WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Aug 23, 2023 – 12:00 PM EDT

PDB ID	:	3BXE
Title	:	Crystal structure of effector binding domain of central glycolytic gene regulator (CggR) from Bacillus subtilis in complex with dihydroxyacetone phosphate
Authors	:	Rezacova, P.; Otwinowski, Z.
Deposited on	:	2008-01-13
Resolution	:	1.80 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

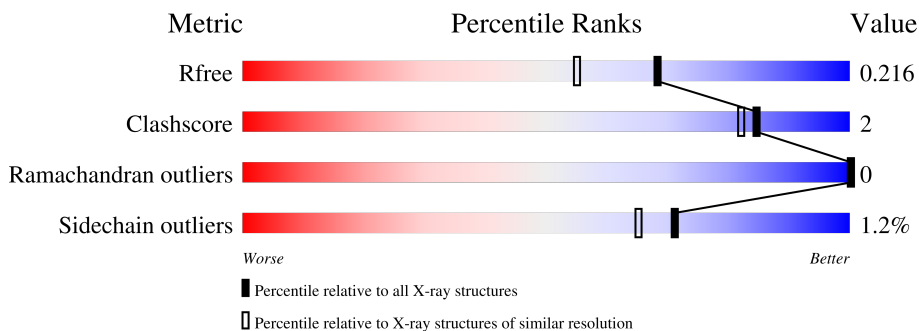|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.216
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		1.2%

Worse → *Better*

▌Percentile relative to all X-ray structures

⫿ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	255	93% · ·
1	B	255	95% · ·

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4493 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Central glycolytic gene regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	10	0
			1924	1206	332	374	12			
1	B	253	Total	C	N	O	S	0	7	0
			1946	1218	337	380	11			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	86	SER	-	expression tag	UNP O32253
A	87	ASN	-	expression tag	UNP O32253
A	88	ALA	-	expression tag	UNP O32253
B	86	SER	-	expression tag	UNP O32253
B	87	ASN	-	expression tag	UNP O32253
B	88	ALA	-	expression tag	UNP O32253

- Molecule 2 is 1,3-DIHYDROXYACETONEPHOSPHATE (three-letter code: 13P) (formula: $C_3H_7O_6P$).
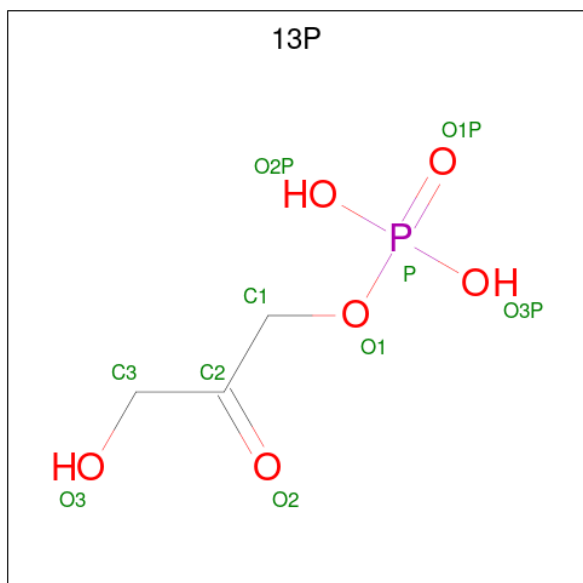
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 10	C 3	O 6	P 1	0	0
2	B	1	Total 10	C 3	O 6	P 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	294	Total 294	O 294	0	0
3	B	309	Total 309	O 309	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Central glycolytic gene regulator

Chain A:

- Molecule 1: Central glycolytic gene regulator

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.54Å　　83.82Å　　113.71Å 90.00°　　90.00°　　90.00°	Depositor
Resolution (Å)	35.90　–　1.80 35.90　–　1.80	Depositor EDS
% Data completeness (in resolution range)	99.0 (35.90-1.80) 99.0 (35.90-1.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.83 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.3.0037	Depositor
R, R_{free}	0.177　,　0.215 0.178　,　0.216	Depositor DCC
R_{free} test set	2517 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	23.3	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 34.8	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4493	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.70% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 13P

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	0.53	0/1979	0.58	0/2665
1	B	0.54	0/1991	0.62	0/2682
All	All	0.54	0/3970	0.60	0/5347

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1924	0	1979	10	0
1	B	1946	0	1999	9	0
2	A	10	0	5	0	0
2	B	10	0	5	0	0
3	A	294	0	0	4	0
3	B	309	0	0	1	0
All	All	4493	0	3988	18	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (18) close contacts within the same asymmetric unit are listed below, sorted by their clash

WORLDWIDE
wwPDB
PROTEIN DATA BANK

magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:ILE:CD1	1:A:134[A]:CYS:SG	2.31	1.19
1:B:323[B]:ASN:ND2	3:B:546:HOH:O	1.71	1.17
1:A:111:ILE:HD13	1:A:134[A]:CYS:SG	2.02	0.98
1:B:111:ILE:CD1	1:B:134[B]:CYS:SG	2.56	0.93
1:B:134[A]:CYS:SG	1:B:138[A]:ARG:NH1	2.46	0.88
1:A:111:ILE:HD12	1:A:134[A]:CYS:SG	2.24	0.78
1:B:111:ILE:HD13	1:B:134[B]:CYS:SG	2.23	0.76
1:A:111:ILE:HD11	1:A:134[A]:CYS:SG	2.31	0.69
1:A:243:GLU:HB3	3:A:474:HOH:O	1.98	0.63
1:B:134[A]:CYS:SG	1:B:138[A]:ARG:CZ	2.89	0.61
1:A:103:GLU:HG3	3:A:421:HOH:O	2.05	0.55
1:B:111:ILE:HD12	1:B:134[B]:CYS:SG	2.45	0.54
1:B:312[B]:GLU:CD	1:B:312[B]:GLU:H	2.13	0.51
1:A:221:GLU:HG3	1:B:210:LEU:HD13	1.96	0.46
1:A:221:GLU:HG2	1:A:222:PRO:HD2	2.01	0.42
1:B:261:ILE:HG23	1:B:266:ALA:HB3	2.01	0.41
1:A:182:VAL:HG23	3:A:578:HOH:O	2.19	0.41
1:A:232:LYS:HE3	3:A:555:HOH:O	2.22	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	256/255 (100%)	250 (98%)	6 (2%)	0	100	100
1	B	258/255 (101%)	253 (98%)	5 (2%)	0	100	100
All	All	514/510 (101%)	503 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	211/207 (102%)	209 (99%)	2 (1%)	78	75
1	B	212/207 (102%)	209 (99%)	3 (1%)	67	59
All	All	423/414 (102%)	418 (99%)	5 (1%)	71	65

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	109	ASP
1	A	210	LEU
1	B	204	LEU
1	B	210	LEU
1	B	262	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	13P	A	401	-	9,9,9	1.19	1 (11%)	10,12,12	0.92	0
2	13P	B	401	-	9,9,9	0.86	1 (11%)	10,12,12	0.89	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	13P	A	401	-	-	0/7/8/8	-
2	13P	B	401	-	-	0/7/8/8	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	13P	O1-C1	3.13	1.45	1.43
2	B	401	13P	O1-C1	2.19	1.44	1.43

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.