WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 07:50 AM EDT

PDB ID	:	2PJU
Title	:	Crystal structure of propionate catabolism operon regulatory protein prpR
Authors	:	Ramagopal, U.A.; Toro, R.; Gilmore, M.; Wu, B.; Koss, J.; Groshong, C.; Gheyi, T.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2007-04-16
Resolution	:	2.10 Å(reported)

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

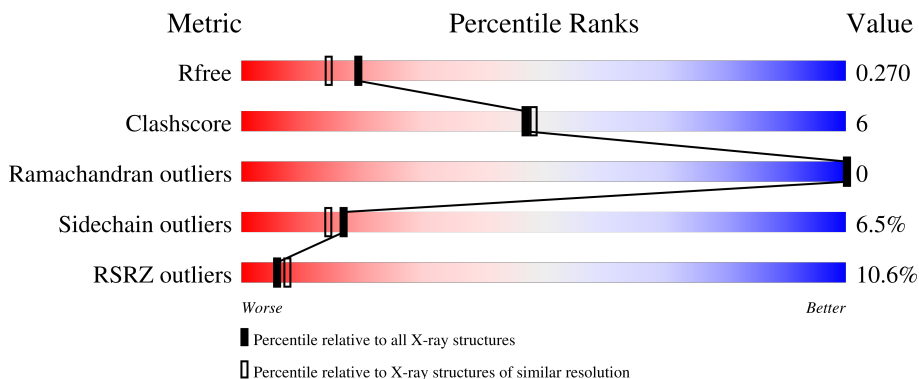|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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		0.270
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		6.5%
RSRZ outliers		10.6%

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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (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 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	225	12% / 72% / 9% · 17%
1	B	225	8% / 74% / 6% · 18%
1	C	225	8% / 69% / 12% · 16%
1	D	225	9% / 76% / 14% · 8%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6218 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 Propionate catabolism operon regulatory protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	186	Total	C	N	O	S	0	1	0
			1440	920	242	274	4			
1	B	184	Total	C	N	O	S	0	0	0
			1417	906	239	268	4			
1	C	188	Total	C	N	O	S	0	1	0
			1462	936	248	274	4			
1	D	208	Total	C	N	O	S	0	0	0
			1627	1033	287	303	4			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	cloning artifact	UNP P77743
A	0	SER	-	cloning artifact	UNP P77743
A	1	LEU	-	cloning artifact	UNP P77743
A	96	PHE	ALA	engineered mutation	UNP P77743
A	216	GLU	-	cloning artifact	UNP P77743
A	217	GLY	-	cloning artifact	UNP P77743
A	218	HIS	-	cloning artifact	UNP P77743
A	219	HIS	-	cloning artifact	UNP P77743
A	220	HIS	-	cloning artifact	UNP P77743
A	221	HIS	-	cloning artifact	UNP P77743
A	222	HIS	-	cloning artifact	UNP P77743
A	223	HIS	-	cloning artifact	UNP P77743
B	-1	MET	-	cloning artifact	UNP P77743
B	0	SER	-	cloning artifact	UNP P77743
B	1	LEU	-	cloning artifact	UNP P77743
B	96	PHE	ALA	engineered mutation	UNP P77743
B	216	GLU	-	cloning artifact	UNP P77743
B	217	GLY	-	cloning artifact	UNP P77743
B	218	HIS	-	cloning artifact	UNP P77743
B	219	HIS	-	cloning artifact	UNP P77743
B	220	HIS	-	cloning artifact	UNP P77743

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	221	HIS	-	cloning artifact	UNP P77743
B	222	HIS	-	cloning artifact	UNP P77743
B	223	HIS	-	cloning artifact	UNP P77743
C	-1	MET	-	cloning artifact	UNP P77743
C	0	SER	-	cloning artifact	UNP P77743
C	1	LEU	-	cloning artifact	UNP P77743
C	96	PHE	ALA	engineered mutation	UNP P77743
C	216	GLU	-	cloning artifact	UNP P77743
C	217	GLY	-	cloning artifact	UNP P77743
C	218	HIS	-	cloning artifact	UNP P77743
C	219	HIS	-	cloning artifact	UNP P77743
C	220	HIS	-	cloning artifact	UNP P77743
C	221	HIS	-	cloning artifact	UNP P77743
C	222	HIS	-	cloning artifact	UNP P77743
C	223	HIS	-	cloning artifact	UNP P77743
D	-1	MET	-	cloning artifact	UNP P77743
D	0	SER	-	cloning artifact	UNP P77743
D	1	LEU	-	cloning artifact	UNP P77743
D	96	PHE	ALA	engineered mutation	UNP P77743
D	216	GLU	-	cloning artifact	UNP P77743
D	217	GLY	-	cloning artifact	UNP P77743
D	218	HIS	-	cloning artifact	UNP P77743
D	219	HIS	-	cloning artifact	UNP P77743
D	220	HIS	-	cloning artifact	UNP P77743
D	221	HIS	-	cloning artifact	UNP P77743
D	222	HIS	-	cloning artifact	UNP P77743
D	223	HIS	-	cloning artifact	UNP P77743

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	68	Total 68	O 68	0	0
2	B	68	Total 68	O 68	0	0
2	C	95	Total 95	O 95	0	0
2	D	41	Total 41	O 41	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.
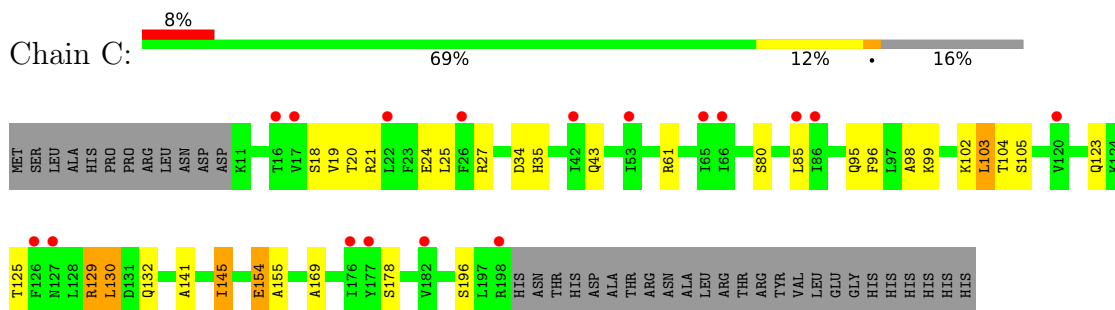
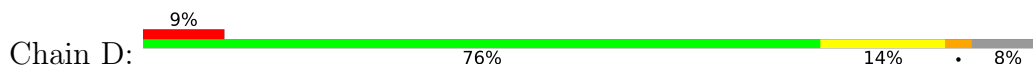• Molecule 1: Propionate catabolism operon regulatory protein
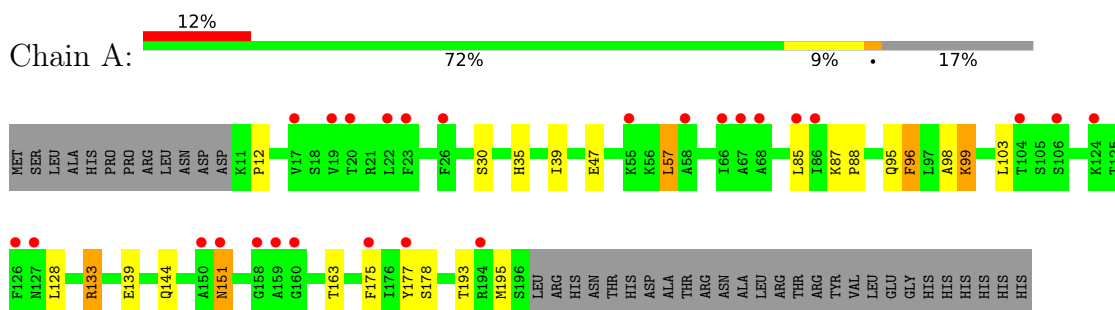
Chain A:

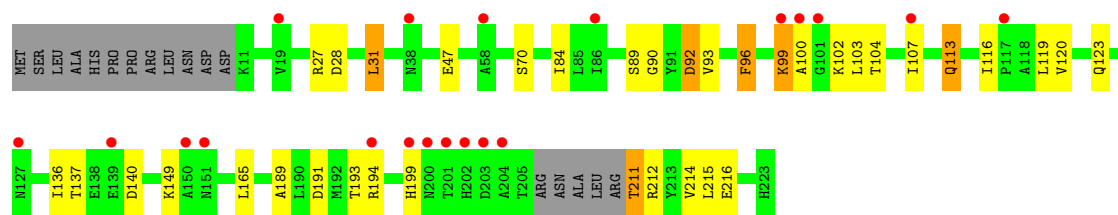• Molecule 1: Propionate catabolism operon regulatory protein

Chain B:

• Molecule 1: Propionate catabolism operon regulatory protein

Chain C:

• Molecule 1: Propionate catabolism operon regulatory protein

Chain D:
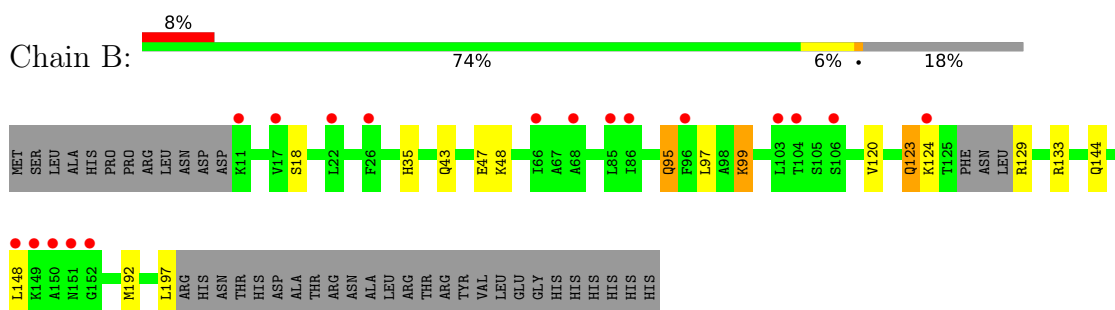
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	35.66Å 58.55Å 103.58Å 85.03° 88.56° 86.94°	Depositor
Resolution (Å)	37.06 – 2.10 30.31 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.8 (37.06-2.10) 97.8 (30.31-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.79 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.201 , 0.252 0.229 , 0.270	Depositor DCC
R_{free} test set	2436 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	36.4	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 51.5	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6218	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.78 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.4748e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 (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	0/1462	0.69	1/1981 (0.1%)
1	B	0.62	2/1437 (0.1%)	0.73	2/1944 (0.1%)
1	C	0.65	0/1484	0.72	0/2010
1	D	0.56	0/1657	0.66	0/2242
All	All	0.62	2/6040 (0.0%)	0.70	3/8177 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	129	ARG	NE-CZ	5.44	1.40	1.33
1	B	129	ARG	CZ-NH1	5.26	1.39	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	129	ARG	NE-CZ-NH2	-10.75	114.93	120.30
1	B	129	ARG	NE-CZ-NH1	6.58	123.59	120.30
1	A	57	LEU	CA-CB-CG	5.04	126.89	115.30

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	1440	0	1465	15	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1417	0	1457	10	0
1	C	1462	0	1506	24	0
1	D	1627	0	1632	28	0
2	A	68	0	0	2	0
2	B	68	0	0	0	0
2	C	95	0	0	1	0
2	D	41	0	0	2	0
All	All	6218	0	6060	74	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 6.

The worst 5 of 74 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:C:20:THR:HG23	1:C:21:ARG:H	1.36	0.90
1:C:21:ARG:HD3	2:C:303:HOH:O	1.82	0.79
1:D:90:GLY:HA2	1:D:119:LEU:HG	1.67	0.76
1:D:27:ARG:HG2	1:D:27:ARG:HH11	1.51	0.75
1:A:95:GLN:O	1:A:99:LYS:HD2	1.86	0.73

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	185/225 (82%)	180 (97%)	5 (3%)	0	100	100
1	B	180/225 (80%)	176 (98%)	4 (2%)	0	100	100
1	C	187/225 (83%)	182 (97%)	5 (3%)	0	100	100
1	D	204/225 (91%)	196 (96%)	8 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	756/900 (84%)	734 (97%)	22 (3%)	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	153/188 (81%)	146 (95%)	7 (5%)	27 26
1	B	151/188 (80%)	141 (93%)	10 (7%)	16 14
1	C	156/188 (83%)	144 (92%)	12 (8%)	13 9
1	D	172/188 (92%)	160 (93%)	12 (7%)	15 12
All	All	632/752 (84%)	591 (94%)	41 (6%)	17 14

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	154	GLU
1	D	113	GLN
1	D	31	LEU
1	D	92	ASP
1	D	149	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	35	HIS
1	C	43	GLN
1	D	199	HIS
1	C	123	GLN
1	D	35	HIS

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(Å^2)	Q<0.9
1	A	186/225 (82%)	0.71	26 (13%)	2	3	31, 38, 48, 55	0
1	B	184/225 (81%)	0.67	18 (9%)	7	10	32, 39, 47, 52	0
1	C	188/225 (83%)	0.69	17 (9%)	9	12	31, 38, 46, 53	0
1	D	208/225 (92%)	0.64	20 (9%)	8	10	30, 39, 51, 58	0
All	All	766/900 (85%)	0.68	81 (10%)	6	7	30, 39, 49, 58	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	104	THR	4.5
1	A	127	ASN	4.3
1	D	202	HIS	4.3
1	B	124	LYS	4.0
1	C	17	VAL	3.9

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.

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

6.5 Other polymers ⓘ

There are no such residues in this entry.