wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 01:57 AM EDT

PDB ID	:	4KSA
Title	:	Crystal Structure of Malonyl-CoA decarboxylase from Rhodopseudomonas palustris, Northeast Structural Genomics Consortium Target RpR127
Authors	:	Forouhar, F.; Neely, H.; Seetharaman, J.; Sahdev, S.; Xiao, R.; Patel, D.J.; Ciccosanti, C.; Wang, D.; Everett, J.K.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2013-05-17
Resolution	:	2.70 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

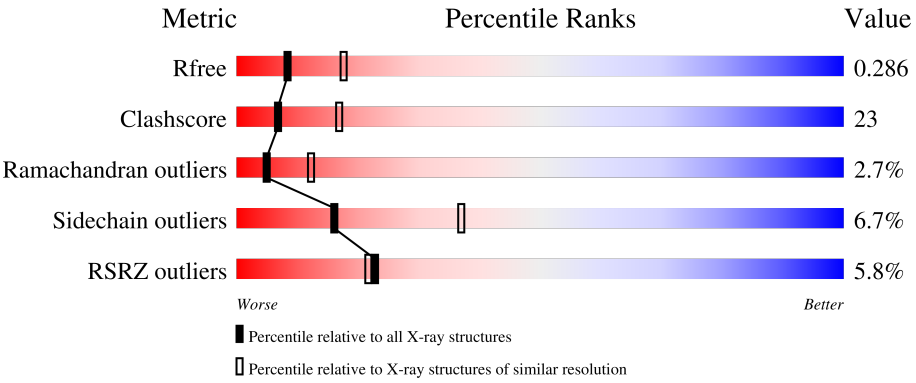
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	453	2% 58% 32% 5% .
1	B	453	10% 51% 34% 6% 9%
1	C	453	5% 57% 36% . .
1	D	453	4% 51% 35% 5% 9%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MG	C	501	-	-	-	X

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 13442 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 Malonyl-CoA decarboxylase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	437	Total	C	N	O	S	Se	0	0	0
			3405	2151	611	638	1	4			
1	B	411	Total	C	N	O	S	Se	0	0	0
			3208	2031	574	598	1	4			
1	C	436	Total	C	N	O	S	Se	0	0	0
			3398	2147	610	636	1	4			
1	D	413	Total	C	N	O	S	Se	0	0	0
			3222	2041	576	600	1	4			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	MSE	-	INITIATING METHIONINE	UNP Q6NCB2
A	452	LEU	-	EXPRESSION TAG	UNP Q6NCB2
A	453	GLU	-	EXPRESSION TAG	UNP Q6NCB2
A	454	HIS	-	EXPRESSION TAG	UNP Q6NCB2
A	455	HIS	-	EXPRESSION TAG	UNP Q6NCB2
A	456	HIS	-	EXPRESSION TAG	UNP Q6NCB2
A	457	HIS	-	EXPRESSION TAG	UNP Q6NCB2
A	458	HIS	-	EXPRESSION TAG	UNP Q6NCB2
A	459	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	7	MSE	-	INITIATING METHIONINE	UNP Q6NCB2
B	452	LEU	-	EXPRESSION TAG	UNP Q6NCB2
B	453	GLU	-	EXPRESSION TAG	UNP Q6NCB2
B	454	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	455	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	456	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	457	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	458	HIS	-	EXPRESSION TAG	UNP Q6NCB2
B	459	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	7	MSE	-	INITIATING METHIONINE	UNP Q6NCB2
C	452	LEU	-	EXPRESSION TAG	UNP Q6NCB2
C	453	GLU	-	EXPRESSION TAG	UNP Q6NCB2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	454	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	455	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	456	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	457	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	458	HIS	-	EXPRESSION TAG	UNP Q6NCB2
C	459	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	7	MSE	-	INITIATING METHIONINE	UNP Q6NCB2
D	452	LEU	-	EXPRESSION TAG	UNP Q6NCB2
D	453	GLU	-	EXPRESSION TAG	UNP Q6NCB2
D	454	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	455	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	456	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	457	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	458	HIS	-	EXPRESSION TAG	UNP Q6NCB2
D	459	HIS	-	EXPRESSION TAG	UNP Q6NCB2

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Mg 1 1	0	0
2	C	1	Total Mg 1 1	0	0

- Molecule 3 is water.

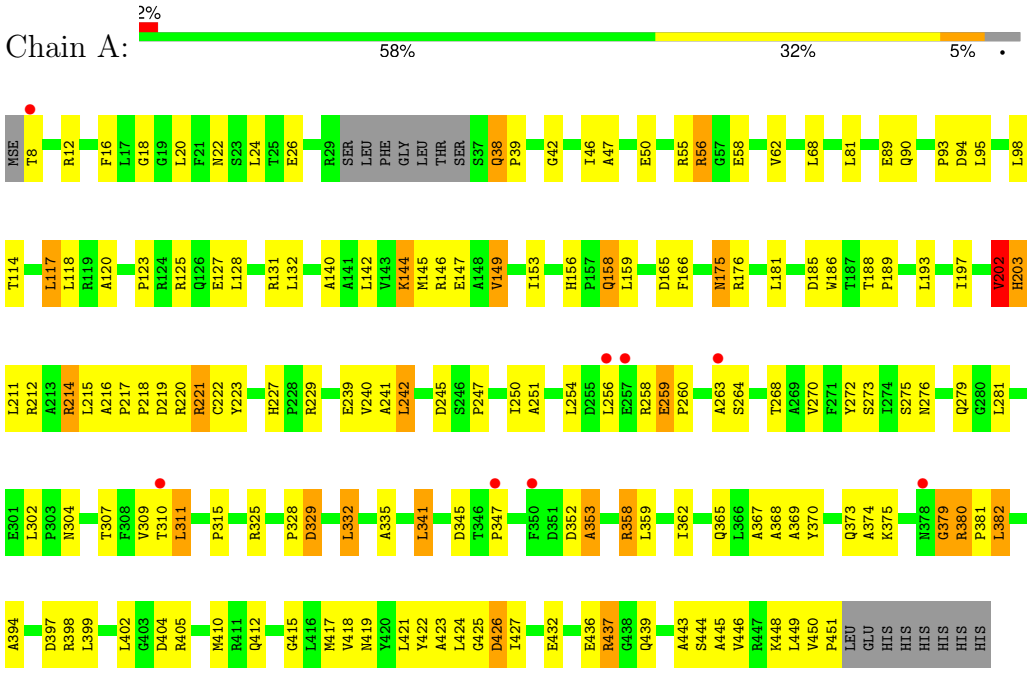
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	54	Total O 54 54	0	0
3	B	40	Total O 40 40	0	0
3	C	58	Total O 58 58	0	0
3	D	55	Total O 55 55	0	0

3 Residue-property plots (i)

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.

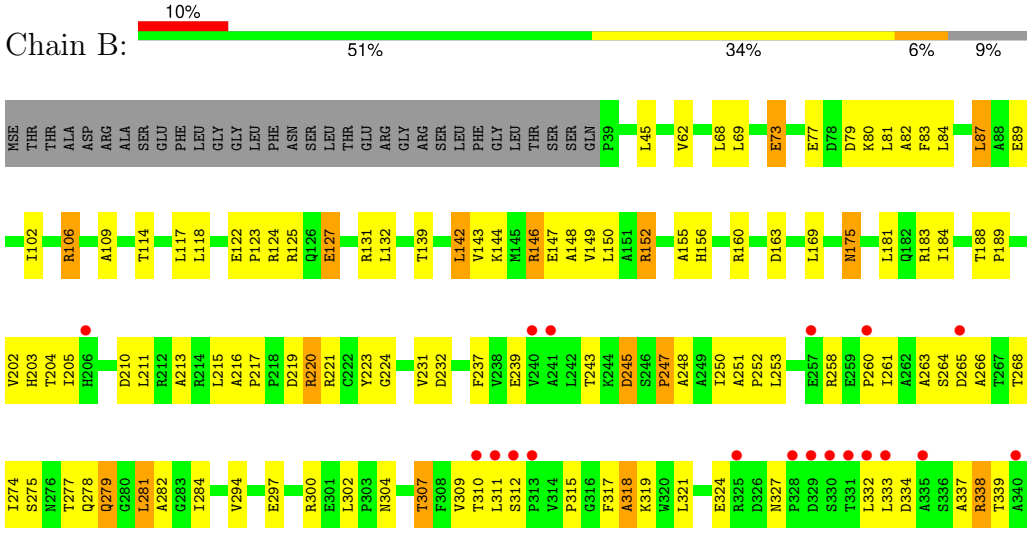
• Molecule 1: Malonyl-CoA decarboxylase

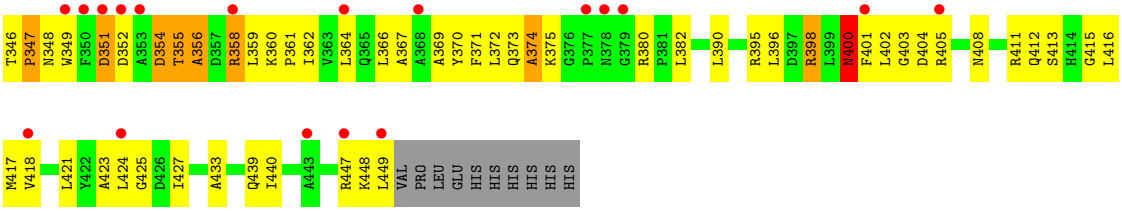
Chain A:



• Molecule 1: Malonyl-CoA decarboxylase

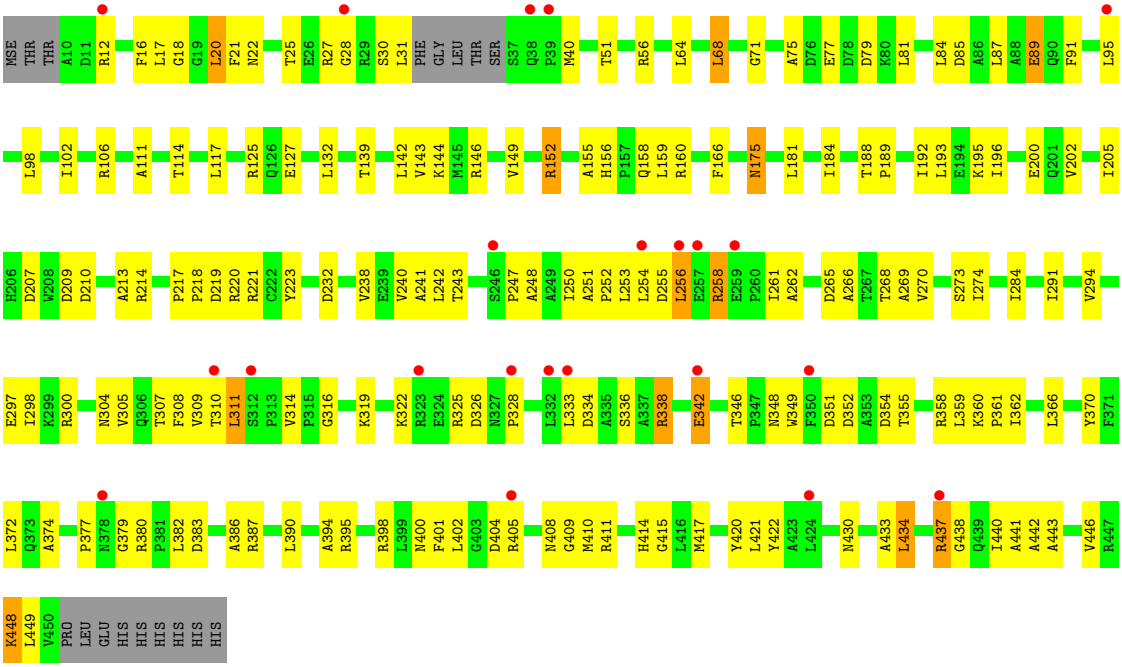
Chain B:





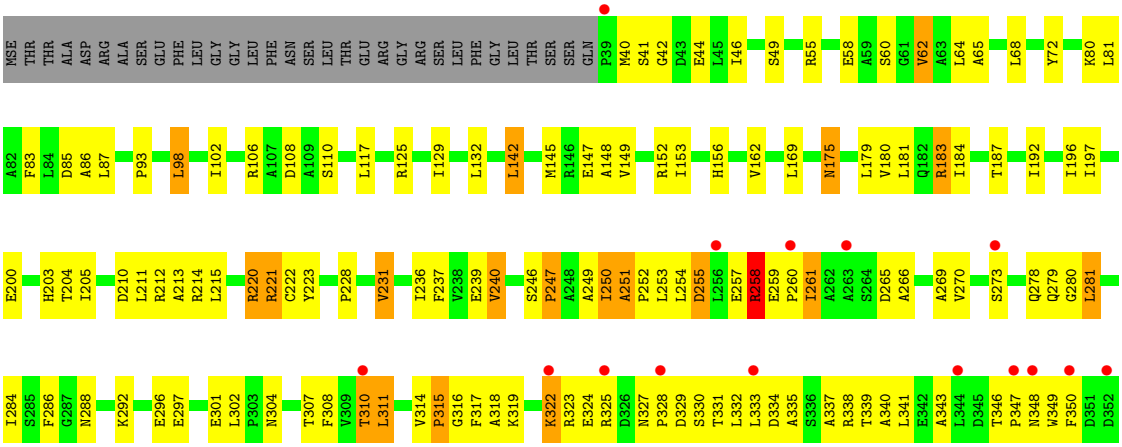
• Molecule 1: Malonyl-CoA decarboxylase

Chain C:



• Molecule 1: Malonyl-CoA decarboxylase

Chain D:




A353
D354
T355
A356
L359
L366
Y370
F371
L372
Q373
A374
K375
N378
L382
D383
P384
V385
A386
R387
F388
H389
L390
G391
N392
G393
A394
R395
L396
D397
R398
L399
L402
G403
D404
R405
M410
S413
H414
G415
L416
M417
V418
N419
Y422
A423
L424
I427
N430
R437
S444
A445
V446
R447
K448
L449
V450
P451
LEU
GLU
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	141.50Å 159.76Å 108.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.89 – 2.70 29.89 – 2.67	Depositor EDS
% Data completeness (in resolution range)	94.7 (29.89-2.70) 96.0 (29.89-2.67)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.64 (at 2.68Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.225 , 0.279 0.232 , 0.286	Depositor DCC
R_{free} test set	12810 reflections (9.70%)	wwPDB-VP
Wilson B-factor (Å^2)	56.9	Xtriage
Anisotropy	0.275	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 55.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	13442	wwPDB-VP
Average B, all atoms (Å^2)	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.63% 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: MG

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	0.42	0/3472	0.63	0/4711
1	B	0.37	0/3273	0.61	1/4443 (0.0%)
1	C	0.38	0/3464	0.58	0/4698
1	D	0.39	0/3288	0.61	1/4465 (0.0%)
All	All	0.39	0/13497	0.61	2/18317 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	277	THR	N-CA-C	5.75	126.51	111.00
1	D	311	LEU	N-CA-C	-5.07	97.30	111.00

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	3405	0	3369	142	0
1	B	3208	0	3180	176	0
1	C	3398	0	3364	142	0
1	D	3222	0	3196	163	0
2	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
3	A	54	0	0	5	0
3	B	40	0	0	0	0
3	C	58	0	0	1	0
3	D	55	0	0	1	0
All	All	13442	0	13109	611	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 23.

The worst 5 of 611 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:8:THR:HG22	1:A:12:ARG:HB2	1.43	0.98
1:B:318:ALA:HB2	1:B:412:GLN:HA	1.43	0.98
1:A:380:ARG:HH11	1:A:380:ARG:HB3	1.26	0.98
1:A:56:ARG:HB2	1:A:56:ARG:HH21	1.31	0.96
1:C:256:LEU:HD23	1:C:256:LEU:H	1.34	0.93

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	433/453 (96%)	388 (90%)	38 (9%)	7 (2%)	9 24
1	B	409/453 (90%)	358 (88%)	37 (9%)	14 (3%)	3 8
1	C	432/453 (95%)	374 (87%)	48 (11%)	10 (2%)	6 16
1	D	411/453 (91%)	359 (87%)	38 (9%)	14 (3%)	3 8
All	All	1685/1812 (93%)	1479 (88%)	161 (10%)	45 (3%)	5 12

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	202	VAL
1	B	354	ASP
1	B	355	THR
1	D	260	PRO
1	D	354	ASP

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	350/360 (97%)	323 (92%)	27 (8%)	13	30
1	B	329/360 (91%)	304 (92%)	25 (8%)	13	30
1	C	349/360 (97%)	330 (95%)	19 (5%)	22	47
1	D	331/360 (92%)	311 (94%)	20 (6%)	19	42
All	All	1359/1440 (94%)	1268 (93%)	91 (7%)	16	37

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	127	GLU
1	D	62	VAL
1	C	144	LYS
1	C	258	ARG
1	D	175	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	419	ASN
1	D	412	GLN
1	C	158	GLN
1	D	419	ASN
1	D	175	ASN

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

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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	433/453 (95%)	-0.02	11 (2%)	57	59	30, 57, 92, 109	0
1	B	407/453 (89%)	0.45	44 (10%)	5	4	32, 70, 152, 158	0
1	C	432/453 (95%)	0.19	22 (5%)	28	26	29, 70, 122, 136	0
1	D	409/453 (90%)	0.15	20 (4%)	29	28	31, 61, 128, 143	0
All	All	1681/1812 (92%)	0.19	97 (5%)	23	22	29, 63, 127, 158	0

The worst 5 of 97 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	332	LEU	11.1
1	B	350	PHE	6.7
1	D	350	PHE	6.5
1	D	347	PRO	6.1
1	C	333	LEU	5.6

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	MG	C	501	1/1	0.41	0.55	75,75,75,75	0
2	MG	A	501	1/1	0.82	0.54	64,64,64,64	0

6.5 Other polymers (i)

There are no such residues in this entry.