WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Nov 3, 2023 – 06:41 AM EDT

PDB ID	:	3W6V
Title	:	Crystal structure of the DNA-binding domain of AdpA, the global transcriptional factor, in complex with a target DNA
Authors	:	Yao, M.D.; Ohtsuka, J.; Nagata, K.; Miyazono, K.; Ohnishi, Y.; Tanokura, M.
Deposited on	:	2013-02-22
Resolution	:	2.95 Å(reported)

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

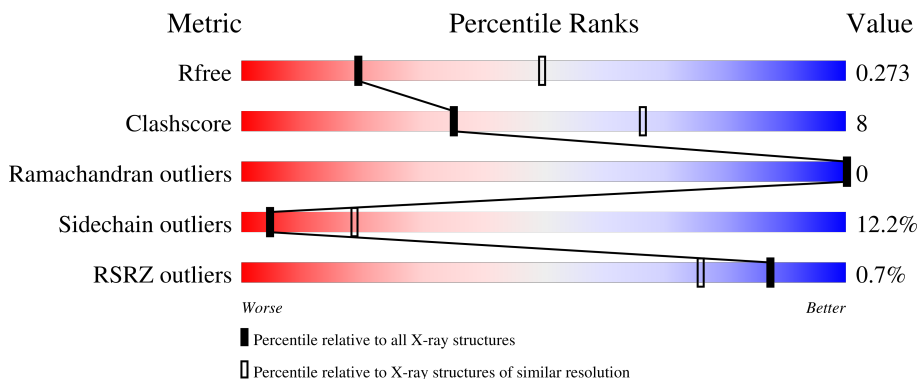|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.95 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		0.273
Clashscore		8
Ramachandran outliers		0
Sidechain outliers		12.2%
RSRZ outliers		0.7%

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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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% 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	147	57% 14% • 24%
2	B	16	88% 12%
3	C	16	50% 50%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 1555 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 AdpA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	111	896	558	177	159	2	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	194	MET	-	expression tag	UNP Q9S166
A	195	GLY	-	expression tag	UNP Q9S166
A	196	SER	-	expression tag	UNP Q9S166
A	197	SER	-	expression tag	UNP Q9S166
A	198	HIS	-	expression tag	UNP Q9S166
A	199	HIS	-	expression tag	UNP Q9S166
A	200	HIS	-	expression tag	UNP Q9S166
A	201	HIS	-	expression tag	UNP Q9S166
A	202	HIS	-	expression tag	UNP Q9S166
A	203	HIS	-	expression tag	UNP Q9S166
A	204	SER	-	expression tag	UNP Q9S166
A	205	SER	-	expression tag	UNP Q9S166
A	206	GLY	-	expression tag	UNP Q9S166
A	207	LEU	-	expression tag	UNP Q9S166
A	208	VAL	-	expression tag	UNP Q9S166
A	209	PRO	-	expression tag	UNP Q9S166
A	210	ARG	-	expression tag	UNP Q9S166
A	211	GLY	-	expression tag	UNP Q9S166
A	212	SER	-	expression tag	UNP Q9S166
A	213	HIS	-	expression tag	UNP Q9S166
A	214	MET	-	expression tag	UNP Q9S166

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*TP*GP*TP*GP*AP*AP*CP*CP*CP* GP*CP*CP*AP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	16	Total 320	C 153	N 60	O 92	P 15	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(*AP*GP*GP*TP*TP*GP*GP*CP*GP*GP* GP*TP*TP*CP*AP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	16	Total 330	C 157	N 62	O 96	P 15	0	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	8	Total 8	O 8	0	0
4	C	1	Total 1	O 1	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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: AdpA

Chain A:
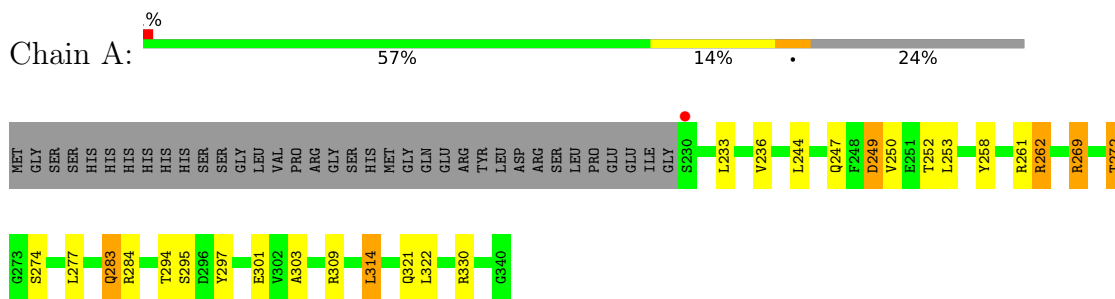
• Molecule 2: DNA (5'-D(*CP*TP*GP*TP*GP*AP*AP*CP*CP*CP*GP*CP*CP*AP*AP*C)-3')

Chain B:
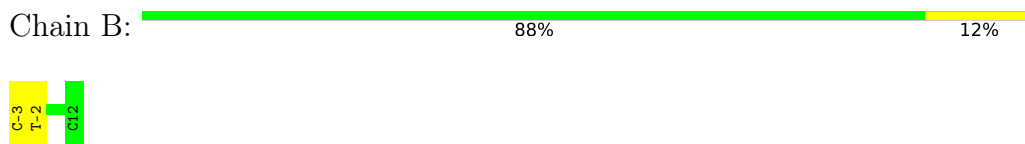
• Molecule 3: DNA (5'-D(*AP*GP*GP*TP*TP*GP*GP*CP*GP*GP*GP*TP*TP*CP*AP*C)-3')

Chain C:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	76.51Å 100.64Å 100.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.50 – 2.95 19.50 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.50-2.95) 99.5 (19.50-2.95)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	68.44 (at 2.93Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.236 , 0.278 0.229 , 0.273	Depositor DCC
R_{free} test set	399 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	69.3	Xtriage
Anisotropy	0.271	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 14.8	EDS
L-test for twinning[2]	$< \lvert L \rvert > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	1555	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< \lvert L \rvert >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (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	0.51	0/916	0.41	0/1237
2	B	0.90	0/358	1.12	0/549
3	C	0.88	0/370	1.30	1/571 (0.2%)
All	All	0.70	0/1644	0.89	1/2357 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	-1	DG	O4'-C4'-C3'	-8.51	100.89	106.00

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	896	0	882	16	0
2	B	320	0	180	2	0
3	C	330	0	182	5	0
4	A	8	0	0	0	0
4	C	1	0	0	0	0
All	All	1555	0	1244	23	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 8.

All (23) 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:272:THR:HG22	1:A:274:SER:H	1.28	0.92
3:C:-3:DA:H4'	3:C:-2:DG:OP1	1.94	0.66
1:A:303:ALA:HB2	1:A:314:LEU:HD12	1.82	0.60
1:A:272:THR:HG22	1:A:274:SER:N	2.08	0.60
3:C:-3:DA:C4'	3:C:-2:DG:OP1	2.55	0.55
1:A:277:LEU:O	1:A:277:LEU:HG	2.05	0.54
1:A:244:LEU:O	1:A:284:ARG:NH1	2.41	0.53
3:C:3:DG:H2"	3:C:4:DC:O5'	2.07	0.53
1:A:262:ARG:O	1:A:262:ARG:HD2	2.08	0.53
1:A:244:LEU:HD11	1:A:283:GLN:HB3	1.91	0.52
1:A:269:ARG:NH2	1:A:274:SER:HA	2.25	0.51
1:A:244:LEU:HD22	1:A:284:ARG:HG2	1.92	0.51
1:A:295:SER:O	1:A:330:ARG:NH2	2.44	0.50
1:A:249:ASP:HB3	1:A:252:THR:H	1.81	0.46
1:A:297:TYR:HB3	1:A:301:GLU:HB3	1.98	0.45
2:B:-2:DT:OP2	2:B:-2:DT:H6	2.00	0.45
1:A:250:VAL:HA	1:A:253:LEU:HD12	1.98	0.44
1:A:321:GLN:HG3	1:A:322:LEU:HD13	1.99	0.44
3:C:8:DT:C2'	3:C:9:DT:H71	2.50	0.42
1:A:233:LEU:O	1:A:236:VAL:HG12	2.19	0.42
1:A:249:ASP:O	1:A:252:THR:HB	2.20	0.42
3:C:9:DT:H2"	3:C:10:DC:O5'	2.20	0.41
2:B:-3:DC:C2'	2:B:-2:DT:H5'	2.51	0.41

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	109/147 (74%)	101 (93%)	8 (7%)	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	90/121 (74%)	79 (88%)	11 (12%)	5	19

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	247	GLN
1	A	249	ASP
1	A	258	TYR
1	A	261	ARG
1	A	262	ARG
1	A	269	ARG
1	A	272	THR
1	A	283	GLN
1	A	294	THR
1	A	309	ARG
1	A	314	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	339	GLN

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 ⓘ

There are no ligands in this entry.

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.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	111/147 (75%)	-0.11	1 (0%)	84	71	45, 61, 72, 77	0
2	B	16/16 (100%)	-0.33	0	100	100	52, 79, 92, 92	0
3	C	16/16 (100%)	-0.30	0	100	100	60, 70, 85, 87	0
All	All	143/179 (79%)	-0.16	1 (0%)	87	76	45, 63, 84, 92	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	230	SER	3.0

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

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

6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5   Other polymers ⓘ

There are no such residues in this entry.