



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:26 PM GMT

PDB ID : 4RWG  
Title : Crystal structure of the CLR:RAMP1 extracellular domain heterodimer with bound high affinity CGRP analog  
Authors : Booe, J.; Pioszak, A.  
Deposited on : 2014-12-03  
Resolution : 2.44 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto: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

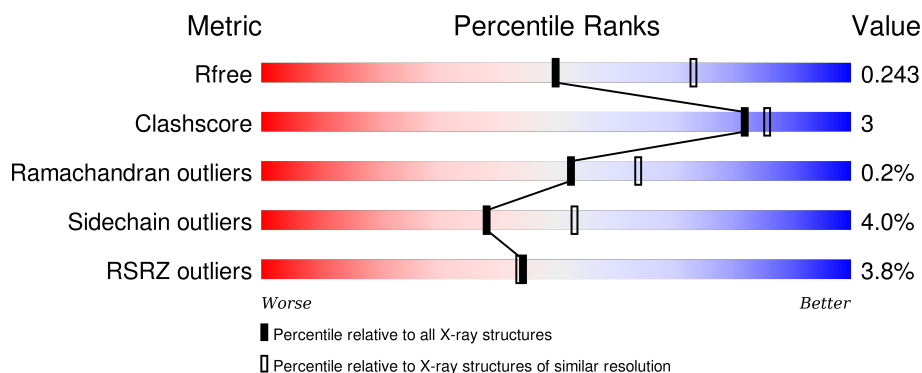
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.44 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	1003 (2.46-2.42)
Clashscore	102246	1071 (2.46-2.42)
Ramachandran outliers	100387	1065 (2.46-2.42)
Sidechain outliers	100360	1065 (2.46-2.42)
RSRZ outliers	91569	1005 (2.46-2.42)

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  $\geq 3$ , 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	593	<div> <div>7%</div> <div>85%</div> <div>7%</div> <div>6%</div> </div>
1	B	593	<div> <div>7%</div> <div>85%</div> <div>9%</div> <div>.</div> </div>
1	C	593	<div> <div>3%</div> <div>83%</div> <div>8%</div> <div>.</div> <div>8%</div> </div>
2	D	12	<div> <div>92%</div> <div>8%</div> </div>
2	E	12	<div> <div>8%</div> <div>92%</div> <div>8%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	12	 <div>92%8%</div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13593 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 Maltose-binding periplasmic protein, Receptor activity-modifying protein 1, Calcitonin gene-related peptide type 1 receptor fusion protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	557	Total	C	N	O	S	0	0	0
			4395	2805	732	835	23			
1	B	569	Total	C	N	O	S	0	0	0
			4503	2876	757	847	23			
1	C	546	Total	C	N	O	S	0	0	0
			4308	2753	716	816	23			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING METHIONINE	UNP P0AEX9
A	1109	ARG	-	LINKER	UNP Q16602
A	1110	ALA	-	LINKER	UNP Q16602
A	2019	VAL	-	EXPRESSION TAG	UNP Q16602
A	2020	GLY	-	EXPRESSION TAG	UNP Q16602
A	2021	SER	-	EXPRESSION TAG	UNP Q16602
A	2022	ALA	-	EXPRESSION TAG	UNP Q16602
A	2023	GLY	-	EXPRESSION TAG	UNP Q16602
A	2024	SER	-	EXPRESSION TAG	UNP Q16602
A	2025	ALA	-	EXPRESSION TAG	UNP Q16602
A	2026	GLY	-	EXPRESSION TAG	UNP Q16602
A	2027	SER	-	EXPRESSION TAG	UNP Q16602
A	2028	ALA	-	EXPRESSION TAG	UNP Q16602
A	2029	GLU	-	EXPRESSION TAG	UNP Q16602
A	2030	ASP	-	EXPRESSION TAG	UNP Q16602
B	1	MET	-	INITIATING METHIONINE	UNP P0AEX9
B	1109	ARG	-	LINKER	UNP Q16602
B	1110	ALA	-	LINKER	UNP Q16602
B	2019	VAL	-	EXPRESSION TAG	UNP Q16602
B	2020	GLY	-	EXPRESSION TAG	UNP Q16602
B	2021	SER	-	EXPRESSION TAG	UNP Q16602
B	2022	ALA	-	EXPRESSION TAG	UNP Q16602

*Continued on next page...*

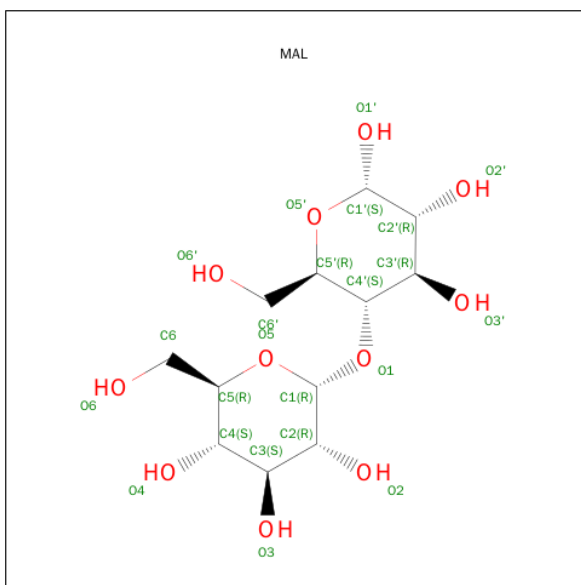
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	2023	GLY	-	EXPRESSION TAG	UNP Q16602
B	2024	SER	-	EXPRESSION TAG	UNP Q16602
B	2025	ALA	-	EXPRESSION TAG	UNP Q16602
B	2026	GLY	-	EXPRESSION TAG	UNP Q16602
B	2027	SER	-	EXPRESSION TAG	UNP Q16602
B	2028	ALA	-	EXPRESSION TAG	UNP Q16602
B	2029	GLU	-	EXPRESSION TAG	UNP Q16602
B	2030	ASP	-	EXPRESSION TAG	UNP Q16602
C	1	MET	-	INITIATING METHIONINE	UNP P0AEX9
C	1109	ARG	-	LINKER	UNP Q16602
C	1110	ALA	-	LINKER	UNP Q16602
C	2019	VAL	-	EXPRESSION TAG	UNP Q16602
C	2020	GLY	-	EXPRESSION TAG	UNP Q16602
C	2021	SER	-	EXPRESSION TAG	UNP Q16602
C	2022	ALA	-	EXPRESSION TAG	UNP Q16602
C	2023	GLY	-	EXPRESSION TAG	UNP Q16602
C	2024	SER	-	EXPRESSION TAG	UNP Q16602
C	2025	ALA	-	EXPRESSION TAG	UNP Q16602
C	2026	GLY	-	EXPRESSION TAG	UNP Q16602
C	2027	SER	-	EXPRESSION TAG	UNP Q16602
C	2028	ALA	-	EXPRESSION TAG	UNP Q16602
C	2029	GLU	-	EXPRESSION TAG	UNP Q16602
C	2030	ASP	-	EXPRESSION TAG	UNP Q16602

- Molecule 2 is a protein called CGRP analog.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	12	Total	C	N	O	0	0	1
			86	60	12	14			
2	E	12	Total	C	N	O	0	0	1
			86	60	12	14			
2	F	11	Total	C	N	O	0	0	1
			75	51	11	13			

- Molecule 3 is SUGAR (MALTOSE) (three-letter code: MAL) (formula: C<sub>12</sub>H<sub>22</sub>O<sub>11</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			23	12	11		
3	B	1	Total	C	O	0	0
			23	12	11		
3	C	1	Total	C	O	0	0
			23	12	11		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		

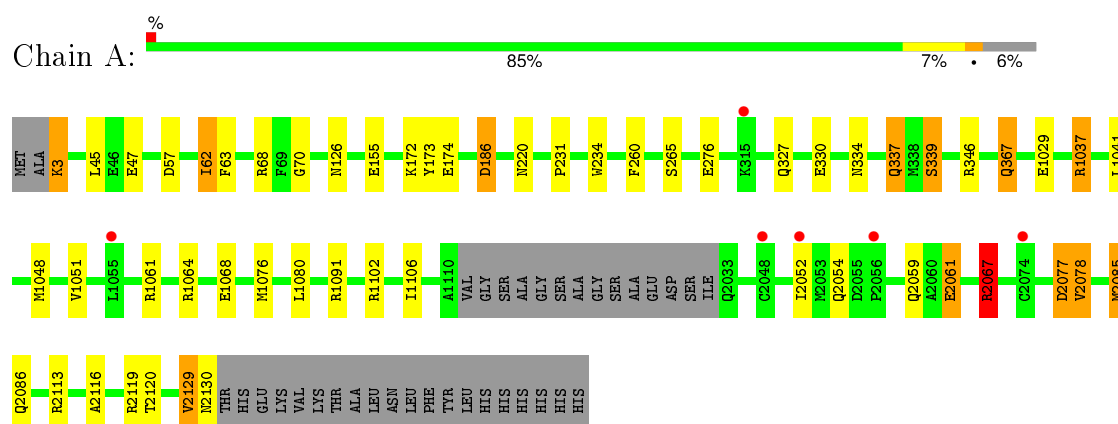
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	39	Total	O	0	0
			39	39		
5	B	23	Total	O	0	0
			23	23		
5	C	6	Total	O	0	0
			6	6		
5	D	1	Total	O	0	0
			1	1		
5	E	1	Total	O	0	0
			1	1		

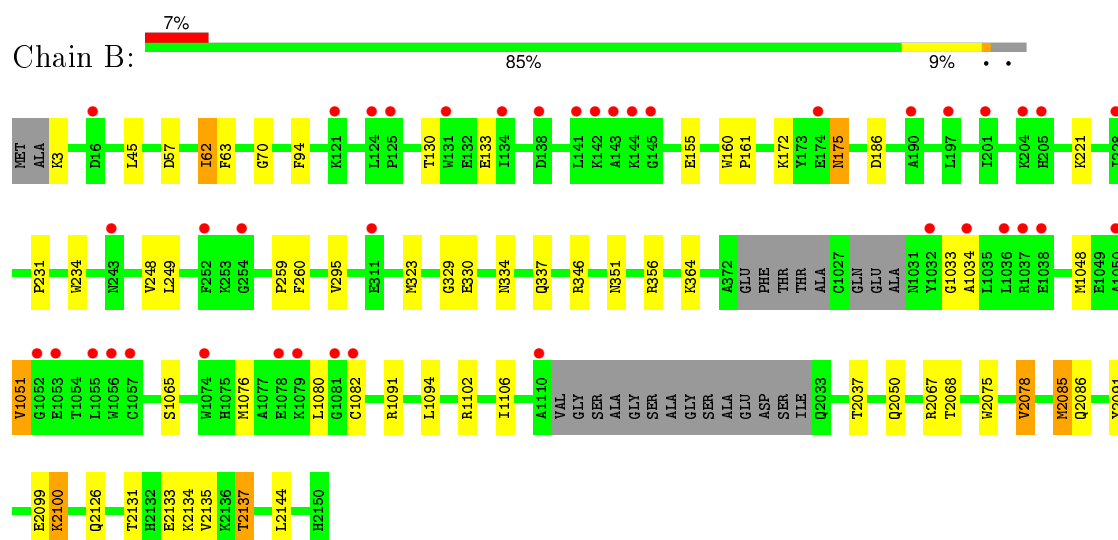
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

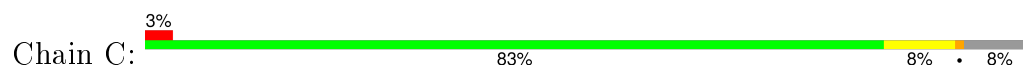
- Molecule 1: Maltose-binding periplasmic protein, Receptor activity-modifying protein 1, Calcitonin gene-related peptide type 1 receptor fusion protein



- Molecule 1: Maltose-binding periplasmic protein, Receptor activity-modifying protein 1, Calcitonin gene-related peptide type 1 receptor fusion protein



- Molecule 1: Maltose-binding periplasmic protein, Receptor activity-modifying protein 1, Calcitonin gene-related peptide type 1 receptor fusion protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	172.81Å 104.62Å 136.48Å 90.00° 122.43° 90.00°	Depositor
Resolution (Å)	50.00 – 2.44 42.51 – 2.44	Depositor EDS
% Data completeness (in resolution range)	99.4 (50.00-2.44) 99.4 (42.51-2.44)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.30 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.200 , 0.243 0.203 , 0.243	Depositor DCC
$R_{free}$ test set	3827 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.6	Xtriage
Anisotropy	0.325	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 47.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 76012 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, MAL, NH2

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	0.76	1/4507 (0.0%)	0.81	5/6122 (0.1%)
1	B	0.65	0/4621	0.75	2/6273 (0.0%)
1	C	0.55	0/4419	0.71	4/6003 (0.1%)
2	D	0.88	0/89	1.10	1/122 (0.8%)
2	E	0.69	0/89	0.77	0/122
2	F	0.67	0/77	0.82	0/106
All	All	0.66	1/13802 (0.0%)	0.76	12/18748 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	339	SER	CB-OG	-6.09	1.34	1.42

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	31	ASP	CB-CG-OD2	7.55	125.09	118.30
1	A	2067	ARG	CG-CD-NE	-6.97	97.16	111.80
1	A	1037	ARG	NE-CZ-NH1	6.40	123.50	120.30
1	B	1082	CYS	CA-CB-SG	6.30	125.34	114.00
1	B	2078	VAL	CB-CA-C	-6.29	99.45	111.40
1	A	2077	ASP	CB-CG-OD1	6.02	123.72	118.30
1	A	2078	VAL	CB-CA-C	-5.98	100.04	111.40
1	C	2078	VAL	CB-CA-C	-5.33	101.27	111.40
1	A	2067	ARG	CB-CG-CD	5.26	125.27	111.60
1	C	2067	ARG	CG-CD-NE	-5.21	100.85	111.80
1	C	365	ASP	CB-CG-OD2	-5.20	113.62	118.30
1	C	2070	ASP	CB-CG-OD1	5.10	122.89	118.30

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	4395	0	4235	32	0
1	B	4503	0	4326	31	0
1	C	4308	0	4149	25	0
2	D	86	0	77	0	0
2	E	86	0	77	0	0
2	F	75	0	68	0	0
3	A	23	0	22	0	0
3	B	23	0	22	0	0
3	C	23	0	22	0	0
4	B	1	0	0	0	0
5	A	39	0	0	3	0
5	B	23	0	0	2	0
5	C	6	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0
All	All	13593	0	12998	84	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 3.

All (84) 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:2137:THR:HG21	1:C:310:GLU:OE2	1.82	0.79
1:A:47:GLU:HG2	5:A:2315:HOH:O	1.83	0.78
1:A:1076:MET:O	1:A:1080:LEU:HD12	1.85	0.76
1:C:1076:MET:O	1:C:1080:LEU:HD12	1.87	0.75
1:A:2129:VAL:O	1:A:2129:VAL:HG12	1.90	0.72
1:B:1076:MET:O	1:B:1080:LEU:HD12	1.97	0.65
1:B:2134:LYS:O	5:B:2308:HOH:O	2.15	0.64
1:B:2135:VAL:O	1:B:2135:VAL:HG12	1.98	0.64
1:B:2137:THR:CG2	1:C:310:GLU:OE2	2.46	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:323:MET:CE	1:B:323:MET:HA	2.29	0.63
1:B:2144:LEU:HB3	1:C:94:PHE:CE2	2.35	0.61
1:C:49:PHE:CE1	1:C:62:ILE:HG12	2.36	0.60
1:A:2052:ILE:HG12	1:A:2067:ARG:HB2	1.83	0.60
1:A:2085:MET:HG2	1:A:2086:GLN:N	2.19	0.57
1:A:173:TYR:O	1:A:174:GLU:HG3	2.04	0.56
1:A:70:GLY:HA3	1:A:334:ASN:O	2.05	0.55
1:A:173:TYR:O	1:A:174:GLU:CG	2.55	0.55
1:B:323:MET:HE2	1:B:323:MET:HA	1.89	0.54
1:C:2106:ASP:OD2	1:C:2110:ASN:HB2	2.07	0.53
1:A:3:LYS:HB2	1:A:57:ASP:OD1	2.08	0.53
1:A:2120:THR:HG22	5:A:2318:HOH:O	2.09	0.53
1:A:186:ASP:OD1	1:A:367:GLN:OE1	2.27	0.52
1:B:1106:ILE:HA	1:B:2050:GLN:HB3	1.90	0.52
1:B:130:THR:OG1	1:B:133:GLU:OE1	2.20	0.52
1:C:94:PHE:CZ	1:C:331:ILE:HD12	2.46	0.51
1:C:2050:GLN:O	1:C:2053:MET:HB3	2.10	0.50
1:B:351:ASN:OD1	1:B:356:ARG:NH1	2.45	0.49
1:B:295:VAL:HG12	5:B:2315:HOH:O	2.13	0.49
1:C:231:PRO:HA	1:C:234:TRP:CE2	2.48	0.49
1:B:221:LYS:HE3	1:C:290:GLU:HG3	1.93	0.49
1:B:259:PRO:O	1:B:329:GLY:HA3	2.13	0.48
1:A:2061:GLU:HB2	5:A:2306:HOH:O	2.12	0.48
1:C:1107:SER:HA	1:C:2047:GLU:OE1	2.13	0.48
1:C:1064:ARG:NH1	1:C:1068:GLU:OE2	2.45	0.48
1:C:2067:ARG:HG2	1:C:2077:ASP:OD1	2.13	0.48
1:A:337:GLN:H	1:A:337:GLN:HE21	1.61	0.48
1:B:62:ILE:HG13	1:B:63:PHE:N	2.29	0.48
1:B:2085:MET:HG2	1:B:2086:GLN:N	2.29	0.48
1:A:68:ARG:NH1	1:A:339:SER:HB2	2.29	0.47
1:C:1106:ILE:HD11	1:C:2054:GLN:HG2	1.97	0.47
1:B:2067:ARG:HG3	1:B:2075:TRP:O	2.14	0.47
1:A:1037:ARG:O	1:A:1041:LEU:HB3	2.15	0.47
1:B:2135:VAL:O	1:B:2135:VAL:CG1	2.62	0.47
1:B:231:PRO:HA	1:B:234:TRP:CE2	2.48	0.47
1:B:2037:THR:HG23	1:B:2091:TYR:HB3	1.97	0.47
1:A:231:PRO:HA	1:A:234:TRP:CE2	2.50	0.46
1:B:2099:GLU:HG2	1:B:2126:GLN:O	2.15	0.46
1:C:70:GLY:HA3	1:C:334:ASN:O	2.15	0.46
1:A:1048:MET:HA	1:A:1051:VAL:HG12	1.98	0.46
1:B:2100:LYS:O	1:B:2100:LYS:HG3	2.14	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:3:LYS:CB	1:A:57:ASP:OD1	2.63	0.46
1:A:339:SER:OG	1:A:2116:ALA:HB2	2.16	0.46
1:A:62:ILE:HG13	1:A:63:PHE:N	2.30	0.46
1:C:2085:MET:HG2	1:C:2086:GLN:N	2.31	0.46
1:C:2094:ASP:O	1:C:2128:ASN:HB3	2.16	0.45
1:B:70:GLY:HA3	1:B:334:ASN:O	2.16	0.45
1:A:337:GLN:H	1:A:337:GLN:NE2	2.15	0.44
1:A:2129:VAL:O	1:A:2129:VAL:CG1	2.63	0.44
1:C:2099:GLU:HG2	1:C:2126:GLN:O	2.17	0.44
1:C:1048:MET:HA	1:C:1051:VAL:HG12	1.98	0.44
1:A:155:GLU:OE1	1:A:346:ARG:NH2	2.51	0.43
1:B:1033:GLY:O	1:B:1034:ALA:C	2.57	0.43
1:C:1109:ARG:O	1:C:1110:ALA:C	2.57	0.42
1:B:1094:LEU:HD23	1:B:1094:LEU:HA	1.70	0.42
1:B:155:GLU:OE1	1:B:346:ARG:NH2	2.52	0.42
1:A:2067:ARG:HG2	1:A:2077:ASP:OD1	2.20	0.42
1:A:1064:ARG:NH1	1:A:1068:GLU:HG3	2.35	0.42
1:A:276:GLU:OE1	1:A:276:GLU:N	2.49	0.42
1:B:1048:MET:HA	1:B:1051:VAL:HG12	2.02	0.42
1:A:1106:ILE:HD11	1:A:2054:GLN:HG3	2.02	0.42
1:C:1107:SER:HA	1:C:2047:GLU:CD	2.40	0.41
1:C:68:ARG:NH1	1:C:339:SER:HB2	2.36	0.41
1:C:94:PHE:CE2	1:C:331:ILE:HD12	2.55	0.41
1:A:2130:ASN:OD1	1:A:2130:ASN:C	2.58	0.41
1:B:2067:ARG:HG2	1:B:2068:THR:N	2.35	0.41
1:A:63:PHE:HA	1:A:265:SER:O	2.21	0.41
1:C:1048:MET:HA	1:C:1051:VAL:CG1	2.51	0.41
1:A:220:ASN:HD22	1:A:220:ASN:N	2.19	0.41
1:A:2113:ARG:HA	1:A:2119:ARG:O	2.21	0.41
1:B:248:VAL:HG12	1:B:249:LEU:O	2.21	0.41
1:A:1048:MET:HA	1:A:1051:VAL:CG1	2.51	0.40
1:B:2099:GLU:CG	1:B:2126:GLN:O	2.70	0.40
1:C:87:PHE:HA	1:C:90:LYS:HE2	2.03	0.40
1:B:160:TRP:N	1:B:161:PRO:CD	2.85	0.40

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	553/593 (93%)	541 (98%)	11 (2%)	1 (0%)	52	64
1	B	562/593 (95%)	553 (98%)	8 (1%)	1 (0%)	52	64
1	C	540/593 (91%)	525 (97%)	15 (3%)	0	100	100
2	D	10/12 (83%)	9 (90%)	1 (10%)	0	100	100
2	E	10/12 (83%)	8 (80%)	1 (10%)	1 (10%)	1	0
2	F	9/12 (75%)	9 (100%)	0	0	100	100
All	All	1684/1815 (93%)	1645 (98%)	36 (2%)	3 (0%)	52	64

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	175	ASN
2	E	29	PRO
1	A	2129	VAL

### 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	456/484 (94%)	436 (96%)	20 (4%)	35	49
1	B	467/484 (96%)	445 (95%)	22 (5%)	32	45
1	C	447/484 (92%)	433 (97%)	14 (3%)	47	64
2	D	9/9 (100%)	9 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	9/9 (100%)	9 (100%)	0	100	100
2	F	8/9 (89%)	8 (100%)	0	100	100
All	All	1396/1479 (94%)	1340 (96%)	56 (4%)	38	53

All (56) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	LYS
1	A	45	LEU
1	A	62	ILE
1	A	126	ASN
1	A	172	LYS
1	A	186	ASP
1	A	260	PHE
1	A	327	GLN
1	A	330	GLU
1	A	337	GLN
1	A	367	GLN
1	A	1029	GLU
1	A	1061	ARG
1	A	1091	ARG
1	A	1102	ARG
1	A	2059	GLN
1	A	2061	GLU
1	A	2067	ARG
1	A	2078	VAL
1	A	2085	MET
1	B	3	LYS
1	B	45	LEU
1	B	57	ASP
1	B	62	ILE
1	B	94	PHE
1	B	172	LYS
1	B	175	ASN
1	B	186	ASP
1	B	260	PHE
1	B	330	GLU
1	B	337	GLN
1	B	364	LYS
1	B	1051	VAL
1	B	1065	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	1091	ARG
1	B	1102	ARG
1	B	2078	VAL
1	B	2085	MET
1	B	2100	LYS
1	B	2131	THR
1	B	2133	GLU
1	B	2137	THR
1	C	45	LEU
1	C	62	ILE
1	C	172	LYS
1	C	186	ASP
1	C	260	PHE
1	C	330	GLU
1	C	337	GLN
1	C	1038	GLU
1	C	1045	GLN
1	C	1091	ARG
1	C	2038	ARG
1	C	2067	ARG
1	C	2078	VAL
1	C	2085	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	51	GLN
1	A	102	ASN
1	A	220	ASN
1	A	337	GLN
1	A	1045	GLN
1	A	2086	GLN
1	B	102	ASN
1	B	175	ASN
1	B	220	ASN
1	B	337	GLN
1	B	1045	GLN
1	B	2045	GLN
1	B	2086	GLN
1	B	2093	GLN
1	C	102	ASN
1	C	220	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	337	GLN
1	C	2086	GLN

### 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 4 ligands modelled in this entry, 1 is 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	MAL	A	2201	-	24,24,24	0.88	1 (4%)	35,35,35	1.64	8 (22%)
3	MAL	B	2201	-	24,24,24	0.56	0	35,35,35	1.64	6 (17%)
3	MAL	C	2201	-	24,24,24	0.61	0	35,35,35	1.24	3 (8%)

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	MAL	A	2201	-	-	0/8/48/48	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MAL	B	2201	-	-	0/8/48/48	0/2/2/2
3	MAL	C	2201	-	-	0/8/48/48	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2201	MAL	O5'-C5'	-2.27	1.38	1.44

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	2201	MAL	C4-C3-C2	-3.05	105.11	110.79
3	A	2201	MAL	O1-C1-O5	-2.61	104.08	110.68
3	A	2201	MAL	C6'-C5'-C4'	-2.55	105.83	113.25
3	B	2201	MAL	C1-O1-C4'	-2.54	111.37	118.01
3	C	2201	MAL	O4-C4-C3	-2.51	104.68	110.34
3	C	2201	MAL	O2-C2-C3	-2.32	105.12	110.34
3	A	2201	MAL	O2-C2-C1	-2.26	105.06	110.02
3	B	2201	MAL	O1'-C1'-C2'	-2.24	103.20	109.21
3	A	2201	MAL	C4-C3-C2	-2.16	106.76	110.79
3	A	2201	MAL	C3-C4-C5	2.06	113.78	110.20
3	B	2201	MAL	C1'-C2'-C3'	2.23	113.74	110.43
3	C	2201	MAL	O5'-C1'-C2'	2.82	114.30	109.80
3	A	2201	MAL	O5'-C1'-C2'	3.14	114.81	109.80
3	A	2201	MAL	C1'-C2'-C3'	3.44	115.55	110.43
3	B	2201	MAL	C1'-O5'-C5'	3.47	119.88	113.47
3	A	2201	MAL	C1'-O5'-C5'	4.44	121.68	113.47
3	B	2201	MAL	O5'-C1'-C2'	5.26	118.19	109.80

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	557/593 (93%)	-0.18	6 (1%) 82 84	35, 67, 106, 132	0
1	B	569/593 (95%)	0.25	40 (7%) 19 17	36, 86, 138, 191	0
1	C	546/593 (92%)	0.11	17 (3%) 52 51	62, 93, 127, 165	0
2	D	11/12 (91%)	-0.34	0 100 100	51, 61, 99, 112	0
2	E	11/12 (91%)	0.44	1 (9%) 11 10	79, 93, 124, 134	0
2	F	10/12 (83%)	-0.21	0 100 100	67, 81, 94, 101	0
All	All	1704/1815 (93%)	0.06	64 (3%) 44 44	35, 83, 127, 191	0

All (64) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1038	GLU	5.3
1	B	141	LEU	4.9
1	C	1106	ILE	4.8
1	B	131	TRP	4.7
1	B	201	ILE	4.6
1	B	1052	GLY	4.4
1	B	1055	LEU	4.2
1	C	35	ILE	3.9
2	E	27	PHE	3.9
1	C	34	GLY	3.9
1	A	315	LYS	3.7
1	B	143	ALA	3.5
1	B	190	ALA	3.4
1	C	1107	SER	3.4
1	C	241	LYS	3.3
1	B	1056	TRP	3.2
1	B	124	LEU	3.2
1	B	1110	ALA	3.1
1	C	2109	GLY	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	175	ASN	3.0
1	B	1079	LYS	3.0
1	B	125	PRO	2.9
1	B	1074	TRP	2.9
1	A	1055	LEU	2.9
1	B	1057	CYS	2.8
1	B	197	LEU	2.8
1	B	1050	ALA	2.8
1	A	2052	ILE	2.7
1	B	1081	GLY	2.7
1	C	174	GLU	2.6
1	B	1032	TYR	2.5
1	C	177	LYS	2.5
1	B	243	ASN	2.5
1	B	254	GLY	2.5
1	B	138	ASP	2.4
1	B	1082	CYS	2.4
1	A	2074	CYS	2.4
1	B	204	LYS	2.4
1	C	36	LYS	2.3
1	C	1028	GLN	2.3
1	B	134	ILE	2.3
1	B	1078	GLU	2.3
1	B	1037	ARG	2.3
1	B	1053	GLU	2.3
1	B	121	LYS	2.3
1	C	2055	ASP	2.3
1	C	2064	TYR	2.3
1	B	16	ASP	2.3
1	B	144	LYS	2.2
1	A	2056	PRO	2.2
1	B	228	ILE	2.2
1	B	205	HIS	2.2
1	A	2048	CYS	2.2
1	B	1034	ALA	2.2
1	B	311	GLU	2.2
1	B	1036	LEU	2.2
1	C	31	LYS	2.2
1	C	277	LEU	2.1
1	B	145	GLY	2.1
1	C	2080	ALA	2.1
1	C	1094	LEU	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	252	PHE	2.0
1	B	174	GLU	2.0
1	B	142	LYS	2.0

## 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	MAL	C	2201	23/23	0.97	0.21	1.98	66,74,83,85	0
3	MAL	A	2201	23/23	0.98	0.22	1.89	42,45,49,56	0
3	MAL	B	2201	23/23	0.97	0.26	0.95	59,65,68,75	0
4	MG	B	2202	1/1	0.94	0.05	-2.60	54,54,54,54	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.