



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 21, 2020 – 04:53 AM BST

PDB ID : 6CQR  
Title : Crystal structure of F24 TCR -DR1-RQ13 peptide complex  
Authors : Farenc, C.; Gras, S.; Rossjohn, J.  
Deposited on : 2018-03-16  
Resolution : 3.04 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 : 2.13.1  
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) : 2.13.1

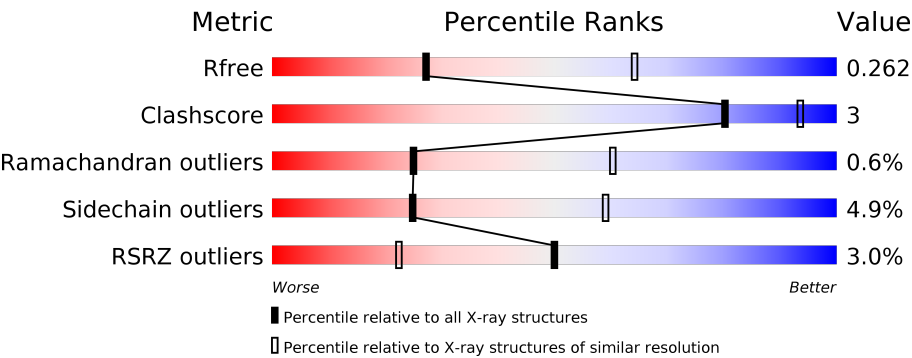
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.04 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	2752 (3.08-3.00)
Clashscore	141614	3096 (3.08-3.00)
Ramachandran outliers	138981	2986 (3.08-3.00)
Sidechain outliers	138945	2988 (3.08-3.00)
RSRZ outliers	127900	2636 (3.08-3.00)

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	182	<div><div>%</div><div><div></div><div>87%</div><div>12%</div><div>.</div></div></div>
1	F	182	<div><div>2%</div><div><div></div><div>91%</div><div>7%</div><div>.</div></div></div>
2	B	190	<div><div></div><div><div></div><div>84%</div><div>15%</div><div>.</div></div></div>
2	G	190	<div><div>17%</div><div><div></div><div>84%</div><div>16%</div></div></div>
3	C	13	<div><div></div><div><div></div><div>92%</div><div>8%</div></div></div>
3	H	13	<div><div></div><div><div></div><div>92%</div><div>8%</div></div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	205	<div><div></div><div>86%11%</div><div></div></div>
4	I	205	<div><div>3%</div><div></div><div>90%7%</div><div></div></div>
5	E	245	<div><div></div><div>88%12%</div><div></div></div>
5	J	245	<div><div>2%</div><div></div><div>88%11%</div><div></div></div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13648 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	179	Total	C	N	O	S	0	0	0
			1473	954	239	275	5			
1	F	179	Total	C	N	O	S	0	0	0
			1473	954	239	275	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	THR	ALA	conflict	UNP P01903
F	182	THR	ALA	conflict	UNP P01903

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	190	Total	C	N	O	S	0	0	0
			1554	978	277	293	6			
2	G	190	Total	C	N	O	S	0	0	0
			1551	977	275	293	6			

- Molecule 3 is a protein called Peptide from Capsid protein p24.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	13	Total	C	N	O	0	0	0
			113	70	22	21			
3	H	13	Total	C	N	O	0	0	0
			113	70	22	21			

- Molecule 4 is a protein called F24 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	200	Total	C	N	O	S	0	1	0
			1569	986	259	316	8			

*Continued on next page...*

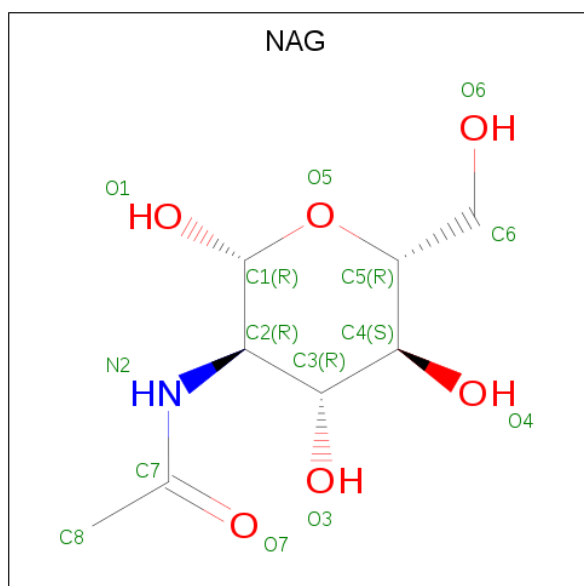
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	I	201	Total	C	N	O	S	0	1	0
			1577	992	260	317	8			

- Molecule 5 is a protein called F24 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	245	Total	C	N	O	S	0	0	0
			1981	1255	337	381	8			
5	J	245	Total	C	N	O	S	0	0	0
			1981	1255	337	381	8			

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	F	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	63	Total 63 O 63	0	0

*Continued on next page...*


*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	35	Total 35	O 35	0	0
7	C	4	Total 4	O 4	0	0
7	D	24	Total 24	O 24	0	0
7	E	28	Total 28	O 28	0	0
7	F	18	Total 18	O 18	0	0
7	G	17	Total 17	O 17	0	0
7	H	2	Total 2	O 2	0	0
7	I	25	Total 25	O 25	0	0
7	J	19	Total 19	O 19	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.

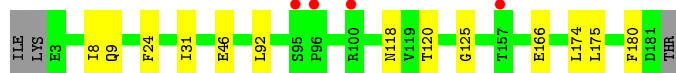
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain A: 




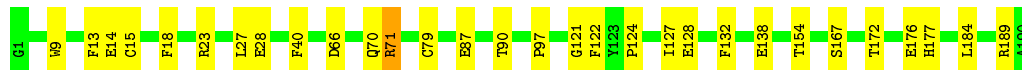
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain F: 




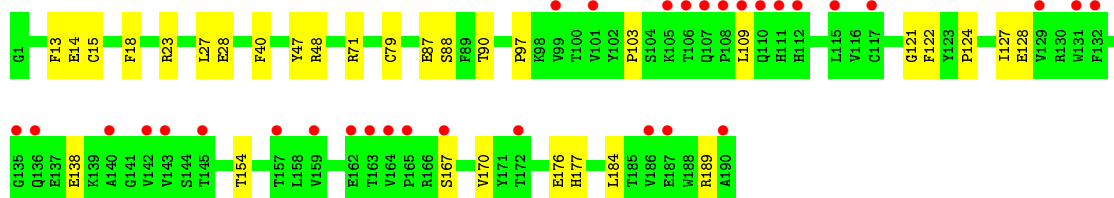
- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain B: 



- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain G: 



- Molecule 3: Peptide from Capsid protein p24

Chain C: 



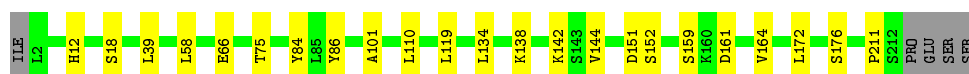
- Molecule 3: Peptide from Capsid protein p24

Chain H: 92% 8%



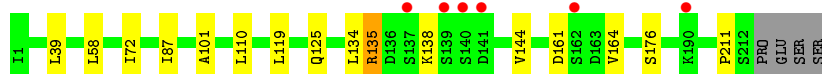
- Molecule 4: F24 alpha chain

Chain D: 86% 11%



- Molecule 4: F24 alpha chain

Chain I: 3% 90% 7%



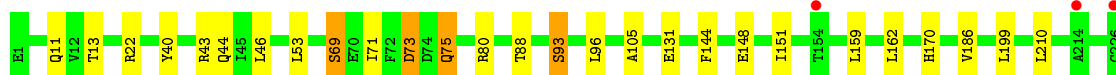
- Molecule 5: F24 beta chain

Chain E: 88% 12%



- Molecule 5: F24 beta chain

Chain J: 2% 88% 11%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	133.93Å 155.32Å 133.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.29 – 3.04 47.39 – 3.04	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.29-3.04) 100.0 (47.39-3.04)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.16 (at 3.07Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.223 , 0.254 0.235 , 0.262	Depositor DCC
$R_{free}$ test set	2796 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	64.8	Xtriage
Anisotropy	0.730	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 62.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.012 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	13648	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.55% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.38	0/1518	0.60	0/2070
1	F	0.39	0/1518	0.60	0/2070
2	B	0.35	0/1594	0.58	0/2165
2	G	0.43	0/1591	0.56	0/2161
3	C	0.33	0/114	0.62	0/149
3	H	0.34	0/114	0.64	0/149
4	D	0.37	0/1609	0.58	0/2180
4	I	0.39	0/1617	0.60	0/2191
5	E	0.35	0/2035	0.58	0/2766
5	J	0.35	0/2035	0.58	0/2766
All	All	0.37	0/13745	0.59	0/18667

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

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	1473	0	1408	7	0
1	F	1473	0	1408	4	0
2	B	1554	0	1478	16	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	G	1551	0	1473	16	0
3	C	113	0	111	1	0
3	H	113	0	111	1	0
4	D	1569	0	1492	8	0
4	I	1577	0	1506	5	0
5	E	1981	0	1880	8	0
5	J	1981	0	1880	13	0
6	A	14	0	13	0	0
6	F	14	0	13	0	0
7	A	63	0	0	0	0
7	B	35	0	0	0	0
7	C	4	0	0	0	0
7	D	24	0	0	0	0
7	E	28	0	0	0	0
7	F	18	0	0	0	0
7	G	17	0	0	0	0
7	H	2	0	0	0	0
7	I	25	0	0	0	0
7	J	19	0	0	0	0
All	All	13648	0	12773	69	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 (69) 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)
2:G:48:ARG:HH11	2:G:48:ARG:HG3	1.39	0.86
1:A:107:CYS:HG	1:A:163:CYS:HG	1.20	0.86
2:G:28:GLU:HB3	2:G:40:PHE:HB3	1.72	0.70
1:F:118:ASN:HB2	1:F:166:GLU:HB2	1.73	0.70
2:B:28:GLU:HB3	2:B:40:PHE:HB3	1.71	0.70
2:G:48:ARG:HG3	2:G:48:ARG:NH1	2.02	0.70
1:A:118:ASN:HB2	1:A:166:GLU:HB2	1.74	0.69
2:B:66:ASP:O	2:B:70:GLN:HG2	1.98	0.63
4:D:159:SER:HB2	4:D:164:VAL:HG13	1.82	0.61
1:F:8:ILE:HG12	2:G:14:GLU:HG2	1.84	0.60
4:D:12:HIS:CE1	4:D:119:LEU:HD23	2.39	0.58
5:J:69:SER:HB3	5:J:71:ILE:HG22	1.86	0.57
5:J:75:GLN:HG3	5:J:93:SER:H	1.68	0.57
2:B:13:PHE:CD1	3:C:94:LEU:HD23	2.41	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:I:161:ASP:HB3	4:I:164:VAL:HG12	1.88	0.55
4:I:72:ILE:HG12	4:I:87:ILE:HG12	1.91	0.53
1:A:9:GLN:HB3	2:B:13:PHE:HB2	1.90	0.53
1:A:8:ILE:HG12	2:B:14:GLU:HG2	1.91	0.53
5:E:186:VAL:HG22	5:E:210:LEU:HD12	1.92	0.52
2:B:15:CYS:SG	2:B:79:CYS:SG	3.03	0.52
5:J:186:VAL:HG22	5:J:210:LEU:HD12	1.92	0.51
4:D:161:ASP:HB3	4:D:164:VAL:HG12	1.93	0.51
2:B:18:PHE:HB2	2:B:23:ARG:HB3	1.93	0.50
2:G:18:PHE:HB2	2:G:23:ARG:HB3	1.93	0.50
5:E:29:HIS:HE1	5:E:120:PHE:CD2	2.30	0.49
1:A:24:PHE:HB3	1:A:31:ILE:HD12	1.93	0.49
1:F:24:PHE:HB3	1:F:31:ILE:HD12	1.94	0.49
2:B:132:PHE:HB2	2:B:172:THR:HG23	1.95	0.48
2:G:15:CYS:SG	2:G:79:CYS:SG	3.07	0.48
5:J:53:LEU:HD23	5:J:71:ILE:HG23	1.95	0.48
4:D:134:LEU:HB3	5:E:144:PHE:HB3	1.96	0.48
5:J:159:LEU:HD13	5:J:210:LEU:HD23	1.95	0.48
1:F:9:GLN:HB3	2:G:13:PHE:HB2	1.94	0.48
4:D:101:ALA:HB1	4:D:110:LEU:HB3	1.96	0.47
4:D:66:GLU:HG3	4:D:75:THR:HG22	1.95	0.47
4:I:135:ARG:H	4:I:135:ARG:HD2	1.79	0.47
2:B:97:PRO:HB3	2:B:122:PHE:HB3	1.96	0.47
5:E:159:LEU:HD13	5:E:210:LEU:HD23	1.96	0.47
5:J:22:ARG:HA	5:J:88:THR:HG22	1.97	0.47
4:I:134:LEU:HB3	5:J:144:PHE:HB3	1.97	0.47
4:D:172:LEU:HB3	5:E:187:CYS:SG	2.55	0.46
4:I:101:ALA:HB1	4:I:110:LEU:HB3	1.96	0.46
2:B:124:PRO:O	2:B:177:HIS:HE1	1.99	0.45
2:G:97:PRO:HB3	2:G:122:PHE:HB3	1.97	0.45
2:B:28:GLU:CD	2:B:71:ARG:HH11	2.20	0.45
2:G:170:VAL:HB	2:G:189:ARG:HH21	1.82	0.45
2:G:128:GLU:HB2	2:G:176:GLU:HB2	1.99	0.45
5:E:22:ARG:HA	5:E:88:THR:HG22	1.99	0.45
5:J:148:GLU:HA	5:J:151:ILE:HD12	1.99	0.45
5:E:148:GLU:HA	5:E:151:ILE:HD12	1.99	0.44
5:E:40:TYR:HB2	5:E:105:ALA:HB3	1.98	0.44
2:G:14:GLU:HB2	2:G:27:LEU:HB3	2.00	0.44
2:G:121:GLY:HA2	2:G:154:THR:HB	2.00	0.44
5:J:40:TYR:HB2	5:J:105:ALA:HB3	1.99	0.44
2:G:124:PRO:O	2:G:177:HIS:HE1	1.99	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:128:GLU:HB2	2:B:176:GLU:HB2	1.99	0.44
2:B:121:GLY:HA2	2:B:154:THR:HB	1.99	0.44
2:B:28:GLU:OE2	2:B:71:ARG:NH1	2.42	0.44
5:J:73:ASP:HB3	5:J:75:GLN:H	1.82	0.43
4:D:18:SER:HB3	4:D:86:TYR:CE1	2.53	0.43
2:B:14:GLU:HB2	2:B:27:LEU:HB3	2.00	0.43
5:J:96:LEU:HD11	5:J:131:GLU:HG3	1.99	0.43
5:J:43:ARG:NH2	5:J:71:ILE:HD11	2.34	0.42
2:G:28:GLU:OE1	2:G:47:TYR:OH	2.37	0.42
2:G:13:PHE:CD1	3:H:94:LEU:HD23	2.55	0.41
5:J:170:HIS:HB3	5:J:231:TYR:HB2	2.02	0.41
1:A:73:MET:HG3	2:B:9:TRP:CZ3	2.55	0.41
1:A:95:SER:HB2	1:A:96:PRO:HD2	2.03	0.41
2:G:109:LEU:HD11	2:G:167:SER:HA	2.02	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	177/182 (97%)	176 (99%)	1 (1%)	0	100	100
1	F	177/182 (97%)	169 (96%)	7 (4%)	1 (1%)	25	60
2	B	188/190 (99%)	177 (94%)	10 (5%)	1 (0%)	29	65
2	G	188/190 (99%)	180 (96%)	7 (4%)	1 (0%)	29	65
3	C	11/13 (85%)	11 (100%)	0	0	100	100
3	H	11/13 (85%)	11 (100%)	0	0	100	100
4	D	199/205 (97%)	185 (93%)	12 (6%)	2 (1%)	15	49
4	I	200/205 (98%)	184 (92%)	14 (7%)	2 (1%)	15	49
5	E	243/245 (99%)	236 (97%)	6 (2%)	1 (0%)	34	69

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	J	243/245 (99%)	237 (98%)	5 (2%)	1 (0%)	34 69
All	All	1637/1670 (98%)	1566 (96%)	62 (4%)	9 (0%)	25 65

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	211	PRO
4	I	211	PRO
5	J	73	ASP
5	E	81	PRO
2	B	87	GLU
4	D	176	SER
4	I	176	SER
2	G	87	GLU
1	F	125	GLY

### 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	164/167 (98%)	154 (94%)	10 (6%)	18 50
1	F	164/167 (98%)	158 (96%)	6 (4%)	34 68
2	B	169/171 (99%)	162 (96%)	7 (4%)	30 65
2	G	168/171 (98%)	161 (96%)	7 (4%)	30 64
3	C	11/11 (100%)	11 (100%)	0	100 100
3	H	11/11 (100%)	11 (100%)	0	100 100
4	D	178/182 (98%)	170 (96%)	8 (4%)	27 62
4	I	179/182 (98%)	172 (96%)	7 (4%)	32 66
5	E	219/219 (100%)	204 (93%)	15 (7%)	16 46
5	J	219/219 (100%)	207 (94%)	12 (6%)	21 54
All	All	1482/1500 (99%)	1410 (95%)	72 (5%)	25 59

All (72) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	46	GLU
1	A	92	LEU
1	A	100	ARG
1	A	105	LEU
1	A	128	VAL
1	A	154	LEU
1	A	170	LEU
1	A	171	ASP
1	A	172	GLU
1	A	175	LEU
2	B	71	ARG
2	B	90	THR
2	B	127	ILE
2	B	138	GLU
2	B	167	SER
2	B	184	LEU
2	B	189	ARG
4	D	39	LEU
4	D	58	LEU
4	D	84	TYR
4	D	138	LYS
4	D	142	LYS
4	D	144	VAL
4	D	151	ASP
4	D	152	SER
5	E	1	GLU
5	E	11	GLN
5	E	13	THR
5	E	44	GLN
5	E	46	LEU
5	E	51	GLU
5	E	71	ILE
5	E	80	ARG
5	E	113	MET
5	E	162	LEU
5	E	170	HIS
5	E	199	LEU
5	E	241	GLN
5	E	258	ARG
5	E	260	ASP
1	F	46	GLU
1	F	92	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	120	THR
1	F	174	LEU
1	F	175	LEU
1	F	180	PHE
2	G	71	ARG
2	G	88	SER
2	G	90	THR
2	G	103	PRO
2	G	127	ILE
2	G	138	GLU
2	G	184	LEU
4	I	39	LEU
4	I	58	LEU
4	I	119	LEU
4	I	125	GLN
4	I	135	ARG
4	I	138	LYS
4	I	144	VAL
5	J	11	GLN
5	J	13	THR
5	J	44	GLN
5	J	46	LEU
5	J	69	SER
5	J	75	GLN
5	J	80	ARG
5	J	93	SER
5	J	162	LEU
5	J	199	LEU
5	J	235	GLU
5	J	241	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	124	ASN
2	B	111	HIS
2	B	136	GLN
2	B	177	HIS
4	D	12	HIS
5	E	48	GLN
5	E	153	HIS
5	E	170	HIS

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	G	113	ASN
2	G	136	GLN
2	G	174	GLN
2	G	177	HIS
3	H	101	GLN
4	I	201	ASN
5	J	48	GLN
5	J	222	ASN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

2 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	$\# Z  > 2$
6	NAG	F	201	1	14,14,15	0.30	0	17,19,21	0.61	0
6	NAG	A	201	1	14,14,15	0.28	0	17,19,21	0.67	0

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
6	NAG	F	201	1	-	1/6/23/26	0/1/1/1
6	NAG	A	201	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	F	201	NAG	O5-C5-C6-O6

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, 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	179/182 (98%)	-0.31	1 (0%) 89 72	43, 60, 88, 119	0
1	F	179/182 (98%)	0.20	4 (2%) 62 33	56, 98, 144, 166	0
2	B	190/190 (100%)	-0.15	0 100 100	35, 62, 91, 106	0
2	G	190/190 (100%)	0.85	32 (16%) 1 0	44, 113, 164, 178	0
3	C	13/13 (100%)	-0.20	0 100 100	45, 54, 73, 91	0
3	H	13/13 (100%)	-0.22	0 100 100	54, 60, 80, 96	0
4	D	200/205 (97%)	-0.12	0 100 100	49, 68, 102, 116	1 (0%)
4	I	201/205 (98%)	0.02	6 (2%) 50 22	44, 84, 131, 146	1 (0%)
5	E	245/245 (100%)	-0.19	1 (0%) 92 79	42, 70, 99, 116	4 (1%)
5	J	245/245 (100%)	0.16	6 (2%) 59 30	47, 95, 132, 143	4 (1%)
All	All	1655/1670 (99%)	0.05	50 (3%) 50 22	35, 74, 140, 178	10 (0%)

All (50) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	111	HIS	7.6
2	G	106	THR	6.8
2	G	110	GLN	6.2
2	G	108	PRO	5.6
4	I	141	ASP	4.9
4	I	137	SER	4.8
2	G	164	VAL	4.3
2	G	109	LEU	3.9
2	G	107	GLN	3.9
2	G	159	VAL	3.6
4	I	140	SER	3.6
1	F	95	SER	3.6
2	G	165	PRO	3.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	G	163	THR	3.5
2	G	136	GLN	3.5
2	G	112	HIS	3.5
2	G	186	VAL	3.2
2	G	132	PHE	3.1
2	G	99	VAL	3.1
2	G	101	VAL	3.0
5	J	154	THR	3.0
2	G	129	VAL	2.8
2	G	135	GLY	2.8
2	G	187	GLU	2.7
2	G	105	LYS	2.7
2	G	162	GLU	2.7
5	J	228	VAL	2.7
1	F	100	ARG	2.6
2	G	172	THR	2.6
2	G	145	THR	2.5
4	I	139	SER	2.5
5	E	236	ASN	2.5
2	G	143	VAL	2.4
1	F	157	THR	2.4
5	J	238	GLU	2.4
2	G	115	LEU	2.4
5	J	214	ALA	2.4
2	G	167	SER	2.4
4	I	190	LYS	2.4
4	I	162	SER	2.4
1	F	96	PRO	2.3
1	A	158	GLU	2.3
2	G	131	TRP	2.2
5	J	247	VAL	2.1
2	G	190	ALA	2.1
5	J	226	CYS	2.1
2	G	142	VAL	2.0
2	G	140	ALA	2.0
2	G	157	THR	2.0
2	G	117	CYS	2.0

## 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 [i](#)

There are no monosaccharides 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
6	NAG	F	201	14/15	0.88	0.13	121,122,123,123	0
6	NAG	A	201	14/15	0.94	0.20	75,76,78,78	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.