W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2017 – 02:17 am GMT

PDB ID	:	1VJ7
Title	:	Crystal structure of the bifunctional catalytic fragment of RelSeq, the RelA/SpoT homolog from Streptococcus equisimilis.
Authors	:	Hogg, T.; Mechold, U.; Malke, H.; Cashel, M.; Hilgenfeld, R.
Deposited on	:	2004-02-03
Resolution	:	2.10 Å(reported)

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

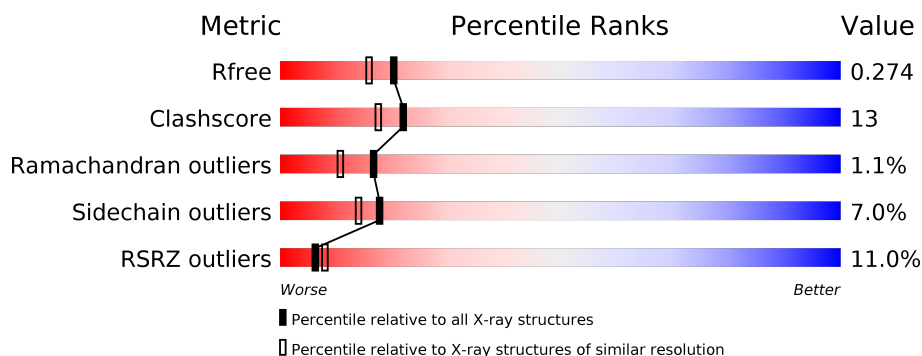|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.10 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		0.274
Clashscore		13
Ramachandran outliers		1.1%
Sidechain outliers		7.0%
RSRZ outliers		11.0%

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$_{free}$	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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 <=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	393	7% — 59% 20% · · 17%
1	B	393	11% — 54% 21% · 21%

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 5424 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 Bifunctional RELA/SPOT.

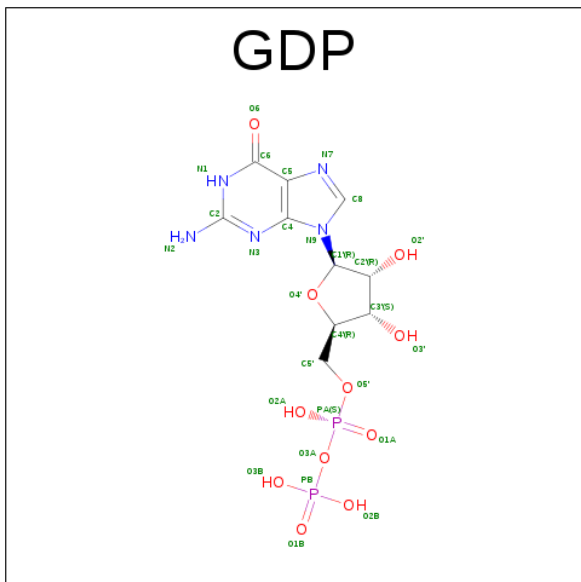| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	326	Total 2653	C 1697	N 461	O 480	S 15	0	0	0
1	B	310	Total 2486	C 1592	N 428	O 454	S 12	6	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	386	LEU	-	EXPRESSION TAG	UNP Q54089
A	387	GLU	-	EXPRESSION TAG	UNP Q54089
A	388	HIS	-	EXPRESSION TAG	UNP Q54089
A	389	HIS	-	EXPRESSION TAG	UNP Q54089
A	390	HIS	-	EXPRESSION TAG	UNP Q54089
A	391	HIS	-	EXPRESSION TAG	UNP Q54089
A	392	HIS	-	EXPRESSION TAG	UNP Q54089
A	393	HIS	-	EXPRESSION TAG	UNP Q54089
B	386	LEU	-	EXPRESSION TAG	UNP Q54089
B	387	GLU	-	EXPRESSION TAG	UNP Q54089
B	388	HIS	-	EXPRESSION TAG	UNP Q54089
B	389	HIS	-	EXPRESSION TAG	UNP Q54089
B	390	HIS	-	EXPRESSION TAG	UNP Q54089
B	391	HIS	-	EXPRESSION TAG	UNP Q54089
B	392	HIS	-	EXPRESSION TAG	UNP Q54089
B	393	HIS	-	EXPRESSION TAG	UNP Q54089

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total 1	Mn 1	0	0
2	A	1	Total 1	Mn 1	0	0

- Molecule 3 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			28	10	5	11	2		
3	B	1	Total	C	N	O	P	0	0
			28	10	5	11	2		

- Molecule 4 is GUANOSINE 5'-DIPHOSPHATE 2':3'-CYCLIC MONOPHOSPHATE (three-letter code: GPX) (formula: $C_{10}H_{14}N_5O_{13}P_3$).
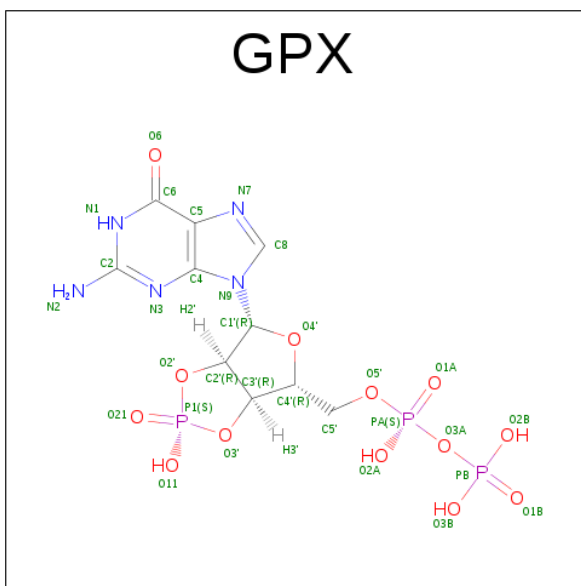
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

• Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	136	Total	O	0	0
			136	136		
5	B	60	Total	O	0	0
			60	60		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: Bifunctional RELA/SPOT

- Molecule 1: Bifunctional RELA/SPOT

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	173.49Å 45.45Å 126.47Å 90.00° 109.83° 90.00°	Depositor
Resolution (Å)	22.70 – 2.10 22.73 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.0 (22.70-2.10) 96.0 (22.73-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.15 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5	Depositor
R, R_{free}	0.236 , 0.272 0.246 , 0.274	Depositor DCC
R_{free} test set	2701 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	44.7	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 67.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5424	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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: GDP, MN, GPX

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	0.69	0/2704	1.17	25/3642 (0.7%)
1	B	0.54	0/2533	1.05	16/3417 (0.5%)
All	All	0.62	0/5237	1.12	41/7059 (0.6%)

There are no bond length outliers.

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	327	ARG	NE-CZ-NH2	-11.82	114.39	120.30
1	A	67	ASP	N-CA-CB	-8.36	95.56	110.60
1	A	214	ARG	NE-CZ-NH2	-8.24	116.18	120.30
1	B	100	ASP	CB-CG-OD2	7.66	125.19	118.30
1	A	327	ARG	NE-CZ-NH1	7.63	124.11	120.30

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	2653	0	2677	71	2
1	B	2486	0	2482	68	1
2	A	1	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	28	0	12	2	0
3	B	28	0	12	0	0
4	B	31	0	10	0	0
5	A	136	0	0	4	0
5	B	60	0	0	0	0
All	All	5424	0	5193	138	2

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 13.

The worst 5 of 138 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:20:MET:CE	1:A:68:ALA:HB2	1.80	1.11
1:A:292:MET:SD	1:A:316:TYR:CE1	2.44	1.10
1:B:20:MET:HE2	1:B:24:ASP:HB3	1.37	1.04
1:A:273:GLU:OE2	1:A:329:LYS:HE2	1.65	0.95
1:B:79:VAL:O	1:B:83:THR:HB	1.68	0.93

All (2) 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)
1:A:226:LYS:NZ	1:A:370:GLU:OE1[1_565]	2.19	0.01
1:A:93:ASP:OD2	1:B:41:TYR:OH[3_445]	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	318/393 (81%)	303 (95%)	14 (4%)	1 (0%)	44 44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	300/393 (76%)	281 (94%)	13 (4%)	6 (2%)	9	4
All	All	618/786 (79%)	584 (94%)	27 (4%)	7 (1%)	17	11

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	67	ASP
1	B	318	PRO
1	B	365	ILE
1	B	366	LYS
1	A	67	ASP

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	282/339 (83%)	263 (93%)	19 (7%)	19	15
1	B	262/339 (77%)	243 (93%)	19 (7%)	16	13
All	All	544/678 (80%)	506 (93%)	38 (7%)	18	14

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	281	MET
1	B	23	THR
1	B	264	ASP
1	B	18	LYS
1	B	41	TYR

Some 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

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	$\#|Z| > 2$
3	GDP	A	998	-	25,30,30	1.16	2 (8%)	26,47,47	2.72	10 (38%)
3	GDP	B	1998	-	25,30,30	1.19	2 (8%)	26,47,47	2.36	8 (30%)
4	GPX	B	1999	-	25,34,34	2.20	4 (16%)	24,55,55	1.36	3 (12%)

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
3	GDP	A	998	-	-	0/12/32/32	0/3/3/3
3	GDP	B	1998	-	-	0/12/32/32	0/3/3/3
4	GPX	B	1999	-	-	1/15/42/42	0/4/4/4

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1999	GPX	C5-C6	-6.82	1.40	1.53
4	B	1999	GPX	C4-N9	-6.56	1.39	1.47
4	B	1999	GPX	C8-N9	-3.44	1.36	1.46
4	B	1999	GPX	C6-N1	2.69	1.37	1.33
3	A	998	GDP	C6-C5	3.13	1.47	1.41

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	998	GDP	C6-C5-C4	-7.89	113.00	120.84
3	B	1998	GDP	C5-C6-N1	-4.96	116.42	123.48
3	A	998	GDP	N3-C2-N1	-4.27	121.22	127.46
3	B	1998	GDP	C6-C5-C4	-4.27	116.59	120.84
4	B	1999	GPX	O6-C6-N1	-4.03	117.33	122.70

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	1999	GPX	PA-O5'-C5'-C4'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	998	GDP	2	0

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(Å^2)	Q<0.9
1	A	326/393 (82%)	0.40	27 (8%)	12	15	22, 33, 47, 63	0
1	B	310/393 (78%)	0.60	43 (13%)	3	4	23, 33, 47, 58	1 (0%)
All	All	636/786 (80%)	0.50	70 (11%)	6	8	22, 33, 47, 63	1 (0%)

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	368	LEU	9.7
1	A	124	HIS	6.0
1	B	112	GLU	5.2
1	A	140	VAL	4.8
1	B	133	LYS	4.6

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 carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors($Å^2$)	Q<0.9
3	GDP	A	998	28/28	0.87	0.17	0.70	29,47,68,70	0
3	GDP	B	1998	28/28	0.86	0.15	0.05	34,48,71,72	0
2	MN	A	997	1/1	0.99	0.15	-0.13	30,30,30,30	0
4	GPX	B	1999	31/31	0.91	0.12	-0.50	27,41,54,58	0
2	MN	B	1997	1/1	1.00	0.10	-0.93	32,32,32,32	0

6.5 Other polymers (i)

There are no such residues in this entry.