



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 10:37 am BST

PDB ID : 5VVH  
Title : Crystal Structure of the Effector Binding Domain of LysR-type Transcriptional Regulator, OccR from *Agrobacterium tumefaciens*  
Authors : Kim, Y.; Chhor, G.; Jedrzejczak, R.; Winans, S.C.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2017-05-19  
Resolution : 2.50 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

<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

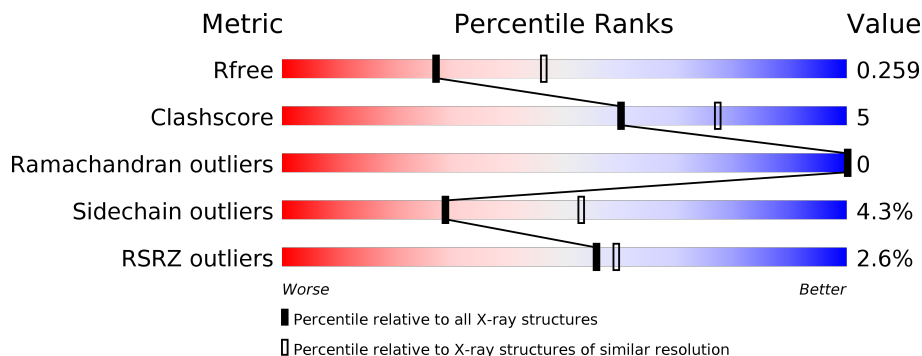
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 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	210	
1	B	210	
1	C	210	
1	D	210	
1	E	210	
1	F	210	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	210	 89% 10% ..
1	H	210	 4% 88% 11% .

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 12985 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 Octopine catabolism/uptake operon regulatory protein OccR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	208	Total 1626	C 1029	N 291	O 296	Se 10	0	4	0
1	B	209	Total 1618	C 1025	N 289	O 294	Se 10	0	2	0
1	C	208	Total 1610	C 1020	N 287	O 294	Se 9	0	2	0
1	D	210	Total 1616	C 1023	N 289	O 295	Se 9	0	1	0
1	E	208	Total 1593	C 1010	N 286	O 289	Se 8	0	0	0
1	F	208	Total 1601	C 1015	N 287	O 290	Se 9	0	1	0
1	G	208	Total 1594	C 1011	N 285	O 290	Se 8	0	0	0
1	H	209	Total 1613	C 1024	N 288	O 293	Se 8	0	1	0

There are 32 discrepancies between the modelled and reference sequences:

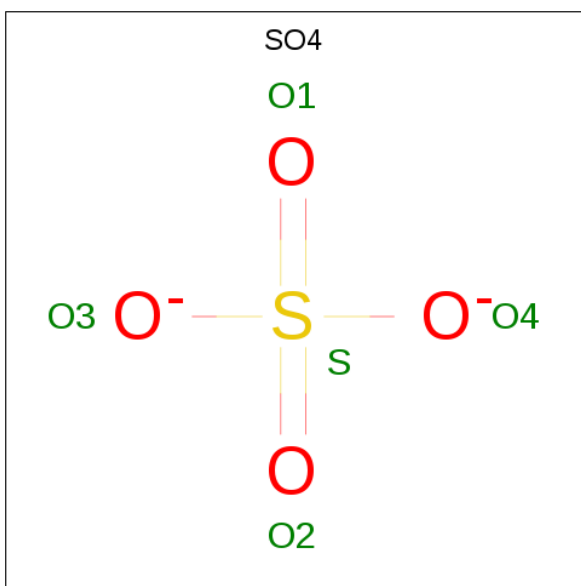
Chain	Residue	Modelled	Actual	Comment	Reference
A	89	SER	-	expression tag	UNP P0A4T3
A	90	ASN	-	expression tag	UNP P0A4T3
A	91	ALA	-	expression tag	UNP P0A4T3
A	261	GLU	GLY	cloning artifact	UNP P0A4T3
B	89	SER	-	expression tag	UNP P0A4T3
B	90	ASN	-	expression tag	UNP P0A4T3
B	91	ALA	-	expression tag	UNP P0A4T3
B	261	GLU	GLY	cloning artifact	UNP P0A4T3
C	89	SER	-	expression tag	UNP P0A4T3
C	90	ASN	-	expression tag	UNP P0A4T3
C	91	ALA	-	expression tag	UNP P0A4T3
C	261	GLU	GLY	cloning artifact	UNP P0A4T3
D	89	SER	-	expression tag	UNP P0A4T3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	90	ASN	-	expression tag	UNP P0A4T3
D	91	ALA	-	expression tag	UNP P0A4T3
D	261	GLU	GLY	cloning artifact	UNP P0A4T3
E	89	SER	-	expression tag	UNP P0A4T3
E	90	ASN	-	expression tag	UNP P0A4T3
E	91	ALA	-	expression tag	UNP P0A4T3
E	261	GLU	GLY	cloning artifact	UNP P0A4T3
F	89	SER	-	expression tag	UNP P0A4T3
F	90	ASN	-	expression tag	UNP P0A4T3
F	91	ALA	-	expression tag	UNP P0A4T3
F	261	GLU	GLY	cloning artifact	UNP P0A4T3
G	89	SER	-	expression tag	UNP P0A4T3
G	90	ASN	-	expression tag	UNP P0A4T3
G	91	ALA	-	expression tag	UNP P0A4T3
G	261	GLU	GLY	cloning artifact	UNP P0A4T3
H	89	SER	-	expression tag	UNP P0A4T3
H	90	ASN	-	expression tag	UNP P0A4T3
H	91	ALA	-	expression tag	UNP P0A4T3
H	261	GLU	GLY	cloning artifact	UNP P0A4T3

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



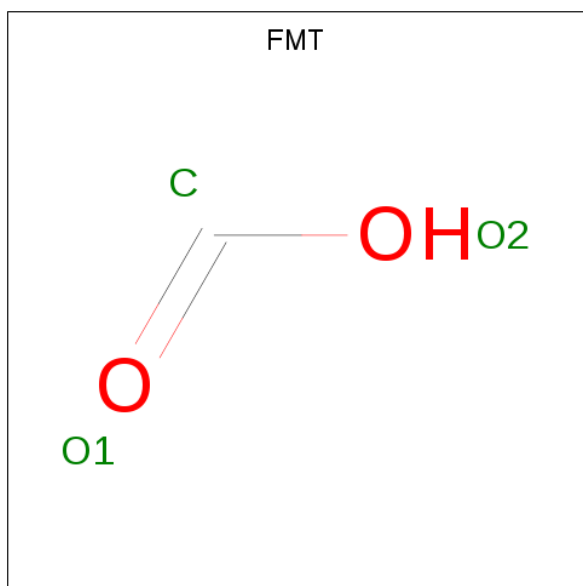
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf		
2	A	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	H	1	Total O S 5 4 1	0	0

- Molecule 3 is FORMIC ACID (three-letter code: FMT) (formula: CH<sub>2</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 3 1 2	0	0
3	B	1	Total C O 3 1 2	0	0
3	B	1	Total C O 3 1 2	0	0
3	D	1	Total C O 3 1 2	0	0
3	G	1	Total C O 3 1 2	0	0
3	G	1	Total C O 3 1 2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	13	Total O 13 13	0	0
4	B	10	Total O 10 10	0	0

Continued on next page...

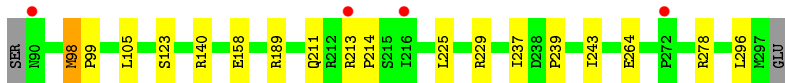
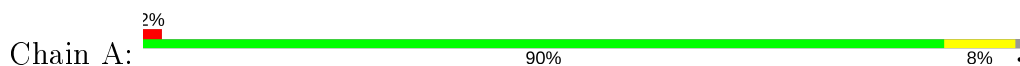
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	C	11	Total O 11 11	0	0
4	D	6	Total O 6 6	0	0
4	E	6	Total O 6 6	0	0
4	F	12	Total O 12 12	0	0
4	G	11	Total O 11 11	0	0
4	H	12	Total O 12 12	0	0

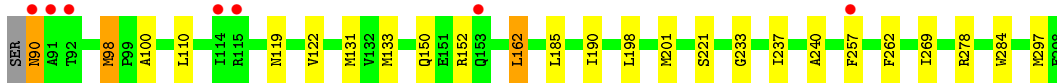
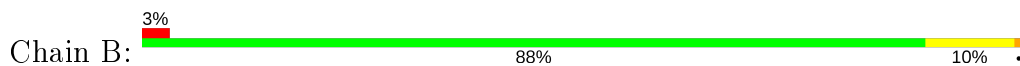
### 3 Residue-property plots [i](#)

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.

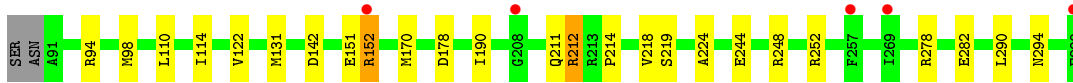
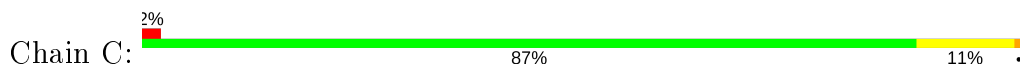
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



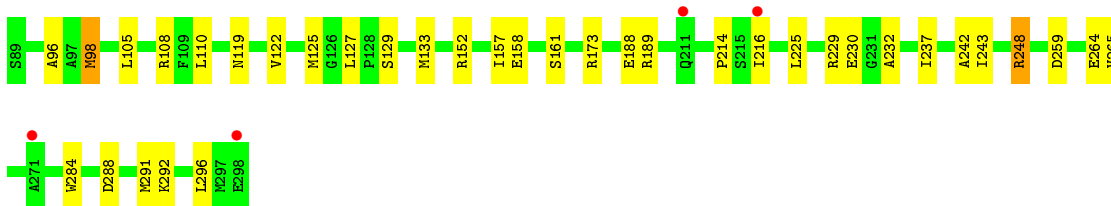
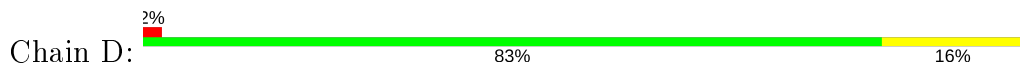
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



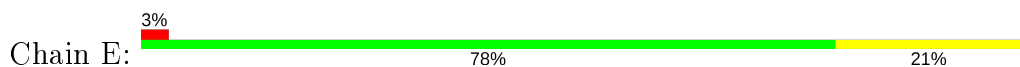
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



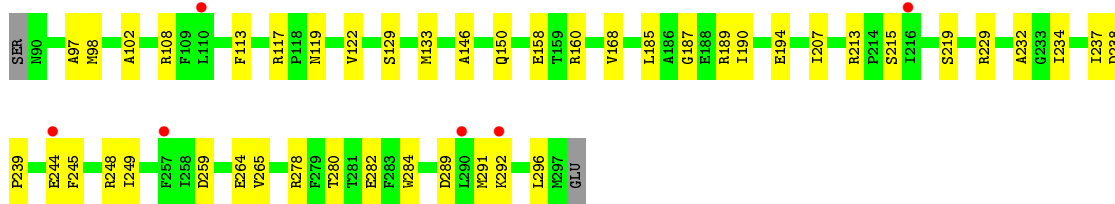
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



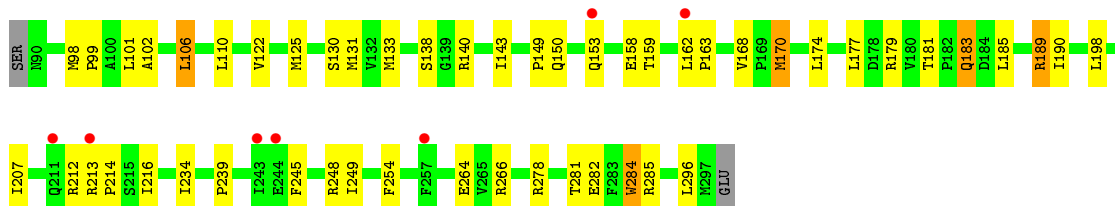
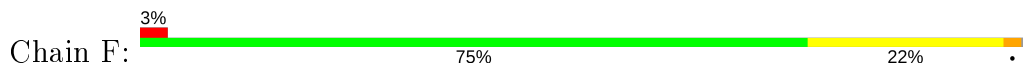
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



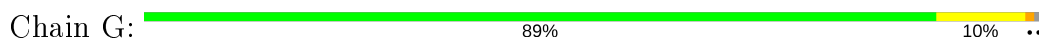




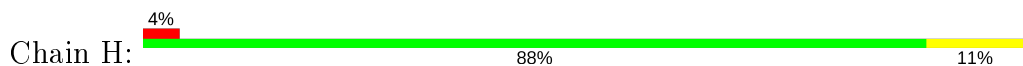
- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



- Molecule 1: Octopine catabolism/uptake operon regulatory protein OccR



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.39Å 106.33Å 215.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.74 – 2.50 47.69 – 2.49	Depositor EDS
% Data completeness (in resolution range)	94.9 (45.74-2.50) 94.3 (47.69-2.49)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.08 (at 2.48Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.218 , 0.259 0.218 , 0.259	Depositor DCC
$R_{free}$ test set	2770 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.2	Xtrriage
Anisotropy	0.157	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 31.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12985	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.90% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: FMT, SO4

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	0.20	0/1646	0.40	0/2215
1	B	0.21	0/1638	0.38	0/2204
1	C	0.21	0/1630	0.39	0/2194
1	D	0.21	0/1636	0.39	0/2202
1	E	0.21	0/1613	0.39	0/2172
1	F	0.20	0/1621	0.39	0/2182
1	G	0.22	0/1614	0.40	0/2173
1	H	0.21	0/1634	0.40	0/2200
All	All	0.21	0/13032	0.39	0/17542

There are no bond length outliers.

There are no bond angle outliers.

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	1626	0	1655	8	0
1	B	1618	0	1651	13	0
1	C	1610	0	1640	13	0
1	D	1616	0	1648	23	0
1	E	1593	0	1629	22	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1601	0	1637	27	0
1	G	1594	0	1629	10	0
1	H	1613	0	1643	12	0
2	A	5	0	0	0	0
2	D	5	0	0	0	0
2	H	5	0	0	0	0
3	A	3	0	1	0	0
3	B	6	0	2	1	0
3	D	3	0	1	0	0
3	G	6	0	2	0	0
4	A	13	0	0	0	0
4	B	10	0	0	0	0
4	C	11	0	0	0	0
4	D	6	0	0	0	0
4	E	6	0	0	0	0
4	F	12	0	0	0	0
4	G	11	0	0	0	0
4	H	12	0	0	0	0
All	All	12985	0	13138	120	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 5.

All (120) 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:G:225:LEU:HD21	1:G:237:ILE:HD13	1.61	0.82
1:B:233:GLY:HA2	3:B:301:FMT:H	1.73	0.71
1:H:98:MSE:HE2	1:H:100:ALA:HB3	1.81	0.63
1:C:114:ILE:HG13	1:D:216:ILE:HD11	1.81	0.62
1:C:190:ILE:HD12	1:C:212:ARG:HG2	1.80	0.62
1:A:189:ARG:HH21	1:A:213:ARG:HH11	1.47	0.62
1:F:174:LEU:HA	1:F:177:LEU:HD13	1.81	0.61
1:B:162:LEU:HD11	1:B:297:MSE:HG3	1.84	0.59
1:F:278:ARG:NH1	1:F:282:GLU:OE2	2.35	0.59
1:D:291:MSE:HG3	1:D:296:LEU:HB2	1.84	0.59
1:D:225:LEU:HD21	1:D:237:ILE:HD13	1.83	0.59
1:E:119:ASN:HA	1:F:214:PRO:HG2	1.86	0.58
1:H:162:LEU:HB3	1:H:239:PRO:HG2	1.84	0.58
1:G:237:ILE:HD12	1:G:242:ALA:HB2	1.84	0.57
1:B:185:LEU:HB3	1:B:190:ILE:HD11	1.86	0.57

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:214:PRO:HG3	1:D:119:ASN:HA	1.86	0.57
1:B:110:LEU:HD11	1:B:122:VAL:HG11	1.86	0.56
1:E:185:LEU:HB3	1:E:190:ILE:HD11	1.86	0.56
1:D:237:ILE:HD12	1:D:242:ALA:HB2	1.88	0.56
1:A:214:PRO:HG3	1:B:119:ASN:HA	1.87	0.56
1:E:189:ARG:NH2	1:E:232:ALA:O	2.39	0.56
1:F:168:VAL:HG12	1:F:234:ILE:HD12	1.86	0.56
1:F:189:ARG:HG2	1:F:213:ARG:HB2	1.88	0.55
1:H:239:PRO:HB3	1:H:296:LEU:HD13	1.88	0.55
1:H:168:VAL:HG12	1:H:234:ILE:HD12	1.89	0.54
1:F:179:ARG:NH1	1:F:254:PHE:O	2.40	0.53
1:E:168:VAL:HG12	1:E:234:ILE:HD12	1.90	0.53
1:G:240:ALA:HA	1:G:291:MSE:HE1	1.90	0.53
1:G:179:ARG:HH21	1:G:255:SER:HA	1.73	0.53
1:D:110:LEU:HD11	1:D:122:VAL:HG11	1.90	0.53
1:H:162:LEU:HD23	1:H:239:PRO:HB2	1.91	0.52
1:G:158:GLU:HG2	1:G:284:TRP:NE1	2.24	0.52
1:C:94:ARG:NH2	1:C:142:ASP:OD1	2.43	0.52
1:E:239:PRO:HB2	1:E:291:MSE:HE1	1.91	0.51
1:D:189:ARG:NH1	1:D:214:PRO:O	2.43	0.51
1:B:90:ASN:HD22	1:B:90:ASN:N	2.08	0.51
1:G:190:ILE:HD12	1:G:212:ARG:HG2	1.92	0.50
1:C:110:LEU:HD11	1:C:122:VAL:HG11	1.93	0.50
1:E:133:MSE:SE	1:E:265:VAL:HG21	2.61	0.50
1:A:105:LEU:HD21	1:A:243:ILE:HG21	1.94	0.50
1:D:173:ARG:NH2	1:D:188:GLU:OE2	2.37	0.49
1:E:219:SER:HB3	1:F:125[A]:MSE:SE	2.62	0.49
1:E:289:ASP:HA	1:E:292:LYS:HE3	1.94	0.49
1:E:129:SER:HB3	1:E:146:ALA:HB1	1.95	0.49
1:E:229:ARG:HG2	1:E:248:ARG:HH21	1.78	0.49
1:B:221:SER:HB2	1:B:237:ILE:HD11	1.94	0.49
1:F:110:LEU:HD11	1:F:122:VAL:HG11	1.93	0.49
1:D:108:ARG:HH12	1:D:243:ILE:HG22	1.77	0.49
1:D:229:ARG:HD3	1:D:248:ARG:CZ	2.43	0.48
1:D:264:GLU:OE1	1:D:284:TRP:NE1	2.46	0.48
1:F:133:MSE:SE	1:F:150:GLN:HG2	2.64	0.48
1:B:198:LEU:HA	1:B:201:MSE:HE3	1.95	0.48
1:F:185:LEU:HB3	1:F:190:ILE:HD11	1.95	0.48
1:F:99:PRO:HA	1:F:102:ALA:HB3	1.96	0.48
1:G:168:VAL:HG12	1:G:234:ILE:HD12	1.94	0.48
1:H:162:LEU:HD22	1:H:262[B]:PHE:HZ	1.78	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:229:ARG:NH2	1:D:230:GLU:OE2	2.43	0.48
1:E:239:PRO:HB2	1:E:296:LEU:HD13	1.96	0.48
1:E:133:MSE:SE	1:E:150:GLN:HG2	2.63	0.48
1:E:97:ALA:HB3	1:E:102:ALA:HB2	1.96	0.47
1:A:239:PRO:HB2	1:A:296:LEU:HD13	1.95	0.47
1:F:143:ILE:HD11	1:F:264:GLU:HG2	1.97	0.47
1:F:284:TRP:CE3	1:F:284:TRP:HA	2.50	0.47
1:C:244:GLU:O	1:C:248:ARG:NH2	2.47	0.47
1:C:278:ARG:NH1	1:C:282:GLU:OE2	2.49	0.46
1:A:225:LEU:HD21	1:A:237:ILE:HD13	1.97	0.46
1:D:96:ALA:HB1	1:D:127:LEU:HD12	1.97	0.46
1:H:167:ALA:HB1	1:H:249:ILE:HD12	1.98	0.45
1:H:229:ARG:HD3	1:H:248:ARG:HD2	1.97	0.45
1:F:162:LEU:HA	1:F:163:PRO:HD3	1.88	0.45
1:F:281:THR:O	1:F:285:ARG:HG3	2.17	0.45
1:A:140:ARG:HH11	1:C:211:GLN:HE22	1.64	0.45
1:E:278:ARG:O	1:E:282:GLU:HB2	2.16	0.45
1:C:218:VAL:HG21	1:C:224:ALA:HB2	2.00	0.44
1:G:162:LEU:HD21	1:G:297:MSE:HG3	1.99	0.44
1:G:133:MSE:SE	1:G:150:GLN:HG2	2.68	0.44
1:D:133:MSE:SE	1:D:265:VAL:HG21	2.67	0.44
1:E:113:PHE:O	1:E:117:ARG:NH1	2.49	0.44
1:F:106:LEU:HD12	1:F:106:LEU:HA	1.83	0.44
1:B:133:MSE:SE	1:B:150:GLN:HG2	2.68	0.43
1:B:284:TRP:HA	1:B:284:TRP:CE3	2.52	0.43
1:C:219:SER:HB3	1:D:125[A]:MSE:SE	2.69	0.43
1:F:245:PHE:HB3	1:F:249:ILE:HD12	2.00	0.43
1:D:189:ARG:HD2	1:D:232:ALA:O	2.18	0.43
1:C:290:LEU:O	1:C:294:ASN:ND2	2.32	0.43
1:D:105:LEU:HD21	1:D:243:ILE:HG21	2.00	0.43
1:D:98:MSE:HA	1:D:127:LEU:O	2.18	0.43
1:C:178:ASP:HA	1:C:252:ARG:HH11	1.84	0.43
1:E:158:GLU:HB3	1:E:264:GLU:HB3	2.01	0.42
1:H:170:MSE:HG2	1:H:250:VAL:HG23	2.01	0.42
1:B:162:LEU:HA	1:B:162:LEU:HD12	1.80	0.42
1:D:225:LEU:HD21	1:D:237:ILE:CD1	2.50	0.42
1:E:245:PHE:O	1:E:249:ILE:HG12	2.20	0.42
1:G:129:SER:O	1:G:133:MSE:HG2	2.18	0.42
1:H:185:LEU:HB3	1:H:190:ILE:HD11	2.02	0.42
1:E:187:GLY:O	1:E:213:ARG:NH1	2.53	0.42
1:E:122:VAL:HG22	1:F:216:ILE:HG12	2.02	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:239:PRO:HB3	1:F:296:LEU:HD22	2.01	0.42
1:B:240:ALA:HB3	1:B:262:PHE:HE2	1.84	0.42
1:F:207:ILE:HD11	1:F:254:PHE:CZ	2.55	0.42
1:D:152:ARG:HD3	1:D:157:ILE:HD12	2.01	0.41
1:E:108:ARG:HD2	1:E:244:GLU:OE2	2.20	0.41
1:C:152:ARG:HG2	1:C:152:ARG:H	1.46	0.41
1:E:160:ARG:HG3	1:E:284:TRP:CZ3	2.54	0.41
1:F:138:SER:OG	1:F:140:ARG:NE	2.54	0.41
1:H:134:GLU:HG3	1:H:140:ARG:NH2	2.36	0.41
1:B:98:MSE:HE2	1:B:100:ALA:HB3	2.02	0.41
1:F:149:PRO:HD3	1:F:198:LEU:HD22	2.02	0.41
1:F:170:MSE:HE1	1:F:248:ARG:HA	2.01	0.41
1:F:207:ILE:HD11	1:F:254:PHE:HZ	1.86	0.41
1:A:98:MSE:HA	1:A:99:PRO:HD3	1.95	0.41
1:F:181:THR:CG2	1:F:183:GLN:HG2	2.51	0.41
1:D:105:LEU:HD11	1:D:291:MSE:HE1	2.03	0.41
1:D:288:ASP:O	1:D:292:LYS:HB2	2.21	0.41
1:A:158:GLU:HB3	1:A:264:GLU:HB3	2.02	0.40
1:D:129:SER:O	1:D:133:MSE:HG2	2.21	0.40
1:E:284:TRP:HA	1:E:284:TRP:CE3	2.56	0.40
1:F:101:LEU:HD23	1:F:101:LEU:HA	1.94	0.40
1:H:101:LEU:O	1:H:106:LEU:HB2	2.21	0.40
1:F:170:MSE:HE2	1:F:170:MSE:HB2	1.96	0.40

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	210/210 (100%)	208 (99%)	2 (1%)	0	100 100
1	B	209/210 (100%)	203 (97%)	6 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	208/210 (99%)	204 (98%)	4 (2%)	0	100	100
1	D	209/210 (100%)	204 (98%)	5 (2%)	0	100	100
1	E	206/210 (98%)	205 (100%)	1 (0%)	0	100	100
1	F	207/210 (99%)	204 (99%)	3 (1%)	0	100	100
1	G	206/210 (98%)	202 (98%)	4 (2%)	0	100	100
1	H	208/210 (99%)	205 (99%)	3 (1%)	0	100	100
All	All	1663/1680 (99%)	1635 (98%)	28 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	171/161 (106%)	166 (97%)	5 (3%)	42	69
1	B	170/161 (106%)	161 (95%)	9 (5%)	22	43
1	C	169/161 (105%)	162 (96%)	7 (4%)	30	55
1	D	170/161 (106%)	165 (97%)	5 (3%)	42	69
1	E	167/161 (104%)	159 (95%)	8 (5%)	25	48
1	F	168/161 (104%)	155 (92%)	13 (8%)	13	25
1	G	167/161 (104%)	161 (96%)	6 (4%)	35	61
1	H	169/161 (105%)	162 (96%)	7 (4%)	30	55
All	All	1351/1288 (105%)	1291 (96%)	60 (4%)	29	52

All (60) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	98	MSE
1	A	123	SER
1	A	211	GLN
1	A	229	ARG

Continued on next page...



*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>
1	A	278	ARG
1	B	90	ASN
1	B	98	MSE
1	B	131[A]	MSE
1	B	131[B]	MSE
1	B	152	ARG
1	B	162	LEU
1	B	257	PHE
1	B	269	ILE
1	B	278	ARG
1	C	98	MSE
1	C	131[A]	MSE
1	C	131[B]	MSE
1	C	151	GLU
1	C	152	ARG
1	C	170	MSE
1	C	212	ARG
1	D	98	MSE
1	D	158	GLU
1	D	161	SER
1	D	248	ARG
1	D	259	ASP
1	E	98	MSE
1	E	194	GLU
1	E	207	ILE
1	E	215	SER
1	E	237	ILE
1	E	238	ASP
1	E	259	ASP
1	E	280	THR
1	F	98	MSE
1	F	106	LEU
1	F	130	SER
1	F	131	MSE
1	F	153	GLN
1	F	158	GLU
1	F	159	THR
1	F	170	MSE
1	F	183	GLN
1	F	189	ARG
1	F	212	ARG
1	F	266	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	284	TRP
1	G	98	MSE
1	G	105	LEU
1	G	158	GLU
1	G	212	ARG
1	G	229	ARG
1	G	243	ILE
1	H	98	MSE
1	H	106	LEU
1	H	108	ARG
1	H	165	VAL
1	H	237	ILE
1	H	259	ASP
1	H	263	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	90	ASN
1	E	103	ASN
1	F	294	ASN
1	G	112	GLN
1	G	119	ASN
1	G	193	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](#)

9 ligands are 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	$\# Z  > 2$
3	FMT	B	302	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	B	301	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	D	302	-	0,2,2	0.00	-	0,1,1	0.00	-
2	SO4	D	301	-	4,4,4	0.14	0	6,6,6	0.04	0
3	FMT	G	302	-	0,2,2	0.00	-	0,1,1	0.00	-
2	SO4	H	301	-	4,4,4	0.14	0	6,6,6	0.06	0
2	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.06	0
3	FMT	A	302	-	0,2,2	0.00	-	0,1,1	0.00	-
3	FMT	G	301	-	0,2,2	0.00	-	0,1,1	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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	301	FMT	1	0

## 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, 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	200/210 (95%)	-0.02	4 (2%) 65 68	19, 38, 66, 83	0
1	B	201/210 (95%)	0.37	7 (3%) 44 47	25, 44, 78, 112	0
1	C	200/210 (95%)	0.08	5 (2%) 57 61	25, 42, 72, 94	0
1	D	202/210 (96%)	0.20	4 (1%) 65 68	25, 47, 75, 92	0
1	E	200/210 (95%)	0.40	6 (3%) 50 53	25, 55, 90, 99	0
1	F	200/210 (95%)	0.15	7 (3%) 44 47	28, 44, 72, 85	0
1	G	200/210 (95%)	-0.05	0 100 100	21, 39, 66, 79	0
1	H	201/210 (95%)	0.15	9 (4%) 33 36	24, 41, 77, 97	0
All	All	1604/1680 (95%)	0.16	42 (2%) 56 59	19, 43, 76, 112	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	298	GLU	4.7
1	H	210	ILE	3.9
1	E	290	LEU	3.9
1	B	91	ALA	3.7
1	C	257	PHE	3.5
1	B	153	GLN	3.3
1	H	211	GLN	3.3
1	H	153	GLN	3.2
1	C	208	GLY	3.2
1	H	152	ARG	3.2
1	E	257	PHE	3.0
1	H	209	GLY	3.0
1	B	257	PHE	3.0
1	B	115	ARG	2.9
1	E	292	LYS	2.8
1	H	212	ARG	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	H	208	GLY	2.7
1	H	215	SER	2.7
1	A	90	ASN	2.7
1	D	216	ILE	2.7
1	B	90	ASN	2.6
1	F	162	LEU	2.6
1	F	257	PHE	2.6
1	E	216	ILE	2.6
1	A	213	ARG	2.5
1	A	216	ILE	2.5
1	D	271	ALA	2.5
1	B	114	ILE	2.5
1	B	92	THR	2.3
1	A	272	PRO	2.3
1	F	243	ILE	2.3
1	F	211	GLN	2.3
1	C	269	ILE	2.2
1	E	244	GLU	2.2
1	F	153	GLN	2.2
1	H	262[A]	PHE	2.2
1	F	213	ARG	2.2
1	C	152	ARG	2.2
1	C	298	GLU	2.2
1	D	211	GLN	2.1
1	F	244	GLU	2.1
1	E	110	LEU	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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
3	FMT	A	302	3/3	0.71	0.31	62,62,62,62	0
3	FMT	B	302	3/3	0.72	0.21	63,63,63,63	0
3	FMT	B	301	3/3	0.76	0.15	47,47,48,48	0
3	FMT	D	302	3/3	0.79	0.21	61,61,61,61	0
2	SO4	H	301	5/5	0.81	0.20	101,101,101,101	0
3	FMT	G	301	3/3	0.82	0.22	70,70,70,70	0
2	SO4	D	301	5/5	0.83	0.27	109,110,110,110	0
3	FMT	G	302	3/3	0.90	0.14	59,59,59,59	0
2	SO4	A	301	5/5	0.92	0.16	75,75,75,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.