Published in final edited form as:
Proteins. 2011 August ; 79(8): 2578–2582. doi:10.1002/prot.23061.

Crystal structure of the novel PaiB transcriptional regulator from *Geobacillus stearothermophilus*

E. V. Filippova[1,2], **J. S. Brunzelle**[1,3], **M. E. Cuff**[2], **H. Li**[2], **A. Joachimiak**[2], and **W. F. Anderson**[1,2,3,*]

[1]Department of Molecular Pharmacology and Biological Chemistry, Northwestern University, Feinberg School of Medicine, Chicago, Illinois 60611, USA

[2]Midwest Center for Structural Genomics, Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

[3]Life Sciences Collaborative Access Team, Advanced Photon Source, Argonne, Illinois 60439, USA

Keywords

transcriptional regulator; *pai* operon; degradative enzyme production; sporulation; *spo* genes; FMN binding

INTRODUCTION

The *pai* operon is transcriptionally regulated by PaiB and has been found to be essential for cell growth and involved in the negative control of sporulation and degradative enzyme production. The *pai* operon was initially cloned as a fragment of the *Bacillus subtilis* chromosome consisting of two genes *paiA* and *paiB*.[1] The presence of both of these genes on a multicopy vector revealed that both decrease the levels of extracellular degradative enzymes such as subtilisin, neutral protease, levansucrase, α-amylase, and alkaline phosphatases and also greatly reduce the frequency of sporulation.[2]

Based on the study by Strauch (1993), *paiB* appears to encode a regulator belonging to the same family as the AbrB, Hpr, and Sin regulators. AbrB, Hpr, Sin, and PaiB control the synthesis of degradative enzymes that occurs during the transition between vegetative growth and the onset of the stationary phase and sporulation in *B. subtilis*.[2] Known sporulation genes have been identified as targets for direct control by AbrB (*spo0E*, *spo0H*, and *spoVG*) and Sin (*spoIIA*, *spoIID*, *spoIIE*, and *spoIIG*).[3] Hpr and PaiB may also regulate *spo* genes directly as evidenced by the Spo phenotype of strains carrying multicopy plasmids that overexpress these two genes.[1] Preliminary bioinformatic data suggested that PaiB might be a split barrel flavin-binding regulator.

Herein, we report the first crystal structure of the Pai transcriptional regulator: PaiB from *Geobacillus stearothermophilus*. The structure of PaiB demonstrates that it is a homodimer and that each monomer consists of three domains: large, intermediate, and C-terminal. The large and intermediate domains fold into a structure similar to pyridoxine 5′-phosphate oxidase (PNPO) and form an analogous flavin mononucleotide (FMN)-binding cavity. The

*Correspondence to: Wayne F. Anderson, Department of Molecular Pharmacology and Biological Chemistry, Northwestern University, Feinberg School of Medicine, 303 East Chicago Avenue, Chicago, IL 60611. wf-anderson@northwestern.edu.

small C-terminal domain of PaiB possesses a helix-turn-helix (HTH) DNA-binding fold. The comparisons of sequence and structure show that PaiB is a representative of a new sequence family of transcriptional regulators. Detailed analysis of the protein structure provides insights into the potential biological function.

MATERIALS AND METHODS

Protein cloning, expression, purification and crystallization

The recombinant PaiB from *G. stearothermophilus* containing a 6His-Tag at the N-terminus was subcloned, expressed, and purified using a protocol developed at Midwest Center for Structural Genomics.[4]

The PaiB protein was crystallized by sitting-drop vapor-diffusion methods at room temperature. One volume of the reservoir solution and one volume of the protein solution were combined in the crystallization drop. The PaiB protein solution contained 10 mg/mL of protein, 500 m*M* sodium chloride, 0.5 m*M* Tris (2-carboxyethyl) phosphine hydrochloride in a 10 m*M* HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) buffer (pH 7.5). The crystals of PaiB grew in a condition containing 0.1*M* ammonium acetate, 0.1*M* Bis-Tris pH 6.0, and 17% (w/v) PEG (polyethylene glycol) 10,000.

Data collection, structure determination and refinement

A low-temperature (100 K) X-ray diffraction dataset was collected from a single crystal of PaiB on beamline 19ID at the Advance Photon Source.[5] The diffraction data were processed with the HKL-2000 program suite.[6] X-ray data-collection statistics are summarized in Table I.

The selenomethionine-labeled PaiB structure was determined using the single-wavelength anomalous dispersion method. The programs SHARP and HYSS were used to locate selenium atoms and calculate experimental phases to the limit of the data.[7,8] SHELXD, SHELXE, RESOLVE, and DM were used to perform density modification and improvement of the phases.[9–13] The initial model of the selenomethionine-labeled PaiB structure was built using ARP/wARP.[14] Gaps, turns, and side chains were fitted manually using the program O.[15] The structure was refined with REFMAC.[16] Solvent atoms were initially built using the program ARP/wARP and later added or removed by manual inspection. The refinement statistics are given in the Table I.

RESULTS AND DISCUSSION

The structure of the apo-form PaiB transcriptional regulator (NCBI GI: 134105234) from *G. stearothermophilus* has been refined to 2.5 Å resolution (Table I). The final model of PaiB contains two protein monomers (residues 11–199 for chain A, residues 10–198 for chain B) and 96 water molecules. Because they lack electron density, 10 (A) or nine (B) N-terminal and three (A) or four (B) C-terminal residues, and the loop regions 135–142 (A) and 138–142 (B) are missing. The His-Tag at the N-terminus also does not appear in the electron density map.

The two PaiB molecules in the asymmetric unit are related by a noncrystallographic twofold axis and form a dimer [Fig. 1(b)]. The ternary structure of the PaiB monomer is composed of three domains: large, intermediate, and C-terminal. The large domain comprises amino-acid residues 11–108 [α1 (14–21), β1 (24–31), β2 (34–41), β3 (43–45), β4 (50–56), β5 (70–81), β6 (94–108)] and 150–165 (β7); it adopts a six-stranded β-barrel flanked by two short α-helixes. The two smaller intermediate and C-terminal domains are composed of residues 109–149 and 166–199, respectively. They resemble a helix-loop-helix motif [α2 (112–126),

α3 (145–148)] and HTH motif [α4 (172–183), α5 (188–198)], respectively. Both molecules of the dimer have the same fold with subtle structural differences. The root mean square deviation (RMSD) between the Cα atoms of two molecules of the dimer is 0.63 Å. The major differences in the molecules are in the loops between β5–β6 and α2–α3. The dimer interface consists of residues from β1, β2, β5, β6, β7 and α2 of both molecules and buries ~1900 Å in surface area. There are 11 hydrogen bonds that contribute to the formation of the dimer interface between following residues: Tyr103-Leu35, Glu34-Arg159, Ala91-Tyr133, Glu127-Tyr80, Glu127-Ser82, Glu188-Glu188, and His39-His101. Among these residues His39, Tyr80, Ser82, and His101 are conserved in the PaiB family of transcriptional regulators [Fig. 1(a)].

In an attempt to elucidate the biological function of PaiB from its three-dimensional structure, a search for close structural homologs of PaiB was performed by DALI, VAST, and ProFunc web servers.[18–21] The PaiB from *G. stearothermophilus* is the first known structure from this transcriptional regulator protein family that is involved in sporulation and degradative enzyme production. Among the list of known proteins with similar secondary structure fold, the best match (with 1.92 Å RMSD) was found with the structure of PNPO from *Mycobacterium tuberculosis* (PDB code 2ASF).[22,23] PNPO is the enzyme that catalyzes the terminal step in the biosynthesis of pyridoxal 5′-phosphate (PLP) by the FMN oxidation of pyridoxine 5′-phosphate (PNP) forming $FMNH_2$ and H_2O_2. There are several structures in the Protein Data Bank of PNPO from different organisms, including the human enzyme, in complexes with PLP, PNP, and FMN.[23–25] Two domains [large (11–108, 150–165) and intermediate domain (109–149)] of PaiB coincide very well with the structure of PNPO [Fig. 2(a)]. This region comprises the FMN-binding cavity of PNPO and suggests that PaiB may also bind FMN. The small C-terminal domain and extended loop between strands β5–β6 [Fig. 2(a)] are the features of the PaiB transcriptional regulator that may be involved in binding of DNA.

The structure of C-terminal domain of PaiB possesses a HTH DNA-binding fold [Fig. 2(a)]. The HTH DNA-binding motif prediction server shows that the C-terminal domain contains a region (173–194 residues) with slight similarity to the HTH motif.[26] Within the PaiB family, the probable DNA-binding domain is not highly conserved, suggesting that it is designed to recognize its specific DNA. Structure analysis of PaiB reveals that the separation between the two recognition helices of the presumed HTH motifs of PaiB is ~32 Å, close to corresponding distances in other known HTH DNA-binding regulators.[27] This suggests that PaiB may also act as a DNA-binding transcriptional regulator.

The loop between β5 and β6 occupies an intermediate position between the C-terminal domain and the PNPO-like domains [Fig. 2(a)]. Additionally, this loop is involved in the extensive dimerization interface of PaiB. Interestingly, the region of β5–loop–β6 is the most conserved region among PaiB regulators [Fig. 1(a)] and could play important roles in binding both the ligand and DNA.

In the PaiB structure, a large cavity is located on the dimer interface and includes amino-acid residues from both subunits [Fig. 2(b)]. The cavity is formed by residues from α4, the extended loop between β5–β6, β6, and β7 from molecule A and by α2, α3, β1, β2 from molecule B. The dimer has two identical cavities each with a volume of 2630 $Å^3$. Based on the crystal structure, it is expected that the PaiB dimer binds two ligand molecules, one per each monomer. Superposition of the PaiB structure with the PNPO from *M. tuberculosis* indicates that the large internal cavity overlaps with the FMN-binding pocket of PNPO [Fig. 2(b)].[22,23] Thus, this cavity is likely to serve as a ligand pocket in PaiB and, like PNPO, PaiB may be capable of binding similar ligands (FMN, PLP, and PNP). There is a cluster of conserved residues among PaiB transcriptional regulators that line the inner wall of the

binding cavity: Phe24, His39, Pro41, His54, Ala56, Asn59, Tyr80, Tyr86, Ala91, Val92, Pro93, Thr94, Trp95, Asn96, Tyr97, His101, Ile150, Lys165, Leu166, and Glu168 [Figs. 1(a) and 2(c)]. This cluster coincides with the FMN and PLP binding regions in the PNPO molecule according to the superposition of PNPO and PaiB structures [Fig. 2(b)].

This study is the first to describe the crystal structure of a PaiB family transcriptional regulator. Although this study indicates a conserved ligand binding site and likely DNA-binding site, the detailed mechanism for its biological activity is still in question. The tertiary structure of PaiB reported in this article now provides a basis for an in-depth study of its structure–function relationship.

Acknowledgments

Grant sponsor: NIH; Grant number: PSI grant GM074942 (to the Midwest Center for Structural Genomics).

The X-ray data collection was performed at the Argonne National Laboratory, at the Advanced Photon Source (APS). Argonne is operated by University of Chicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. The authors would like to thank Shu Moy, George Minasov, Ludmila Shuvalova for help and discussions.

References

1. Honjo M, Nakayama A, Fukazawa K, Kawamura K, Ando K, Hori M, Furutani Y. A novel *Bacillus subtilis* gene involved in negative control of sporulation and degradative-enzyme production. J Bacteriol. 1990; 172:1783–1790. [PubMed: 2108124]
2. Strauch MA. Regulation of *Bacillus subtilis* gene expression during the transition from exponential growth to stationary phase. Prog Nucleic Acid Res Mol Biol. 1993; 46:121–153. [PubMed: 8234782]
3. Strauch MA, Hoch JA. Transition-state regulators: sentinels of *Baciiius subtiiis* post-exponential gene expression. Mol Microbiol. 1993; 7:337–342. [PubMed: 8459762]
4. Zhang RG, Skarina T, Katz JE, Beasley S, Khachatryan A, Vyas S, Arrowsmith CH, Clarke S, Edwards A, Joachimiak A, Savchenko A. Structure of *Thermotoga maritima* stationary phase survival protein SurE: a novel acid phosphatase. Structure. 2001; 9:1095–1106. [PubMed: 11709173]
5. Rosenbaum G, Alkire R, Evans G, Rotella FJ, Lazarski K, Zhang R, Ginell SL, Duke N, Naday I, Lazarz J, Molitsky MJ, Keefe L, Gonczy J, Rock L, Sanishvili R, Walsh MA, Westbrook E, Joachimiak A. The Structural Biology Center 19ID undulator beamline: facility specifications and protein crystallographic results. J Synchrotron Radiat. 2006; 13:30–45. [PubMed: 16371706]
6. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Macromol Crystal. 1997; A276:307–326.
7. De La Fortelle E, Bricogne G. Maximum-likelihood heavy-atom refinement for multiple isomorphous replacement and multiwave-length anomalous diffraction methods. Methods Enzymol. 1997; 276:472–494.
8. Grosse-Kunstleve RW, Adams PD. Substructure search procedures for macromolecular structures. Acta Crystallogr. 2003; D59:1966–1973.
9. Schneider TR, Sheldrick GM. Substructure solution with SHELXD. Acta Crystallogr. 2002; D58:1772–1779.
10. Sheldrick GM. Macromolecular phasing with SHELXE. Kristallogr. 2002; 217:644–650.
11. Terwilliger TC. Automated structure solution, density modification and model building. Acta Crystallogr. 2002; D58:1937–1940.
12. Cowtan KD, Main P. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. Acta Crystallogr. 1993; D49:148–157.
13. Cowtan KD, Zhang KYJ. Density modification for macromolecular phase improvement. Progr Biophys Mol Biol. 1999; 72:245–270.

14. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nat Struct Biol. 1999; 6:458–463. [PubMed: 10331874]
15. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved Methods for building protein models in electron-density maps and the location of errors in these models. Acta Crystallogr. 1991; A47:110–119.
16. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. 1997; D53:240–255.
17. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000; 28:235–242. [PubMed: 10592235]
18. Holm L, Sander C. 3-D lookup: fast protein structure database searches at 90% reliability. Proc Int Conf Intell Syst Mol Biol. 1995; 3:179–187. [PubMed: 7584435]
19. Gibrat JF, Madej T, Bryant SH. Surprising similarities in structure comparison. Curr Opin Struct Biol. 1996; 6:377–385. [PubMed: 8804824]
20. Madej T, Gibrat JF, Bryant SH. Threading a database of protein cores. Proteins. 1995; 23:356–3690. [PubMed: 8710828]
21. Laskowski RA, Watson JD, Thornton JM. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res. 2005; 33:89–93.
22. Biswal BK, Cherney MM, Wang M, Garen C, James MNG. Structures of *Mycobacterium tuberculosis* pyridoxine 5′-phosphate oxidase and its complexes with flavin mononucleotide and pyridoxal 5′-phosphate. Acta Crystallogr. 2005; D61:1492–1499.
23. Biswal BK, Au K, Cherney MM, Garen C, James MNG. The molecular structure of Rv2074, a probable pyridoxine 5′-phosphate oxidase from *Mycobacterium tuberculosis*, at 1. 6 Å resolution. Acta Crystallogr. 2006; F62:735–742.
24. Safo MK, Musayev FK, di Salvo ML. X-ray structure of *Escherichia coli* pyridoxine 5′-phosphate oxidase complexed with pyridoxal 5′-phosphate at 2. 0 Å resolution. J Mol Biol. 2001; 310:817–826. [PubMed: 11453690]
25. Musayev FN, Di Salvo ML, Ko TP, Schirch V, Safo MK. Structure and properties of recombinant human pyridoxine 5′-phosphate oxidase. Protein Sci. 2003; 12:1455–1463. [PubMed: 12824491]
26. Dodd IB, Egan JB. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. Nucleic Acids Res. 1990; 18:5019–5026. [PubMed: 2402433]
27. Ramos JL, Martinez-Bueno M, Molina-Henares AJ, Teran W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R. The TetR family of transcriptional repressors. Microbiol Mol Biol Rev. 2005; 69:326–356. [PubMed: 15944459]

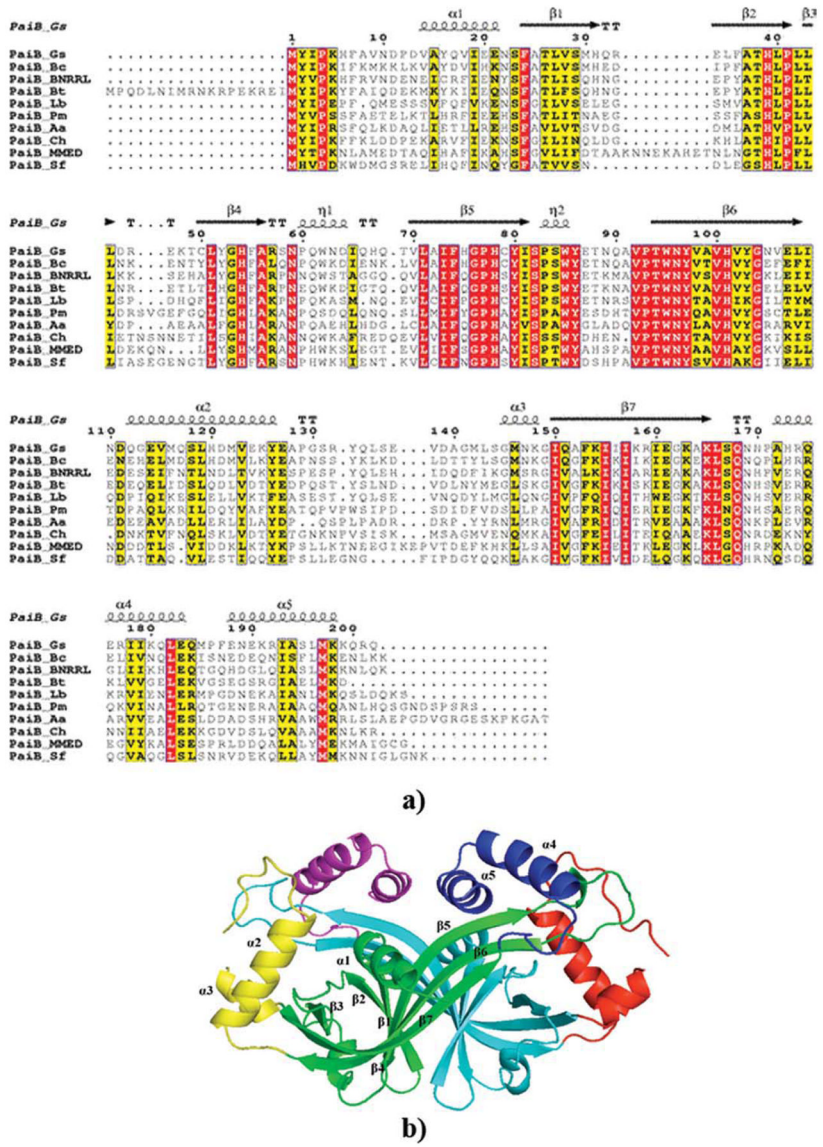

Figure 1.

(**a**) Sequence alignment of PaiB from *Geobacillus stearothermophilus* (PaiB_Gs) and other PaiB family members: PaiB from *Bacillus cereus subsp. cytotoxis* (PaiB_Bc), PaiB from *Bacillus sp. NRRLB-1491* (PaiB_BNRRL), PaiB from *Bacillus thuringiensis* (PaiB_Bt), PaiB from *Leptospira biflexa serovar Patoc* (PaiB_Lb), PaiB from *Planctomyces maris DSM8797* (PaiB_Pm), PaiB from *Alicyclobacillus acidocaldarius LAA1* (PaiB_Aa), PaiB from *Cytophaga hutchinsonii* (PaiB_Ch), PaiB from *Marinomonas sp. HED121* (PaiB_Ms), PaiB from *Shewanella frigidimarina* (PaiB_Sf). Secondary structure elements of PaiB_Gs are indicated above the sequence. Sequence homologies are highlighted by red background (identities) and yellow (similarities). (**b**) Ribbon diagram of PaiB transcriptional regulator. The domains of the dimer are colored in cyan/green (for large domains), in red/yellow (for small intermediate domains) and violet/blue (for small C-terminal domains). The secondary structure elements are presented for one molecule of the dimer.

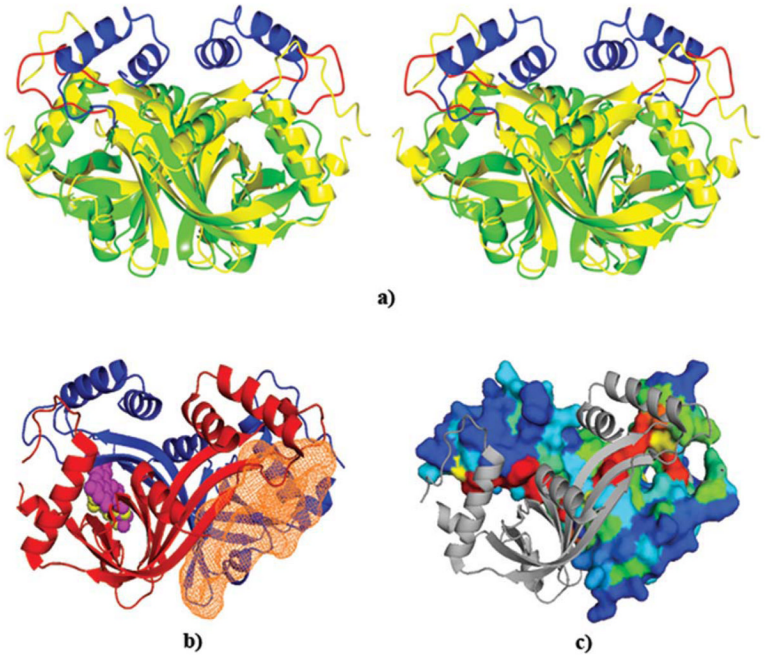

Figure 2.

(**a**) Stereo representation of superposition of the PaiB dimer (multicolor) and PNPO dimer (green). The small predicted DNA-binding domains and regions β5-loop-β6 are shown in blue and red, respectively. (**b**) The proposed FMN-binding pocket in the PaiB structure is colored in orange and shown for molecule B of the dimer. FMN (magenta) and PLP (yellow) molecules are shown as spheres models in the FMN-binding site for molecule A of the dimer. Their positions were defined by the superposition of the binding pockets of PaiB and PNPO. (**c**) PaiB dimer: The B molecule of the dimer is presented as a surface and colored according to the sequence conservation [from dark blue (highly variable) to red (invariant)], the A molecule is shown as a ribbon model (grey). Clusters of residues that coincide with the FMN and PLP binding regions [seen on 2(b)] are highly conserved and shown in red.

Table I

Crystallographic Parameters, Data Collection and Refinement Statistics

Data collection	
Resolution (Å)	38.8–2.5 (2.56–2.50)
Space group	P212121
Unit cell parameters	
a (Å)	62.6
b (Å)	77.9
c (Å)	89.4
Completeness (%)	97.7 (86.1)
Observed/unique reflections	34,879/13,772
$I/\sigma(I)$	8.7 (2.3)
R_{merge} (%)	6.7 (34.1)
Wilson B-factor (Å^2)	48.7
Refinement	
R (%)/R_{free} (%)[a]	21.3/27.6
RMS deviations from idealized geometry	
Bond lengths (Å)	0.016
Bond angles (°)	1.8
Mean *B* value (Å^2)	40.5
Number of atoms	3012
Ramachandran analysis	
Favored (%)	343 (95.6)
Allowed (%)	16 (4.4)
Generously allowed (%)	2 (0.6)
PDB ID[17]	2OL5

Data for the highest resolution shell are given in parentheses.

[a] R factor = $\Sigma(|F_{obs}|-k|F_{calc}|)/\Sigma|F_{obs}|$ and R_{free} is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.