to OmpR, and transmembrane signaling. Interestingly, several residues (242-248) at and near the active site His-243 encounter the highest structure flexibility within the four-helix bundle. By virtue of the dimeric bundle structure, the inter-subunit surface provides an acidic patch near His-243, which may play a role in the interaction with OmpR.

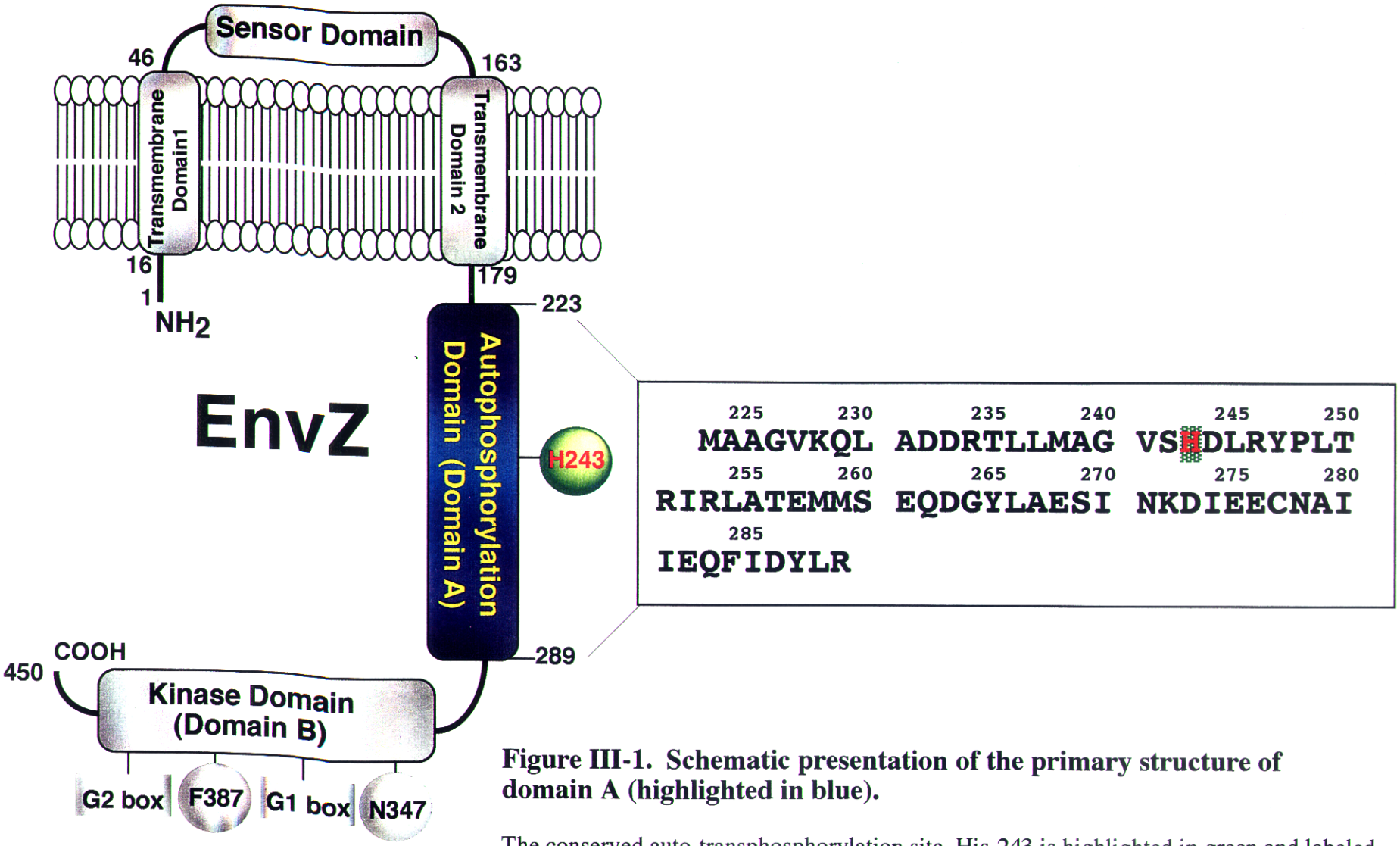

Figure III-1. Schematic presentation of the primary structure of domain A (highlighted in blue).

The conserved auto-transphosphorylation site, His-243 is highlighted in green and labeled in red.

Table III-1.

^{15}N, ^{13}C, ^{13}CO, and ^{1}H Resonance Assignments for Domain A at pH 7.2 and 25 °C[a]

residue	N	CO	C^{α}	C^{β}	others
M223					
A224					
A225		177.9	52.5 (4.41)	19.1 (1.42)	
G226	109.0 (8.49)	173.9	45.2 (3.98)		
V227	119.7 (8.00)	176.2	62.2 (4.12)	32.7 (2.11)	$C^{\gamma 1}$ 21.1 (0.96); $C^{\gamma 2}$ 20.4 (0.96)
K228	125.6 (8.44)	176.2	56.4 (4.30)	32.9 (1.81)	C^{γ} 24.8 (1.44); C^{δ} 28.9 (1.71); C^{ϵ} 42.0 (3.02)
Q229	122.3 (8.41)	176.3	55.9 (4.34)	29.5 (2.12, 2.02)	C^{γ} 33.8 (2.39)
I230	124.1 (8.33)	176.9	55.3 (4.34)	42.4 (1.67, 1.57)	C^{γ} 28.0 (1.67) $C^{\delta 1}$ 25.0 (0.88); $C^{\delta 2}$ 23.5 (0.93)
A231	124.4 (8.27)	173.7	52.5 (4.31)	19.3 (1.44)	
D232	120.0 (8.20)	176.0	54.3 (4.58)	41.2 (2.76)	
D233	120.8 (8.28)	177.2	54.9 (4.57)	41.0 (2.76)	
R234	122.1 (8.48)	177.3	58.2 (4.18)	30.1 (1.97)	C^{γ} 27.4 (1.70); C^{γ} 27.4 (1.77); C^{δ} 42.8 (3.22)
T235	115.3 (8.22)	176.0	65.2 (4.04)	68.6 (4.27)	C^{γ} 21.8 (1.28)
L236	122.9 (7.92)	179.0	56.9 (4.21)	41.6 (1.75)	C^{γ} 27.1 (1.62); $C^{\delta 1}$ 24.6 (0.85) ; $C^{\delta 2}$ 23.4 (0.86)
L237	121.9 (7.99)	178.6	57.0 (4.18)	41.9 (1.65)	C^{γ} 27.3 (1.87); $C^{\delta 1}$ 26.1 (0.85) ; $C^{\delta 2}$ 24.4 (0.85)
M238	118.2 (8.20)	177.5	56.9 (4.24)	32.1 (2.07, 2.12)	C^{γ} 32.5 (2.48); C^{γ} 32.4 (2.58)
A239	123.3 (8.15)	179.6	54.2 (4.16)	18.2 (1.48)	
G240	107.5 (8.12)	174.9	46.6 (3.96)		
V241	121.1 (7.85)	176.8	64.9 (3.84)	31.8 (1.98)	$C^{\gamma 1}$ 20.9 (0.71) ; $C^{\gamma 2}$ 22.1 (0.74)
S242	114.5 (8.05)	175.4	61.1 (4.05)	62.7 (3.89)	
H243	119.9 (7.83)	177.1	58.6 (4.39)	30.2 (3.26, 3.30)	
D244	121.0 (8.48)	178.0	56.4 (4.53)	40.1 (2.86, 2.78)	
L245	120.0 (8.32)	176.8	56.6 (4.22)	43.5 (1.50, 1.90)	$C^{\delta 2}$ 25.0 (0.95)
R246	115.1 (7.77)	178.8	59.5 (3.97)	30.1 (1.90)	C^{γ} 28.4 (1.71); C^{δ} - (3.19)
T247	116.1 (7.76)		69.0 (4.09)	66.4 (4.41)	C^{γ} 22.2 (1.26)
P248		178.1	65.6 (4.17)	31.2 (2.32, 1.89)	C^{γ} 28.4 (1.87, 2.32); C^{δ} 49.8 (3.39, 4.18)

L249	115.2 (7.04)	179.2	58.3 (3.97)	41.3 (1.95, 1.37)	C^{γ} 26.6 (1.96); $C^{\delta 1}$ 26.6 (0.82); $C^{\delta 2}$ 23.2 (0.86)
T250	116.7 (8.19)	176.0	66.7 (3.92)	68.4 (4.49)	C^{γ} 21.5 (1.27)
R251	121.4 (7.62)	180.1	61.0 (3.96)	30.9 (1.83)	
I 252	122.6 (8.11)	177.0	65.9 (3.78)	37.4 (1.98)	$C^{\gamma Me}$ 18.1 (0.85); C^{δ} 13.5 (0.75)
R253	120.6 (7.82)	179.6	59.3 (3.97)	29.6 (1.71)	C^{γ} 26.7 (1.77, 1.69); C^{δ} 43.0 (3.28, 3.15)
L254	120.3 (8.39)	179.5	57.9 (4.07)	41.5 (1.83, 1.55)	C^{γ} 27.0 (1.83); $C^{\delta 1}$ 25.1 (0.91); $C^{\delta 2}$ 23.0 (0.91)
A255	122.6 (7.91)	180.6	55.0 (4.16)	17.9 (1.54)	
T256	107.9 (8.03)	177.3	64.9 (4.06)	68.8 (4.28)	C^{γ} 22.0 (1.39)
E257	120.7 (7.44)	177.2	57.9 (4.13)	29.4 (2.19)	C^{γ} 36.3 (2.32, 2.53)
M258	115.8 (7.54)	176.0	54.9 (4.54)	33.5 (2.38, 2.27)	C^{γ} 32.4 (2.64); C^{γ} 32.3 (2.77)
M259	119.8 (7.07)	175.6	57.1 (3.97)	34.2 (1.98)	C^{γ} 33.4 (2.55, 2.75)
S260	119.2 (9.78)	176.3	58.8 (4.38)	64.3 (4.16, 4.08)	
E261	124.6 (8.92)	179.2	59.0 (4.12)	29.6 (2.14)	C^{γ} 36.2 (2.42); C^{γ} 36.2 (2.35)
Q262	119.7 (8.69)	176.3	58.2 (4.26)	28.1 (2.18)	C^{γ} 33.6 (2.51)
D263	119.1 (7.69)	176.7	53.6 (5.15)	40.8 (3.10, 2.67)	
G264	110.6 (8.00)	176.1	47.9 (4.03)		C^{α} 47.9 (4.03)
Y265	120.0 (8.57)	177.5	60.0 (4.48)	36.9 (3.15)	C^{δ} - (6.98); C^{ε} - (7.19)
L266	124.2 (7.14)	178.1	56.5 (3.97)	41.9 (1.15, 1.66)	C^{γ} 26.6 (1.67); $C^{\delta 1}$ 25.3 (0.79); $C^{\delta 2}$ 23.4 (0.77);
A267	123.3 (7.93)	179.6	55.8 (3.82)	18.4 (1.61)	
E268	117.1 (8.21)	178.9	59.3 (4.06)	29.6 (2.11)	C^{γ} 36.3 (2.26); C^{γ} 36.2 (2.41)
S269	115.7 (7.59)	176.3	61.4 (4.11)	62.8 (3.96, 4.04)	
I 270	121.9 (8.46)	178.0	65.4 (3.81)	38.1 (1.83)	$C^{\gamma 2}$ 17.5 (0.74); $C^{\delta 1}$ 13.0 (0.62)
N271	117.8 (8.32)	178.0	57.3 (4.35)	38.4 (2.87, 2.72)	N^{δ} 111.1 (6.98, 7.24)
K272	121.5 (7.99)	179.2	59.6 (4.12)	31.7 (2.02)	C^{γ} 25.0 (1.64, 1.52); C^{δ} 29.0 (1.73); C^{ε} 42.4 (3.00)
D273	123.0 (7.97)	179.3	58.3 (4.71)	41.1 (2.78, 2.48)	
I 274	123.3 (8.79)	178.2	64.9 (3.63)	37.1 (2.22)	$C^{\gamma 1}$ 28.0 (1.22, 1.73); $C^{\gamma 2}$ 17.7 (0.86); C^{δ} 12.8 (0.85)
E275	120.4 (8.10)	179.7	59.6 (4.12)	28.8 (2.37)	C^{γ} 36.3 (2.52)

E276	122.1 (8.08)	178.8	59.8 (4.19)	29.2 (2.30)	C^{γ} 36.1 (2.47, 2.25)
C277	117.6 (8.20)	176.5	64.3 (4.01)	27.2 (3.39, 2.53)	
N278	117.5 (8.56)	177.0	56.4 (4.38)	38.4 (2.84)	
A279	123.1 (8.13)	180.5	55.0 (4.18)	17.9 (1.56)	
I 280	120.7 (8.00)	178.4	64.7 (3.81)	38.1 (2.04)	$C^{\gamma 1}$ 28.9 (1.87, 1.21); $C^{\gamma 2}$ 18.0 (0.97); C^{δ} 14.0 (0.88)
I 281	119.7 (8.10)	177.6	65.1 (3.50)	38.1 (1.98)	$C^{\gamma 1}$ 29.4 (1.87); $C^{\gamma 2}$ 17.5 (0.91); C^{δ} 14.2 (0.80)
E282	118.7 (8.16)	178.5	59.1 (3.98)	29.3 (2.08)	C^{γ} 36.2 (2.31, 2.41)
Q283	117.4 (7.77)	177.6	58.1 (4.10)	28.9 (2.05, 2.11)	C^{γ} 34.0 (2.24, 2.41); N^{ε} 111.8 (6.86, 7.27)
F284	118.6 (8.00)	176.7	59.4 (4.61)	39.2 (3.30, 3.16)	C^{δ} - (7.26)
I 285	118.7 (7.98)	177.3	63.7 (3.83)	37.9 (2.00)	$C^{\gamma 1}$ 28.5 (1.65, 1.24); $C^{\gamma 2}$ 17.5 (0.93); $C^{\delta 1}$ 13.3 (0.85)
D286	120.0 (8.00)	177.4	56.3 (4.48)	40.8 (2.59)	
Y287	118.3 (7.84)	176.1	58.8 (4.48)	38.4 (3.17, 3.12)	C^{δ} - (6.84); C^{ε} - (7.14)
L288	120.1 (7.69)	176.0	54.9 (4.26)	42.3 (1.71, 1.56)	C^{γ} 26.1 (1.65); $C^{\delta 1}$ 25.5 (0.76); $C^{\delta 2}$23.1 (0.78)
R289	125.6 (7.49)		58.1 (4.07)	31.3 (1.82, 1.77)	C^{γ} 27.4 (1.68); C^{δ} 43.2 (3.22)

a ^{15}N and ^{13}C chemical shifts are given first, and the attached ^{1}H chemical shifts are in parenthesis. The chemical shift reference used for ^{1}H and ^{13}C is 3-(trimethylsilyl)propionate, sodium salt. ^{15}N chemical shift are reported relative to external liquid NH_3.

Table III-2.

Structural Statistics of the 30 Structures of EnvZ Domain A (223-289)[a]

Rms deviations from experimental distance restraints (Å)	
All (1797)	0.049 ± 0.003
Interresidue sequential ($\|i - j\| = 1$) (494)	0.042 ± 0.003
Interresidue short-range ($1 < \|i - j\| \leq 5$) (534)	0.059 ± 0.009
Interresidue long-range ($\|i - j\| > 5$) (50)	0.070 ± 0.004
Intraresidue (616)	0.036 ± 0.005
H bond (100)	0.055 ± 0.004
Inter-subunit (27)	0.090 ± 0.004
Rms deviations from experimental dihedral restraints (deg) (297)	0.60 ± 0.096
Rms deviations from idealized covalent geometry	
Bonds (Å)	0.006 ± 0.0001
Angles (deg)	0.66 ± 0.014
Impropers (deg)	0.41 ± 0.012
Energies (kcal mol^{-1})	
F_{NOE}[b]	213.9 ± 35.4
F_{cdih}[b]	3.62± 1.01
F_{repel}[c]	46.1 ± 7.1
E_{L-J}[d]	-100.7± 36.4
Average rms difference (Å)[e]	
Homodimer: Residues in Helix I (235-255) and Helix II(268-280)	1.10 (1.90)
One subunit: Residues in Helix I (235-255) and Helix II (268-280)	0.92 (1.80)
Residues in Helix I (235-255)	0.82 (1.66)
Residues in Helix II(268-280)	0.56 (1.54)

[a]The number of each type of restraints used in the structure calculation is given in parenthesis. [b]F_{NOE} and F_{cdih} were calculated using force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively. [c]F_{repel} was calculated using a final value of 4.0 kcal mol^{-1} Å^{-4} with the van der Waals hard sphere radii set to 0.75 times those in the parameter set PARALLHSA supplied with X-PLOR (Brünger, 1993). [d]E_{L-J} is the Lennard-Jones van der Waals energy calculated with the CHARMM empirical energy function, and is not included in the target function for simulated annealing calculation. [e] The average rms differences from the energy-minimized average structure are given for selected residues. The value for backbone atoms (N, Cα, and C') is followed by that for all heavy atoms in parenthesis.

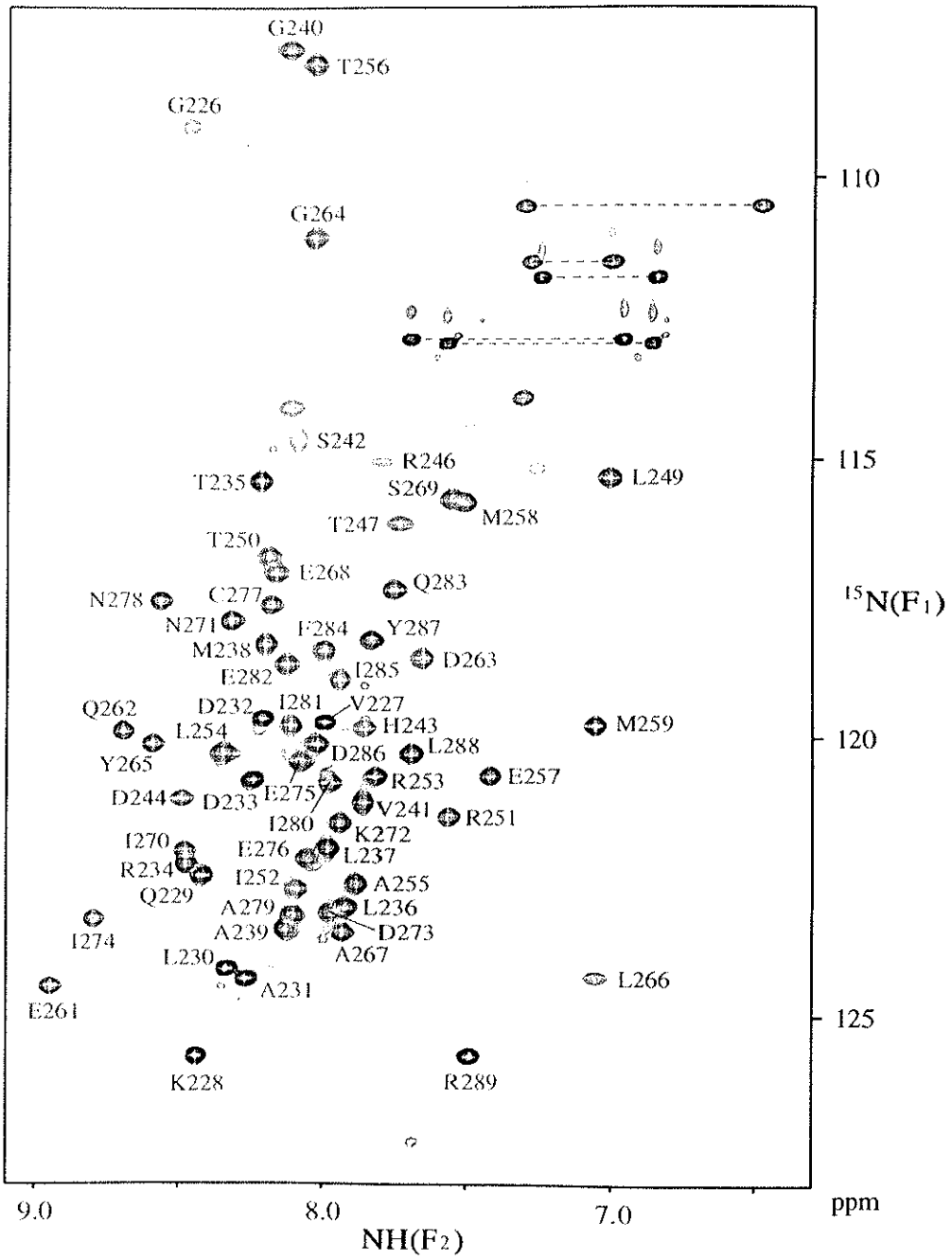

Figure III-2. ^{1}H- ^{15}N HSQC spectrum of domain A.

The spectrum is labeled with a sequence-specific assignment.
Cross peaks connected by dotted lines correspond to Gln and Asn side-chain NH_2 groups.

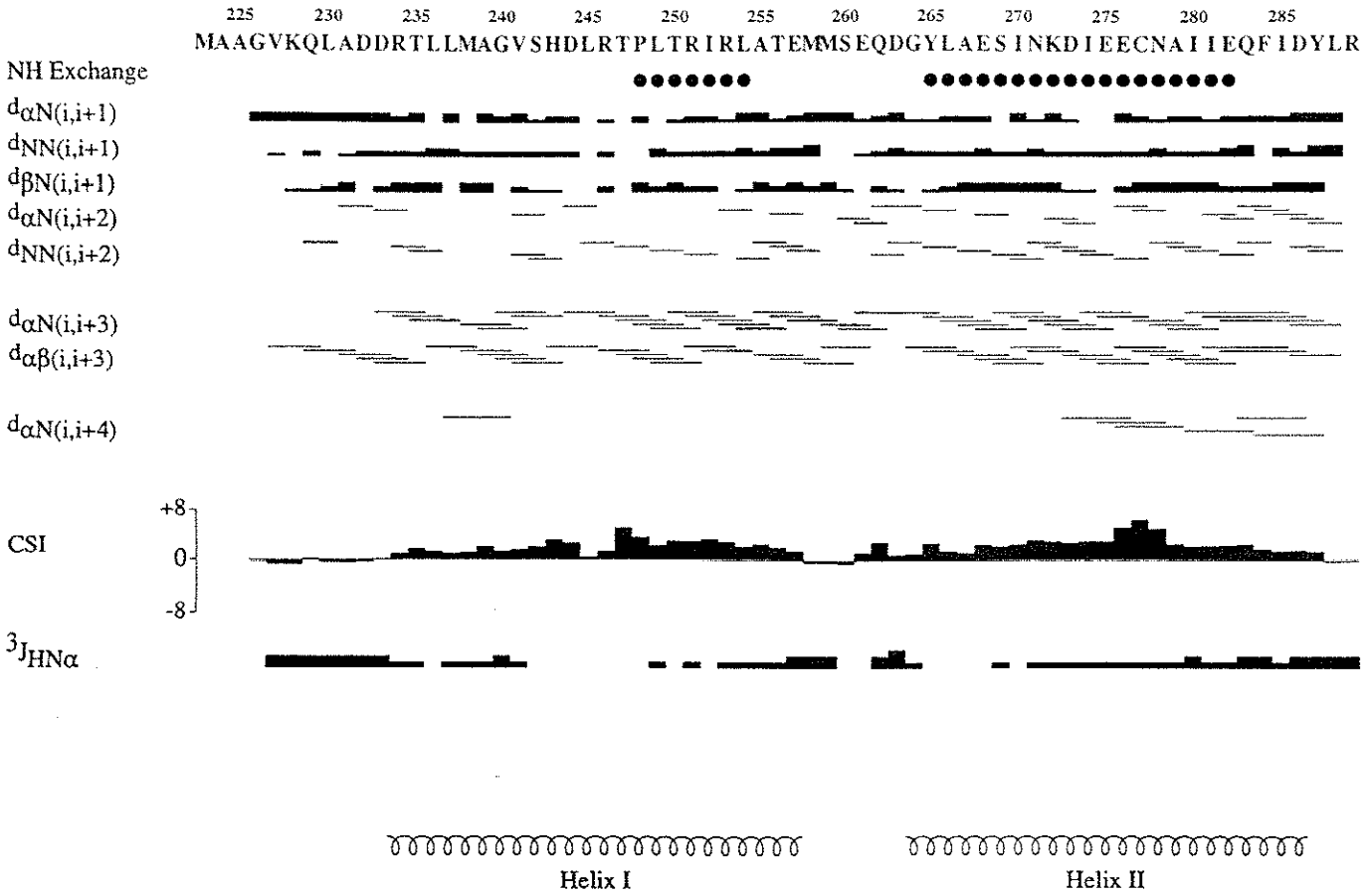

Figure III-3. Summary of sequential and medium-range NOEs involving NH, Hα, and Hβ resonances, amide hydrogen exchange data, $^3J_{HN\alpha}$ coupling, and CSI observed for domain A.

Backbone amide hydrogens that exchange slowly are indicated with filled circles. NOE connectivities [dNN(i, i+1), dαN(i, i+1), and dβN(i, i+1)] are represented as strong, medium, weak, or very weak intensity in bar graphs. $^3J_{HN\alpha}$ data with small couplings (< 5 Hz) are shown as bar graphs. At the bottom, cartoon representation of the secondary structure is shown. CSI indicated in bar graphs were calculated as Figure II-2.

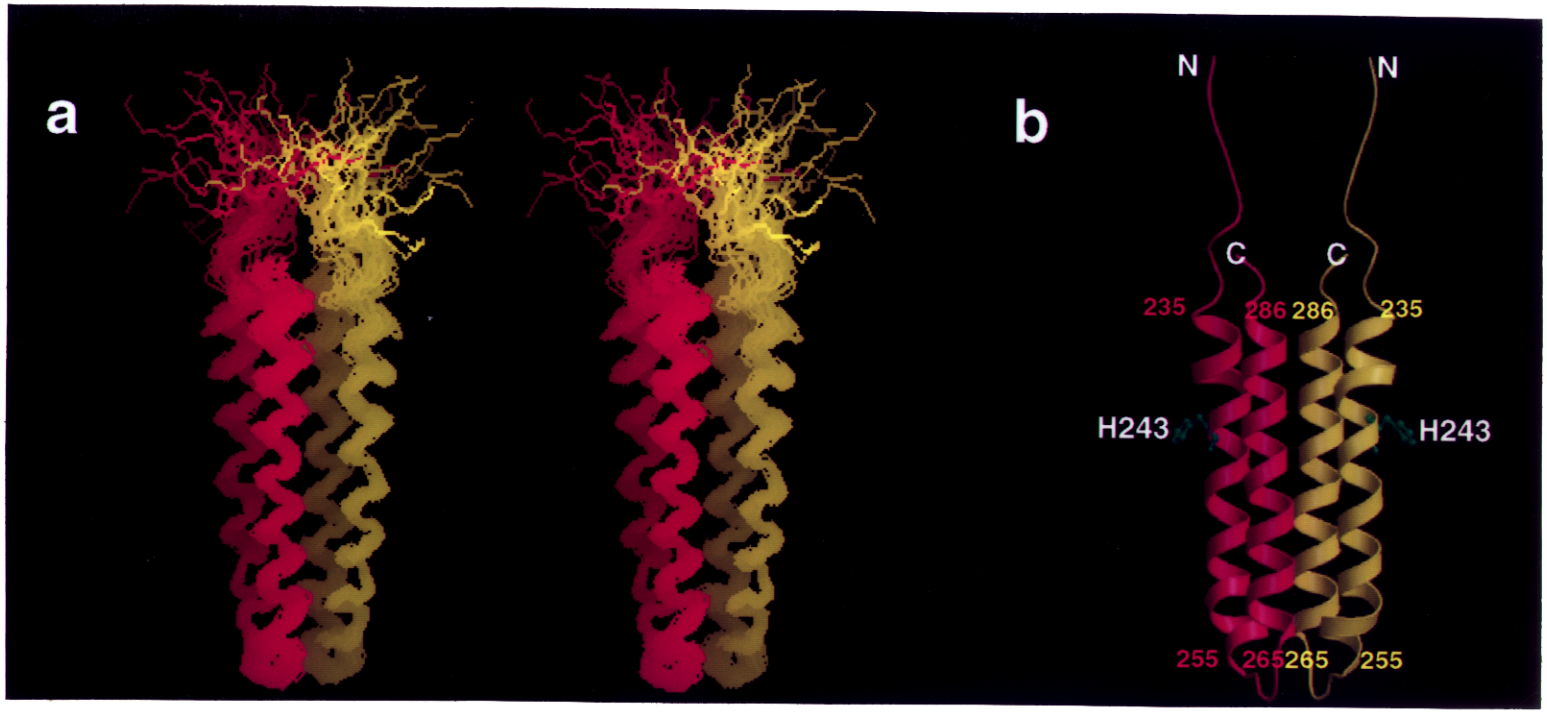

Figure III-4a. Stereoview of a best fit superposition of the backbone atoms (N, Cα, and C') of the 30 NMR-derived structures of domain A.

The main-chain atoms of the 30 structures are superimposed against the one with the lowest NOE violations using the residues that are part of helices 235-255 and 265-286. To clarify the orientation of the dimer, one (subunit A) and another (subunit B) are colored in magenta and yellow, respectively. The figure was generated using Insight II.

III-4b. Schematic ribbon drawing of the selected structure of domain A.

The conserved His-243 is labeled and shown as a ball-and-stick model. The start and end of helices are also labeled. Each subunit is colored as Figure III-4a and the N- and C- termini of the protein are indicated. This figure was generated using MOLSCRIPT and Raster3D.

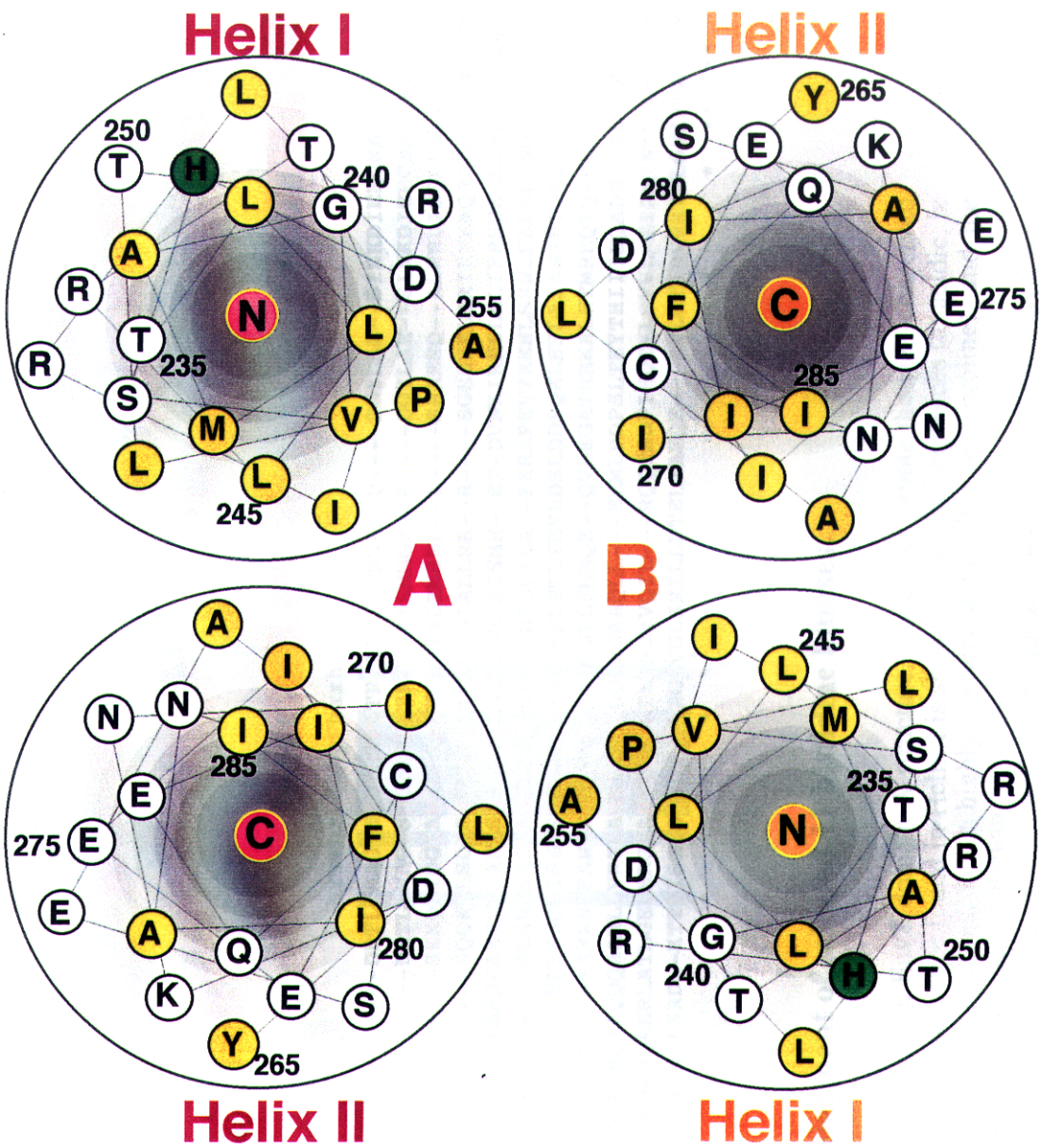

Figure III-5. Helical wheel representation of domain A.

Wheels are orientated according to the NOE connectivities. At the center of each wheels, the N- or C- terminal helices are indicated as N or C, respectively. Subunits A and B are colored as in Figure III-4a. The hydrophobic residues are shown in yellow, and the auto-transphosphorylation site His-243 in green.

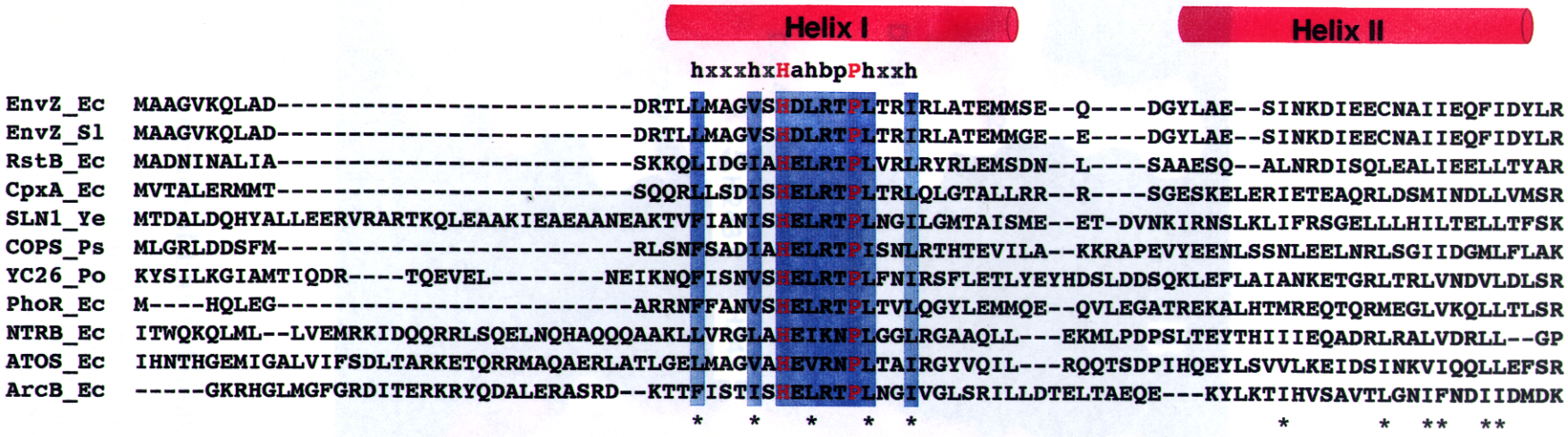

Figure III-6. Sequence alignment of selected 11 histidine kinases.

The cylinders represent α-helices found in EnvZ domain A structure. Asterisks denote the well-conserved hydrophobic residues. The consensus sequence found in Helix I is shown above the sequences (a: acidic, b: basic, h: hydrophobic, p: polar, x: any residues, H, P: highly conserved histidine and proline), and the corresponding residues in each sequence is shaded in cyan.
EnvZ_Ec (EnvZ from *E.coli*), EnvZ_Sl (EnvZ from *Salmonella typhimurium*), RstB from *E.coli*, CpxA from *E.coli*, SLN1 from *Saccharomyces cerevisiar*, COPS from *Pseudomonas syringae (pv. tomato)*, YC26 from *Porphyra purpurea* (plant), PhoR from *E.coli*, NTRB from *E.coli*, ATOS from *E.coli*, and ArcB from *E.coli*.

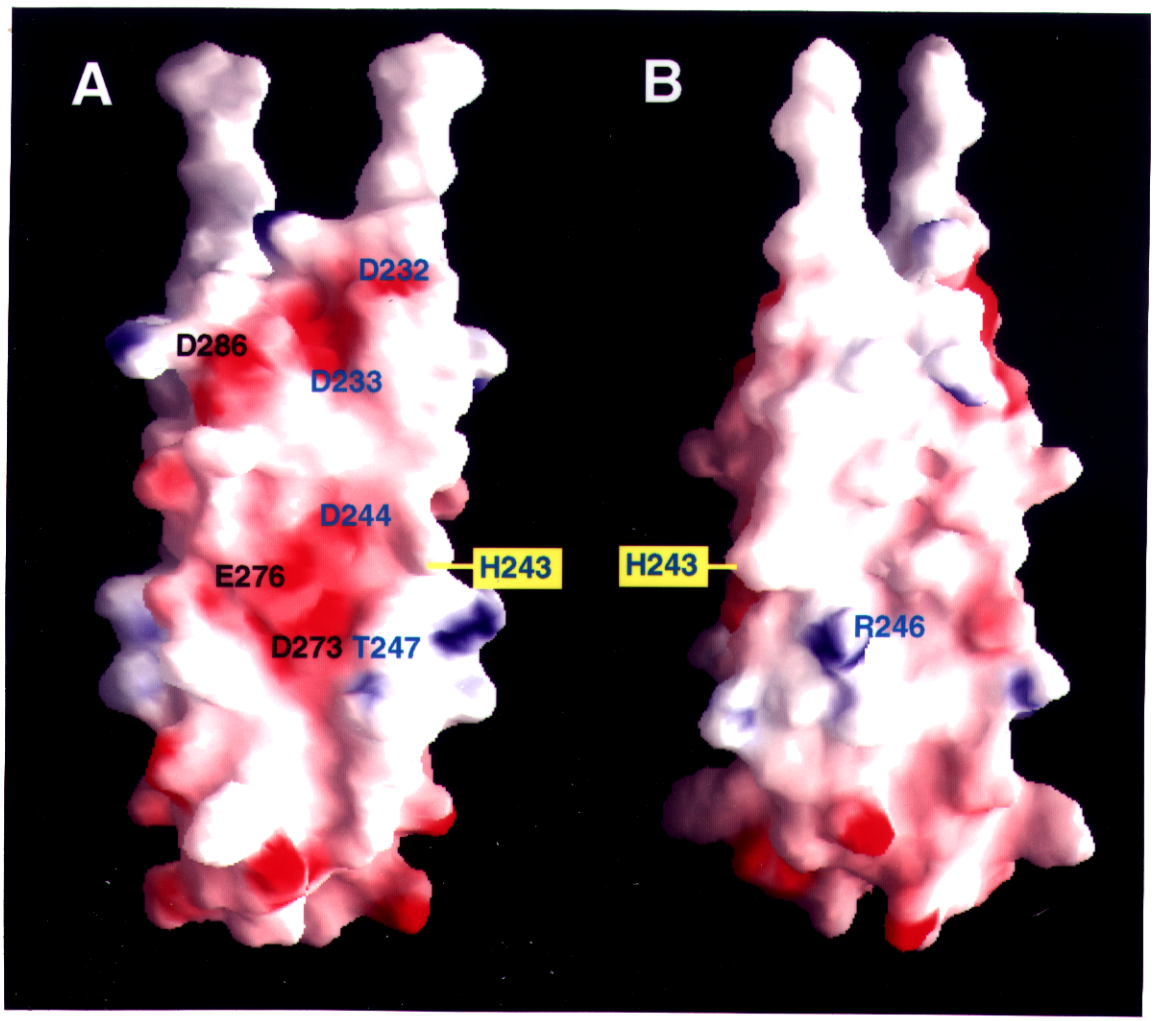

Figure III-7. Molecular surface of EnvZ domain A.

Surface electrostatic potential of EnvZ domain A was calculated and rendered using GRASP (negative potentials in red and positive in blue). Residue numbers of one subunit is shown in blue and those of another subunit in black to highlight the inter-subunit surface. Some outstanding residues of negative or positive potentials as well as the conserved His-243 and Thr-247 are labeled. The image in (B) is related to that in (A) by a 90° rotation along the vertical axis.

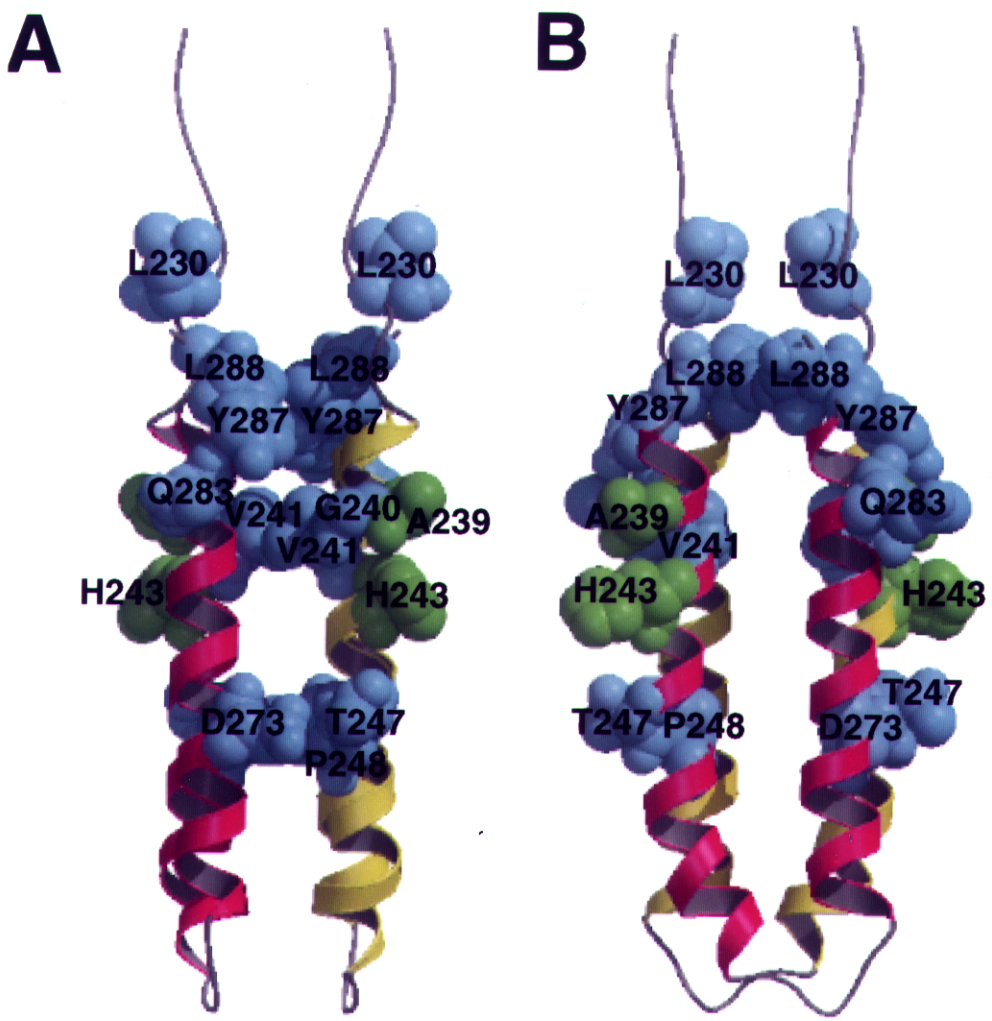

Figure III-8. Summary of mutations in EnvZ domain A.

(A, B) Mutations in domain A that affected EnvZ functions are summarized in the space filling model. Mutations resulted in kinase-/phosphatase+ are colored in green, and kinase+/phosphatase- in cyan. Subunits A and B are colored as in Figure III-4a. The inter-subunit surface is shown in A and intra-subunit surface is in B. This figure was generated using MOLSCRIPT.