



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:47 PM GMT

PDB ID : 4AIJ
Title : Crystal structure of RovA from Yersinia in complex with a rovA promoter fragment
Authors : Quade, N.; Mendonca, C.; Herbst, K.; Heroven, A.K.; Heinz, D.W.; Dersch, P.
Deposited on : 2012-02-10
Resolution : 2.05 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

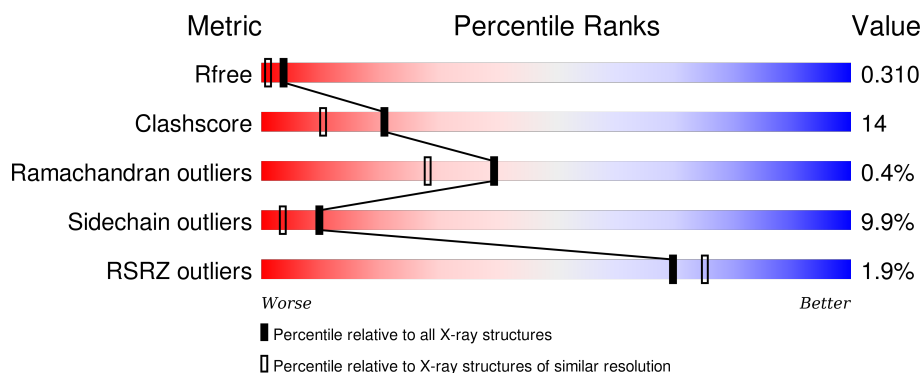
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.05 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	1192 (2.04-2.04)
Clashscore	102246	1269 (2.04-2.04)
Ramachandran outliers	100387	1258 (2.04-2.04)
Sidechain outliers	100360	1258 (2.04-2.04)
RSRZ outliers	91569	1194 (2.04-2.04)

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	151	<div> <div>2%</div> <div>60% 28% 8%</div> </div>
1	B	151	<div> <div>2%</div> <div>60% 31% 6%</div> </div>
2	C	21	<div> <div>38% 43% 19%</div> </div>
3	D	21	<div> <div>29% 38% 33%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3158 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 TRANSCRIPTIONAL REGULATOR SLYA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	A	139	Total	C	N	O	0	0	0
			1101	690	202	209			
1	B	142	Total	C	N	O	0	0	0
			1126	705	206	215			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	144	LEU	-	EXPRESSION TAG	UNP B1JJ73
A	145	GLU	-	EXPRESSION TAG	UNP B1JJ73
A	146	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	147	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	148	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	149	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	150	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	151	HIS	-	EXPRESSION TAG	UNP B1JJ73
A	81	SER	CYS	ENGINEERED MUTATION	UNP B1JJ73
A	108	SER	CYS	ENGINEERED MUTATION	UNP B1JJ73
B	144	LEU	-	EXPRESSION TAG	UNP B1JJ73
B	145	GLU	-	EXPRESSION TAG	UNP B1JJ73
B	146	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	147	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	148	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	149	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	150	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	151	HIS	-	EXPRESSION TAG	UNP B1JJ73
B	81	SER	CYS	ENGINEERED MUTATION	UNP B1JJ73
B	108	SER	CYS	ENGINEERED MUTATION	UNP B1JJ73

- Molecule 2 is a DNA chain called ROVA PROMOTER FRAGMENT, 5'-D(*AP*AP*TP*T
P*AP*TP*AP*TP *TP*AP*TP*TP*TP*GP*AP*AP*TP*TP*AP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	21	Total 428	C 210	N 72	O 126	P 20	0	0	0

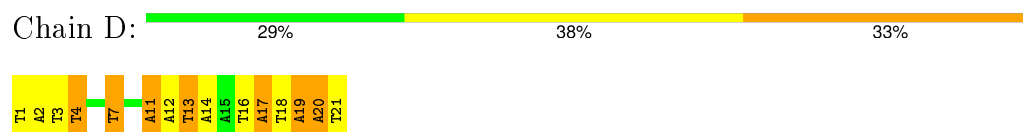
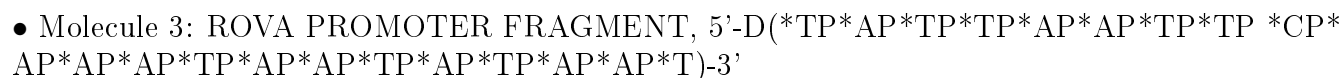
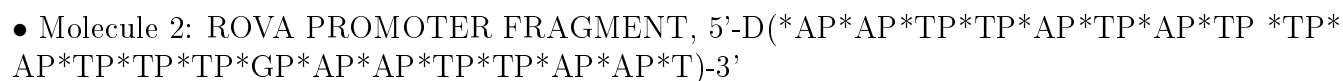
- Molecule 3 is a DNA chain called ROVA PROMOTER FRAGMENT, 5'-D(*TP*AP*TP*T
P*AP*AP*TP*TP *CP*AP*AP*AP*TP*AP*AP*TP*AP*TP*AP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	21	Total 427	C 209	N 76	O 122	P 20	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	40	Total 40	O 40	0	0
4	B	31	Total 31	O 31	0	0
4	C	1	Total 1	O 1	0	0
4	D	4	Total 4	O 4	0	0

• Molecule 1: TRANSCRIPTIONAL REGULATOR SLYA



4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, α , β , γ	88.78 Å 88.78 Å 67.73 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	76.90 – 2.05 37.13 – 2.05	Depositor EDS
% Data completeness (in resolution range)	98.6 (76.90-2.05) 99.2 (37.13-2.05)	Depositor EDS
R_{merge}	0.01	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.41 (at 2.05 Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.203 , 0.251 0.264 , 0.310	Depositor DCC
R_{free} test set	1911 reflections (5.15%)	DCC
Wilson B-factor (Å ²)	39.9	Xtriage
Anisotropy	0.671	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 24.0	EDS
Estimated twinning fraction	0.502 for H,K,L 0.498 for K,H,-L 0.006 for -h,-k,l 0.177 for h,-h-k,-l 0.012 for -k,-h,l	Xtriage
Reported twinning fraction	0.502 for H,K,L 0.498 for K,H,-L	Depositor
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 37171 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	3158	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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	0.65	0/1113	0.83	1/1505 (0.1%)
1	B	0.69	1/1138 (0.1%)	0.82	0/1538
2	C	0.78	0/479	1.55	5/738 (0.7%)
3	D	0.71	0/479	1.60	11/737 (1.5%)
All	All	0.70	1/3209 (0.0%)	1.13	17/4518 (0.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	10	ALA	CA-CB	5.25	1.63	1.52

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	D	11	DA	P-O3'-C3'	9.17	130.71	119.70
2	C	14	DG	P-O3'-C3'	7.88	129.15	119.70
3	D	17	DA	P-O3'-C3'	7.62	128.85	119.70
3	D	18	DT	O4'-C1'-N1	7.06	112.94	108.00
2	C	12	DT	P-O3'-C3'	6.93	128.01	119.70

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	1101	0	1161	45	0
1	B	1126	0	1187	33	0
2	C	428	0	244	9	0
3	D	427	0	242	8	0
4	A	40	0	0	2	0
4	B	31	0	0	0	0
4	C	1	0	0	0	0
4	D	4	0	0	1	0
All	All	3158	0	2834	83	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 14.

The worst 5 of 83 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:B:121:ASP:O	1:B:125:VAL:HG12	1.61	1.00
1:A:129:LEU:HD11	1:B:122:GLU:HB3	1.53	0.88
2:C:10:DA:H2''	2:C:11:DT:O5'	1.80	0.82
1:A:131:ASP:O	1:A:135:LYS:HG2	1.80	0.81
1:A:65:ARG:HD3	1:A:65:ARG:O	1.87	0.74

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	137/151 (91%)	131 (96%)	6 (4%)	0	100	100
1	B	140/151 (93%)	135 (96%)	4 (3%)	1 (1%)	26	15
All	All	277/302 (92%)	266 (96%)	10 (4%)	1 (0%)	39	28

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	46	GLU

5.3.2 Protein sidechains [i](#)

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	125/137 (91%)	114 (91%)	11 (9%)	12	5
1	B	128/137 (93%)	114 (89%)	14 (11%)	8	2
All	All	253/274 (92%)	228 (90%)	25 (10%)	10	3

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2	GLU
1	B	34	TRP
1	B	141	GLN
1	B	32	THR
1	B	50	ILE

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

Mol	Chain	Res	Type
1	A	31	GLN

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 [i](#)

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	139/151 (92%)	0.33	3 (2%) 65 71	35, 46, 105, 135	1 (0%)
1	B	142/151 (94%)	0.31	3 (2%) 67 72	34, 49, 89, 112	1 (0%)
2	C	21/21 (100%)	-0.39	0 100 100	82, 95, 112, 119	0
3	D	21/21 (100%)	-0.46	0 100 100	86, 100, 112, 123	0
All	All	323/344 (93%)	0.23	6 (1%) 70 75	34, 51, 108, 135	2 (0%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	82	ALA	6.8
1	B	82	ALA	5.2
1	B	87	ALA	3.2
1	B	86	ARG	2.7
1	A	83	ASN	2.5

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.