



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 08:33 am BST

PDB ID : 3HA4
Title : Crystal structure of the type one membrane protein MIX1 from Leishmania
Authors : Gorman, M.A.; Walsh, P.J.; Parker, M.W.
Deposited on : 2009-05-01
Resolution : 2.40 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

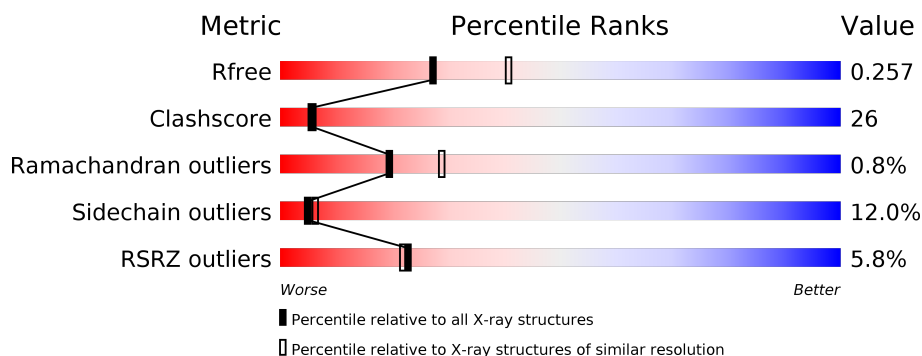
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 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 $\leq 5\%$. 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	154	<div> <div>2%</div> <div> <div></div> <div>51%</div> <div>25%</div> <div>6%</div> <div>18%</div> </div> </div>
1	B	154	<div> <div>4%</div> <div> <div></div> <div>47%</div> <div>25%</div> <div>5%</div> <div>22%</div> </div> </div>
1	C	154	<div> <div>6%</div> <div> <div></div> <div>42%</div> <div>23%</div> <div>8%</div> <div>27%</div> </div> </div>
1	D	154	<div> <div>0%</div> <div> <div></div> <div>51%</div> <div>21%</div> <div>•</div> <div>26%</div> </div> </div>
1	E	154	<div> <div>5%</div> <div> <div></div> <div>38%</div> <div>26%</div> <div>7%</div> <div>28%</div> </div> </div>
1	F	154	<div> <div>2%</div> <div> <div></div> <div>53%</div> <div>26%</div> <div>5%</div> <div>16%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	154	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>9%34%32%6%27%</div></div>
1	H	154	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>6%44%23%5%29%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7963 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 MIX1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	127	Total	C	N	O	S	0	1	0
			1055	665	187	198	5			
1	B	120	Total	C	N	O	S	0	0	0
			984	624	172	183	5			
1	C	113	Total	C	N	O	S	0	2	0
			939	598	160	176	5			
1	D	114	Total	C	N	O	S	0	1	0
			937	598	159	175	5			
1	E	111	Total	C	N	O	S	0	0	0
			907	580	155	167	5			
1	F	129	Total	C	N	O	S	0	1	0
			1070	675	186	204	5			
1	G	113	Total	C	N	O	S	0	1	0
			932	595	158	174	5			
1	H	109	Total	C	N	O	S	0	0	0
			889	569	150	165	5			

There are 32 discrepancies between the modelled and reference sequences:

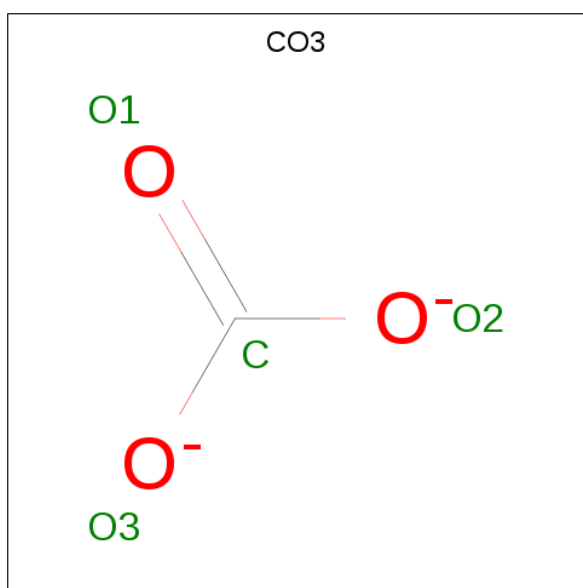
Chain	Residue	Modelled	Actual	Comment	Reference
A	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
A	43	SER	-	EXPRESSION TAG	UNP Q4QI49
A	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
A	45	MET	-	EXPRESSION TAG	UNP Q4QI49
B	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
B	43	SER	-	EXPRESSION TAG	UNP Q4QI49
B	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
B	45	MET	-	EXPRESSION TAG	UNP Q4QI49
C	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
C	43	SER	-	EXPRESSION TAG	UNP Q4QI49
C	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
C	45	MET	-	EXPRESSION TAG	UNP Q4QI49
D	42	GLY	-	EXPRESSION TAG	UNP Q4QI49

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	43	SER	-	EXPRESSION TAG	UNP Q4QI49
D	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
D	45	MET	-	EXPRESSION TAG	UNP Q4QI49
E	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
E	43	SER	-	EXPRESSION TAG	UNP Q4QI49
E	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
E	45	MET	-	EXPRESSION TAG	UNP Q4QI49
F	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
F	43	SER	-	EXPRESSION TAG	UNP Q4QI49
F	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
F	45	MET	-	EXPRESSION TAG	UNP Q4QI49
G	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
G	43	SER	-	EXPRESSION TAG	UNP Q4QI49
G	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
G	45	MET	-	EXPRESSION TAG	UNP Q4QI49
H	42	GLY	-	EXPRESSION TAG	UNP Q4QI49
H	43	SER	-	EXPRESSION TAG	UNP Q4QI49
H	44	HIS	-	EXPRESSION TAG	UNP Q4QI49
H	45	MET	-	EXPRESSION TAG	UNP Q4QI49

- Molecule 2 is CARBONATE ION (three-letter code: CO3) (formula: CO₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	1	3		

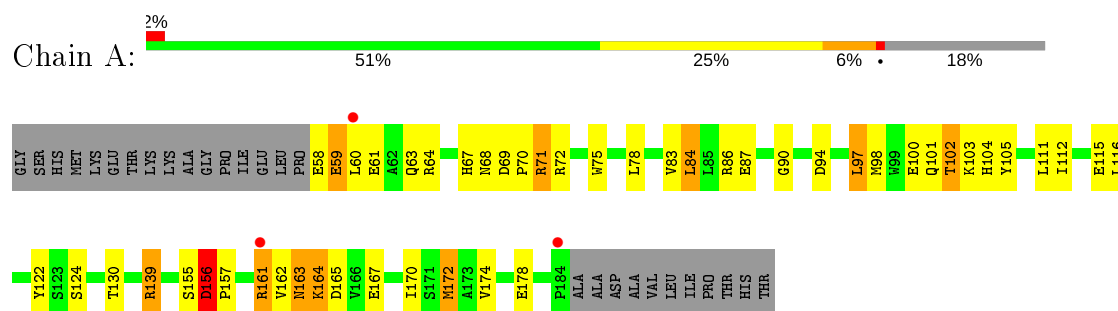
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	57	Total 57	O 57	0	0
3	B	53	Total 53	O 53	0	0
3	C	21	Total 21	O 21	0	0
3	D	32	Total 32	O 32	0	0
3	E	10	Total 10	O 10	0	0
3	F	41	Total 41	O 41	0	0
3	G	14	Total 14	O 14	0	0
3	H	18	Total 18	O 18	0	0

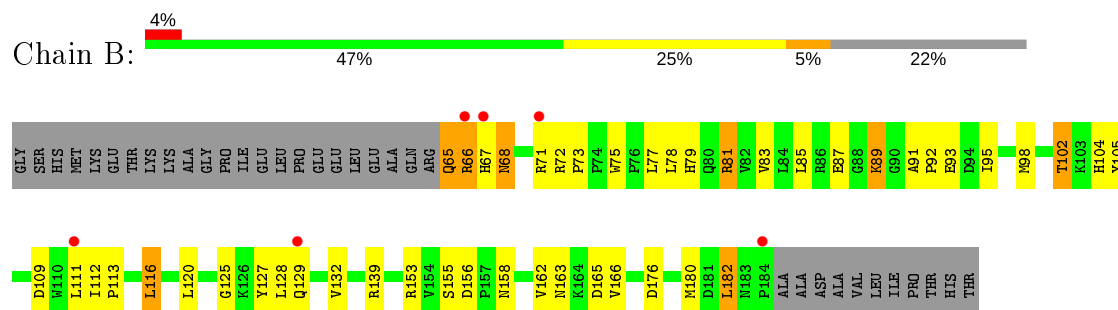
3 Residue-property plots [i](#)

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.

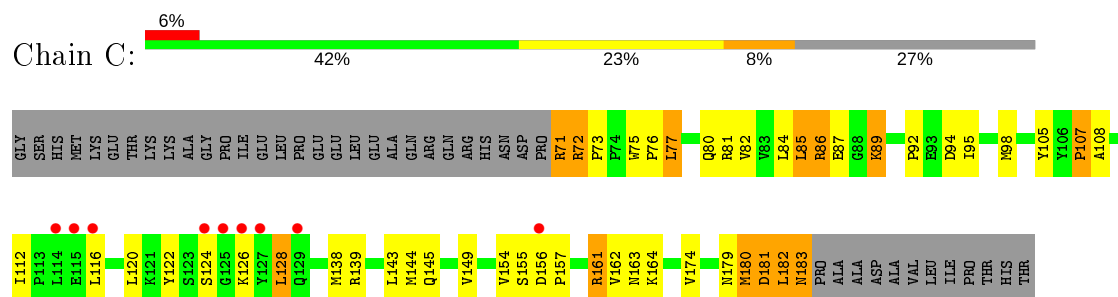
• Molecule 1: MIX1



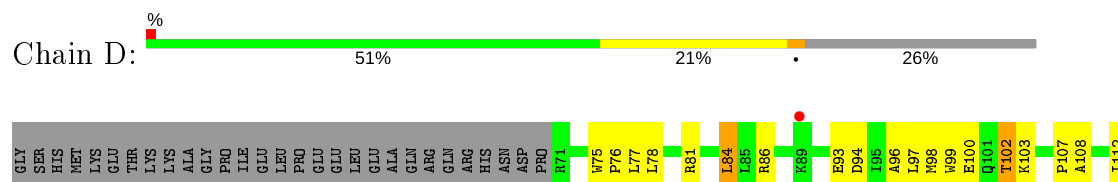
• Molecule 1: MIX1

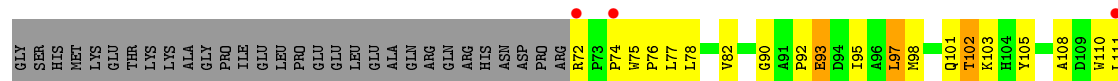


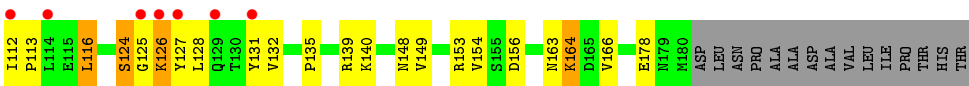
• Molecule 1: MIX1



• Molecule 1: MIX1







4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	92.52Å 100.39Å 140.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 2.40 25.83 – 1.99	Depositor EDS
% Data completeness (in resolution range)	99.3 (19.96-2.40) 94.9 (25.83-1.99)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.40 (at 1.99Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.212 , 0.261 0.212 , 0.257	Depositor DCC
R_{free} test set	4377 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å ²)	35.3	Xtriage
Anisotropy	0.280	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 48.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7963	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: CO3

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	1.19	0/1079	0.73	1/1464 (0.1%)
1	B	0.98	0/1008	0.68	0/1370
1	C	1.07	0/963	0.71	0/1307
1	D	0.79	0/959	0.61	0/1303
1	E	0.81	0/929	0.71	2/1261 (0.2%)
1	F	0.89	0/1094	0.78	2/1485 (0.1%)
1	G	0.88	0/954	0.64	0/1295
1	H	0.58	0/910	0.52	0/1236
All	All	0.92	0/7896	0.68	5/10721 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
1	F	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	156	ASP	C-N-CD	-6.62	106.05	120.60
1	A	84	LEU	CA-CB-CG	6.35	129.90	115.30
1	F	89	LYS	N-CA-C	6.06	127.36	111.00
1	F	182	LEU	CA-CB-CG	-5.81	101.94	115.30
1	E	157	PRO	N-CA-C	5.28	125.82	112.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	180	MET	Peptide
1	F	89	LYS	Peptide

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	1055	0	1042	75	1
1	B	984	0	975	46	1
1	C	939	0	936	74	0
1	D	937	0	931	32	0
1	E	907	0	908	78	0
1	F	1070	0	1052	45	0
1	G	932	0	930	69	1
1	H	889	0	889	36	1
2	A	4	0	0	0	0
3	A	57	0	0	11	0
3	B	53	0	0	3	0
3	C	21	0	0	3	0
3	D	32	0	0	4	0
3	E	10	0	0	1	0
3	F	41	0	0	1	0
3	G	14	0	0	2	0
3	H	18	0	0	1	0
All	All	7963	0	7663	397	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 26.

The worst 5 of 397 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:156:ASP:OD2	3:C:230:HOH:O	1.59	1.16
1:A:100:GLU:OE2	3:A:247:HOH:O	1.64	1.15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:108:ALA:HB3	1:C:156:ASP:HB3	1.27	1.15
1:C:86:ARG:HH11	1:C:86:ARG:HG3	0.94	1.10
1:C:86:ARG:HH11	1:C:86:ARG:CG	1.62	1.10

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:178:GLU:OE1	1:B:153:ARG:NH2[2_654]	1.96	0.24
1:G:153:ARG:NH1	1:H:178:GLU:OE2[3_655]	2.05	0.15

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	126/154 (82%)	122 (97%)	2 (2%)	2 (2%)	9	13
1	B	118/154 (77%)	113 (96%)	5 (4%)	0	100	100
1	C	113/154 (73%)	105 (93%)	7 (6%)	1 (1%)	17	25
1	D	113/154 (73%)	112 (99%)	1 (1%)	0	100	100
1	E	109/154 (71%)	100 (92%)	8 (7%)	1 (1%)	17	25
1	F	128/154 (83%)	122 (95%)	4 (3%)	2 (2%)	9	13
1	G	112/154 (73%)	105 (94%)	6 (5%)	1 (1%)	17	25
1	H	107/154 (70%)	101 (94%)	6 (6%)	0	100	100
All	All	926/1232 (75%)	880 (95%)	39 (4%)	7 (1%)	19	29

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	163	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	181	ASP
1	E	157	PRO
1	F	59	GLU
1	G	124	SER

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	116/136 (85%)	104 (90%)	12 (10%)	7	10
1	B	109/136 (80%)	93 (85%)	16 (15%)	3	3
1	C	104/136 (76%)	91 (88%)	13 (12%)	4	5
1	D	103/136 (76%)	95 (92%)	8 (8%)	12	19
1	E	100/136 (74%)	84 (84%)	16 (16%)	2	3
1	F	118/136 (87%)	106 (90%)	12 (10%)	7	10
1	G	103/136 (76%)	88 (85%)	15 (15%)	3	3
1	H	98/136 (72%)	88 (90%)	10 (10%)	7	10
All	All	851/1088 (78%)	749 (88%)	102 (12%)	5	6

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	154	VAL
1	E	126	LYS
1	H	102	THR
1	D	168	GLU
1	E	84	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	129	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	158	ASN
1	F	80	GLN
1	B	104	HIS
1	F	65	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](#)

1 ligand is modelled in this entry.

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	$\# Z > 2$
2	CO3	A	1	-	0,3,3	0.00	-	0,3,3	0.00	-

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 [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 ⓘ

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	127/154 (82%)	-0.36	3 (2%) 59 57	24, 37, 79, 143	0
1	B	120/154 (77%)	-0.06	6 (5%) 28 27	24, 43, 101, 128	0
1	C	113/154 (73%)	0.02	9 (7%) 12 11	27, 43, 83, 100	0
1	D	114/154 (74%)	-0.18	2 (1%) 68 66	20, 37, 75, 128	0
1	E	111/154 (72%)	0.30	7 (6%) 20 18	34, 66, 108, 144	0
1	F	129/154 (83%)	-0.31	3 (2%) 60 58	24, 41, 90, 126	0
1	G	113/154 (73%)	0.42	14 (12%) 4 3	31, 63, 101, 132	0
1	H	109/154 (70%)	0.19	10 (9%) 9 8	32, 62, 94, 114	0
All	All	936/1232 (75%)	-0.01	54 (5%) 23 22	20, 48, 95, 144	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	158	ASN	4.9
1	G	181	ASP	4.7
1	A	161	ARG	4.6
1	B	66	ARG	4.4
1	G	182	LEU	4.4

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 [i](#)

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, 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	B-factors(\AA^2)	Q<0.9
2	CO3	A	1	4/4	0.89	0.22	67,67,67,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.