



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:36 PM GMT

PDB ID : 4M6F
Title : Dimer of the G-Segment Invertase bound to a DNA substrate
Authors : Ritacco, C.J.; Wang, J.; Steitz, T.A.
Deposited on : 2013-08-09
Resolution : 4.99 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

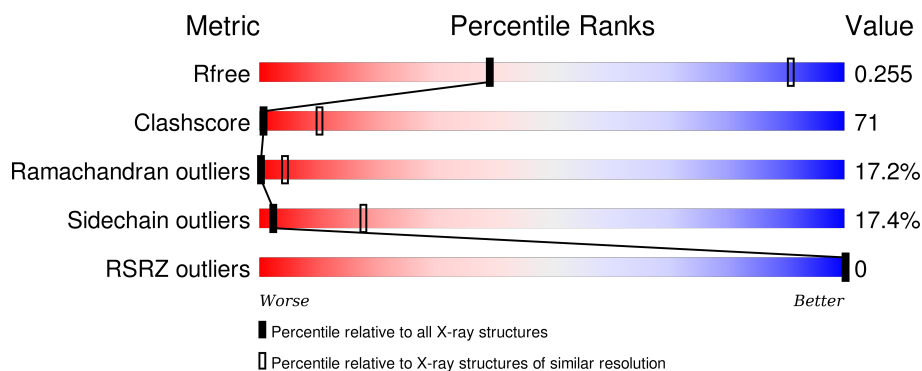
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.99 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1119 (6.22-3.60)
Clashscore	102246	1019 (6.22-3.66)
Ramachandran outliers	100387	1158 (6.22-3.60)
Sidechain outliers	100360	1136 (6.22-3.60)
RSRZ outliers	91569	1122 (6.22-3.60)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	193	<div> <div></div> <div>23%51%20%••</div> </div>
2	B	17	<div> <div></div> <div>94%6%</div> </div>
3	C	18	<div> <div></div> <div>94%6%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2196 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 DNA-invertase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	188	Total	C	N	O	S	0	0	0
			1476	921	284	265	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	114	VAL	MET	CONFLICT	UNP P03015
A	148	GLN	TRP	CONFLICT	UNP P03015

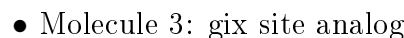
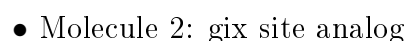
- Molecule 2 is a DNA chain called gix site analog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	17	Total	C	N	O	P	0	0	0
			352	170	58	107	17			

- Molecule 3 is a DNA chain called gix site analog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	18	Total	C	N	O	P	0	0	0
			368	177	69	104	18			

- Molecule 1: DNA-invertase



4 Data and refinement statistics

Property	Value	Source
Space group	P 64 2 2	Depositor
Cell constants a, b, c, α , β , γ	119.80 Å 119.80 Å 343.40 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 4.99 29.95 – 5.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-4.99) 91.5 (29.95-5.00)	Depositor EDS
R_{merge}	11.10	Depositor
R_{sym}	12.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.99 (at 5.06 Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.229 , 0.288 0.242 , 0.255	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	358.6	Xtriage
Anisotropy	0.230	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.18 , 424.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.35$, $\langle L^2 \rangle = 0.18$	Xtriage
Outliers	0 of 6274 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	2196	wwPDB-VP
Average B, all atoms (Å ²)	344.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.23	0/1495	0.57	0/2006
2	B	0.51	1/393 (0.3%)	0.69	0/604
3	C	0.48	1/413 (0.2%)	0.67	0/632
All	All	0.35	2/2301 (0.1%)	0.61	0/3242

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	DA	OP3-P	-7.27	1.52	1.61
3	C	14	DA	OP3-P	-7.22	1.52	1.61

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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	1476	0	1536	244	0
2	B	352	0	197	60	1
3	C	368	0	204	55	0
All	All	2196	0	1937	291	1

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 71.

The worst 5 of 291 close contacts within the same asymmetric unit are listed below, sorted by

their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:141:PRO:HG3	2:B:7:DT:H4'	1.40	1.01
1:A:182:LYS:H	3:C:19:DT:H5"	1.25	0.98
1:A:131:ARG:HE	1:A:132:ASN:H	1.07	0.94
1:A:42:ARG:NE	1:A:46:LYS:HA	1.87	0.90
1:A:170:VAL:HG13	1:A:171:ALA:H	1.35	0.89

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:17:DT:O2	2:B:17:DT:O2[4_775]	2.19	0.01

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	186/193 (96%)	110 (59%)	44 (24%)	32 (17%)	0 4

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	19	ARG
1	A	64	ASP
1	A	70	MET
1	A	88	ARG
1	A	131	ARG

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	155/161 (96%)	128 (83%)	27 (17%)	2 18

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	88	ARG
1	A	97	SER
1	A	178	LYS
1	A	92	ASP
1	A	49	LEU

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	188/193 (97%)	-1.00	0 100 100	110, 341, 487, 558	0
2	B	17/17 (100%)	-0.82	0 100 100	209, 336, 431, 474	0
3	C	18/18 (100%)	-0.97	0 100 100	318, 361, 447, 490	0
All	All	223/228 (97%)	-0.98	0 100 100	110, 344, 482, 558	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.