



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2016 – 05:52 AM EDT

PDB ID : 5KS9  
Title : Bel502-DQ8-glia-alpha1 complex  
Authors : Petersen, J.; Rossjohn, J.; Reid, H.H.  
Deposited on : 2016-07-08  
Resolution : 2.55 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:

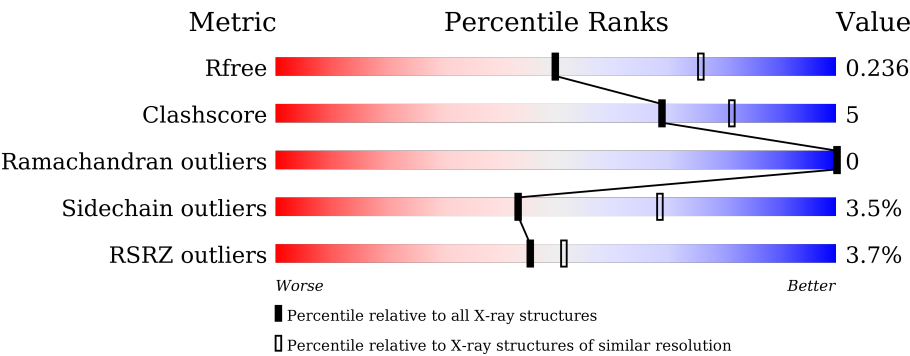
MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027939

# 1 Overall quality at a glance i

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

The reported resolution of this entry is 2.55 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	4549 (2.58-2.50)
Clashscore	102246	5292 (2.58-2.50)
Ramachandran outliers	100387	5194 (2.58-2.50)
Sidechain outliers	100360	5196 (2.58-2.50)
RSRZ outliers	91569	4561 (2.58-2.50)

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. 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	192	<div><div>2%</div><div></div><div>84%</div><div>10%</div><div>5%</div></div>
1	C	192	<div><div>3%</div><div></div><div>84%</div><div>11%</div><div>5%</div></div>
2	B	230	<div><div>3%</div><div></div><div>67%</div><div>11%</div><div>21%</div></div>
2	D	230	<div><div>4%</div><div></div><div>63%</div><div>11%</div><div>25%</div></div>
3	E	207	<div><div>4%</div><div></div><div>79%</div><div>13%</div><div>7%</div></div>
3	G	207	<div><div>6%</div><div></div><div>78%</div><div>14%</div><div>7%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	242	<div><div>%</div><div><div></div><div>91%</div><div>7% ..</div></div></div>
4	H	242	<div><div>3%</div><div><div></div><div>89%</div><div>9% ..</div></div></div>
5	I	16	<div><div>6%</div><div><div></div><div>88%</div><div>13%</div></div></div>
5	J	16	<div><div>6%</div><div><div></div><div>100%</div><div></div></div></div>

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13515 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, DQ alpha 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	182	Total	C	N	O	S	0	0	0
			1464	944	240	278	2			
1	C	182	Total	C	N	O	S	0	0	0
			1464	944	240	278	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	THR	-	expression tag	UNP Q30063
A	183	SER	-	expression tag	UNP Q30063
A	184	GLY	-	expression tag	UNP Q30063
A	185	ASP	-	expression tag	UNP Q30063
A	186	ASP	-	expression tag	UNP Q30063
A	187	ASP	-	expression tag	UNP Q30063
A	188	ASP	-	expression tag	UNP Q30063
A	189	LYS	-	expression tag	UNP Q30063
C	182	THR	-	expression tag	UNP Q30063
C	183	SER	-	expression tag	UNP Q30063
C	184	GLY	-	expression tag	UNP Q30063
C	185	ASP	-	expression tag	UNP Q30063
C	186	ASP	-	expression tag	UNP Q30063
C	187	ASP	-	expression tag	UNP Q30063
C	188	ASP	-	expression tag	UNP Q30063
C	189	LYS	-	expression tag	UNP Q30063

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	181	Total	C	N	O	S	0	1	0
			1493	946	263	277	7			
2	D	173	Total	C	N	O	S	0	0	0
			1395	890	244	254	7			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-13	SER	-	linker	PDB ?
B	-12	GLY	-	linker	PDB ?
B	-11	GLY	-	linker	PDB ?
B	-10	SER	-	linker	PDB ?
B	-9	ILE	-	linker	PDB ?
B	-8	GLU	-	linker	PDB ?
B	-7	GLY	-	linker	PDB ?
B	-6	ARG	-	linker	PDB ?
B	-5	GLY	-	linker	PDB ?
B	-4	GLY	-	linker	PDB ?
B	-3	SER	-	linker	PDB ?
B	-2	GLY	-	linker	PDB ?
B	-1	ALA	-	linker	PDB ?
B	0	SER	-	linker	PDB ?
B	193	THR	-	expression tag	UNP U3PYM0
B	194	GLY	-	expression tag	UNP U3PYM0
B	195	GLY	-	expression tag	UNP U3PYM0
B	196	ASP	-	expression tag	UNP U3PYM0
B	197	ASP	-	expression tag	UNP U3PYM0
B	198	ASP	-	expression tag	UNP U3PYM0
B	199	ASP	-	expression tag	UNP U3PYM0
B	200	LYS	-	expression tag	UNP U3PYM0
D	-13	SER	-	linker	PDB ?
D	-12	GLY	-	linker	PDB ?
D	-11	GLY	-	linker	PDB ?
D	-10	SER	-	linker	PDB ?
D	-9	ILE	-	linker	PDB ?
D	-8	GLU	-	linker	PDB ?
D	-7	GLY	-	linker	PDB ?
D	-6	ARG	-	linker	PDB ?
D	-5	GLY	-	linker	PDB ?
D	-4	GLY	-	linker	PDB ?
D	-3	SER	-	linker	PDB ?
D	-2	GLY	-	linker	PDB ?
D	-1	ALA	-	linker	PDB ?
D	0	SER	-	linker	PDB ?
D	193	THR	-	expression tag	UNP U3PYM0
D	194	GLY	-	expression tag	UNP U3PYM0
D	195	GLY	-	expression tag	UNP U3PYM0
D	196	ASP	-	expression tag	UNP U3PYM0
D	197	ASP	-	expression tag	UNP U3PYM0
D	198	ASP	-	expression tag	UNP U3PYM0

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	199	ASP	-	expression tag	UNP U3PYM0
D	200	LYS	-	expression tag	UNP U3PYM0

- Molecule 3 is a protein called Bel502 TCR alpha TRAV20\*01.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	193	Total	C	N	O	S	0	0	0
			1484	929	252	294	9			
3	G	192	Total	C	N	O	S	0	0	0
			1466	918	247	292	9			

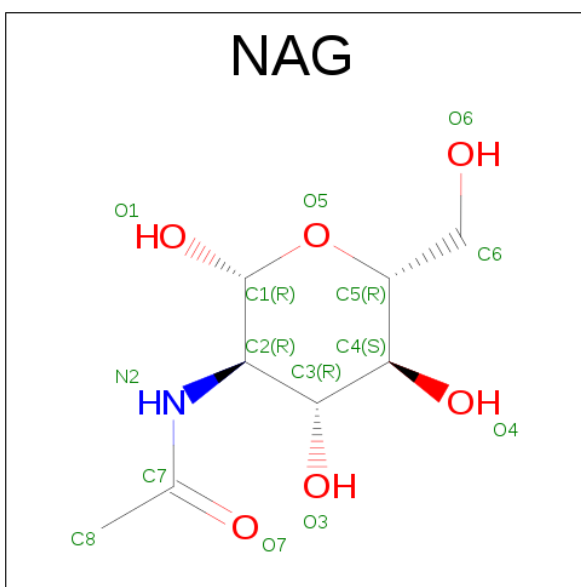
- Molecule 4 is a protein called Bel502 TCR beta TRBV9\*01.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	240	Total	C	N	O	S	0	0	0
			1887	1193	327	362	5			
4	H	240	Total	C	N	O	S	0	0	0
			1886	1193	330	358	5			

- Molecule 5 is a protein called DQ8-glia-alpha1 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	16	Total	C	N	O	0	0	0
			112	67	19	26			
5	J	16	Total	C	N	O	0	0	0
			112	67	19	26			

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	G	1	Total	Ca	0	0
			1	1		
7	A	1	Total	Ca	0	0
			1	1		
7	C	1	Total	Ca	0	0
			1	1		
7	E	1	Total	Ca	0	0
			1	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	79	Total	O	0	0
			79	79		

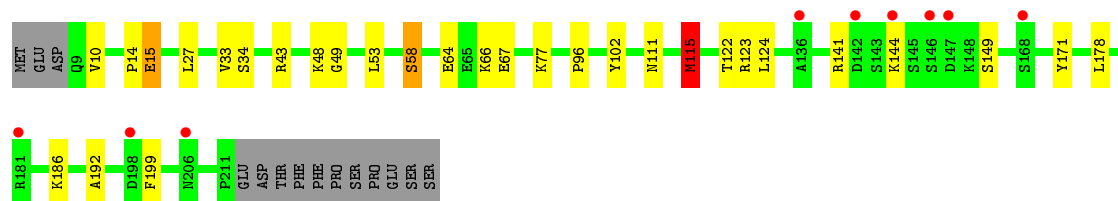
*Continued on next page...*

*Continued from previous page...*

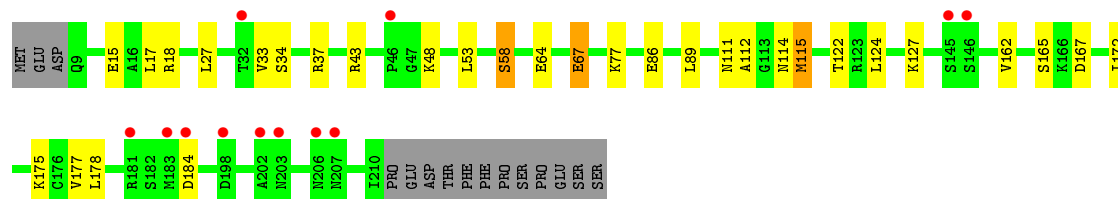
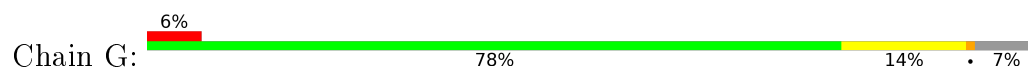
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	60	Total 60	O 60	0	0
8	C	105	Total 105	O 105	0	0
8	D	90	Total 90	O 90	0	0
8	E	75	Total 75	O 75	0	0
8	F	145	Total 145	O 145	0	0
8	G	44	Total 44	O 44	0	0
8	H	81	Total 81	O 81	0	0
8	I	9	Total 9	O 9	0	0
8	J	4	Total 4	O 4	0	0



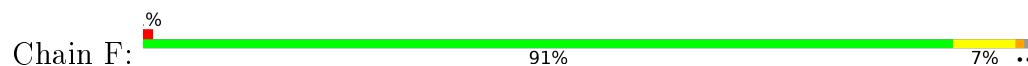




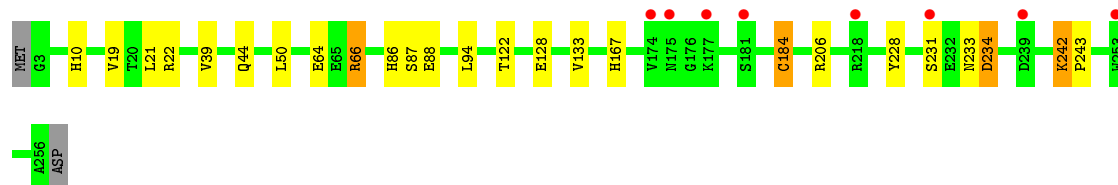
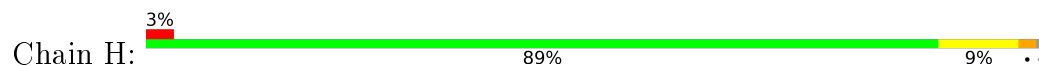
- Molecule 3: Bel502 TCR alpha TRAV20\*01



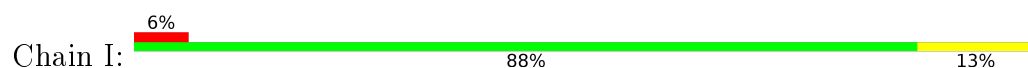
- Molecule 4: Bel502 TCR beta TRBV9\*01



- Molecule 4: Bel502 TCR beta TRBV9\*01



- Molecule 5: DQ8-glia-alpha1 peptide



- Molecule 5: DQ8-glia-alpha1 peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.56Å 56.87Å 232.05Å 90.00° 92.77° 90.00°	Depositor
Resolution (Å)	46.90 – 2.55 46.84 – 2.55	Depositor EDS
% Data completeness (in resolution range)	98.8 (46.90-2.55) 98.8 (46.84-2.55)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.54Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.199 , 0.224 0.214 , 0.236	Depositor DCC
$R_{free}$ test set	1263 reflections (2.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.1	Xtriage
Anisotropy	0.548	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 57.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.025 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13515	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 40.45 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.7305e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: CA, 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	0.66	1/1508 (0.1%)	0.66	0/2061
1	C	0.83	3/1508 (0.2%)	0.73	1/2061 (0.0%)
2	B	0.82	3/1534 (0.2%)	0.79	3/2092 (0.1%)
2	D	0.79	0/1430	0.72	0/1951
3	E	0.76	2/1514 (0.1%)	0.69	0/2050
3	G	0.60	0/1495	0.64	0/2026
4	F	0.60	0/1938	0.66	0/2644
4	H	0.56	0/1937	0.64	1/2642 (0.0%)
5	I	0.43	0/115	0.63	0/156
5	J	0.40	0/115	0.57	0/156
All	All	0.70	9/13094 (0.1%)	0.69	5/17839 (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
2	B	0	1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	174[A]	HIS	CA-C	5.81	1.68	1.52
2	B	174[B]	HIS	CA-C	5.81	1.68	1.52
1	C	31	GLU	CD-OE1	-5.79	1.19	1.25
1	A	114	PRO	N-CD	5.49	1.55	1.47
3	E	14	PRO	N-CD	5.48	1.55	1.47
3	E	115	MET	SD-CE	-5.32	1.48	1.77
1	C	31	GLU	CD-OE2	-5.20	1.20	1.25
2	B	165	PRO	N-CD	5.12	1.55	1.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	18	PRO	N-CD	5.01	1.54	1.47

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	173	CYS	CA-CB-SG	7.79	128.02	114.00
1	C	142	ASP	CB-CG-OD1	6.09	123.78	118.30
4	H	242	LYS	C-N-CD	5.51	139.98	128.40
2	B	96	GLU	C-N-CD	5.43	139.80	128.40
2	B	164	THR	C-N-CD	5.41	139.77	128.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	174[B]	HIS	Mainchain

## 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	1464	0	1395	15	0
1	C	1464	0	1395	12	0
2	B	1493	0	1426	20	1
2	D	1395	0	1309	18	0
3	E	1484	0	1420	18	0
3	G	1466	0	1388	18	0
4	F	1887	0	1784	11	1
4	H	1886	0	1789	15	0
5	I	112	0	94	2	0
5	J	112	0	94	0	0
6	A	14	0	13	0	0
6	B	14	0	13	0	0
6	C	14	0	13	0	0
6	D	14	0	13	0	0
7	A	1	0	0	0	0
7	C	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	E	1	0	0	0	0
7	G	1	0	0	0	0
8	A	79	0	0	0	0
8	B	60	0	0	0	0
8	C	105	0	0	0	0
8	D	90	0	0	0	0
8	E	75	0	0	0	0
8	F	145	0	0	0	0
8	G	44	0	0	0	0
8	H	81	0	0	0	0
8	I	9	0	0	0	0
8	J	4	0	0	0	0
All	All	13515	0	12146	115	1

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 (115) 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:B:132:PHE:CD1	2:B:174[B]:HIS:CD2	2.17	1.33
2:B:117:CYS:SG	2:B:173:CYS:SG	1.36	1.22
3:E:15:GLU:OE1	3:E:123:ARG:NH2	1.86	1.07
4:H:234:ASP:O	4:H:242:LYS:NZ	1.88	1.07
2:B:132:PHE:CE1	2:B:174[B]:HIS:CD2	2.48	1.02
2:B:132:PHE:CD1	2:B:174[B]:HIS:HD2	1.68	0.97
2:B:132:PHE:CD1	2:B:174[B]:HIS:NE2	2.35	0.95
2:B:132:PHE:HD1	2:B:174[B]:HIS:CD2	1.70	0.95
2:B:117:CYS:CB	2:B:173:CYS:SG	2.54	0.94
4:H:231:SER:HB2	4:H:233:ASN:OD1	1.76	0.85
3:E:111:ASN:HB3	3:E:115:MET:CE	2.11	0.81
2:B:132:PHE:HD1	2:B:174[B]:HIS:HD2	1.11	0.79
3:G:111:ASN:HB3	3:G:115:MET:HE3	1.64	0.78
2:B:172:THR:HG21	2:B:174[B]:HIS:CE1	2.20	0.76
3:G:178:LEU:HB3	4:H:184:CYS:HB3	1.66	0.76
2:D:139:THR:O	2:D:142:VAL:HG12	1.86	0.74
3:E:67:GLU:HG3	3:E:77:LYS:HG3	1.71	0.73
3:G:33:VAL:HG12	3:G:34:SER:N	2.03	0.73
3:G:33:VAL:HG12	3:G:34:SER:H	1.54	0.72
1:A:95:SER:HB2	1:A:96:PRO:HD2	1.70	0.71
2:B:172:THR:CG2	2:B:174[B]:HIS:CE1	2.73	0.71

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:G:111:ASN:HB3	3:G:115:MET:CE	2.20	0.70
3:G:58:SER:O	3:G:64:GLU:HB2	1.93	0.69
4:H:21:LEU:HD22	4:H:122:THR:HG21	1.74	0.68
3:G:67:GLU:HG3	3:G:77:LYS:HG3	1.79	0.64
1:A:50:ARG:HG2	1:A:51:PHE:CE1	2.33	0.64
3:E:111:ASN:HB3	3:E:115:MET:HE2	1.81	0.63
3:G:67:GLU:HG3	3:G:77:LYS:CG	2.28	0.62
3:E:178:LEU:HB3	4:F:184:CYS:HB3	1.80	0.61
2:B:132:PHE:CE1	2:B:174[B]:HIS:HD2	2.05	0.60
1:C:50:ARG:HD2	1:C:51:PHE:CZ	2.36	0.60
2:D:169:ASP:N	2:D:169:ASP:OD1	2.34	0.60
1:A:50:ARG:HG2	1:A:51:PHE:CD1	2.36	0.60
1:C:48:PHE:CE1	2:D:89:THR:HB	2.37	0.60
3:E:58:SER:O	3:E:64:GLU:HB2	2.02	0.59
2:D:130:ARG:HB2	2:D:174:HIS:HB3	1.85	0.59
4:H:22:ARG:HG2	4:H:88:GLU:HG2	1.85	0.58
1:A:57:GLN:OE1	4:H:66:ARG:NH2	2.36	0.58
3:E:67:GLU:HG3	3:E:77:LYS:CG	2.34	0.58
2:D:85:LEU:HD21	5:I:-2:SER:HB3	1.86	0.57
2:B:10:GLN:HB2	2:B:31:ILE:HB	1.85	0.57
3:E:33:VAL:HG12	3:E:34:SER:N	2.20	0.57
3:G:18:ARG:HH21	3:G:127:LYS:HE2	1.71	0.56
2:D:142:VAL:O	2:D:142:VAL:HG13	2.04	0.56
2:B:133:ARG:HD3	2:B:171:TYR:CE2	2.41	0.55
3:G:165:SER:HB2	3:G:172:ILE:HG12	1.88	0.55
3:G:27:LEU:HD22	3:G:122:THR:HG21	1.88	0.54
3:E:27:LEU:HD22	3:E:122:THR:HG21	1.90	0.54
1:A:118:ASN:HB2	1:A:166:GLU:HB2	1.90	0.54
1:C:118:ASN:HB2	1:C:166:GLU:HB2	1.88	0.53
2:D:176:GLU:HG3	2:D:183:PRO:HB3	1.89	0.53
2:D:117:CYS:HB2	2:D:131:TRP:CZ2	2.43	0.53
4:H:44:GLN:HB2	4:H:50:LEU:HD13	1.91	0.53
1:C:50:ARG:HD2	1:C:51:PHE:CE2	2.44	0.53
2:B:172:THR:HG22	2:B:174[B]:HIS:CE1	2.46	0.51
1:A:50:ARG:NE	1:A:51:PHE:CE1	2.79	0.51
3:G:37:ARG:HH21	3:G:114:ASN:HD21	1.58	0.50
1:A:97:VAL:HG11	1:A:180:PRO:HB3	1.93	0.50
4:H:234:ASP:OD1	4:H:234:ASP:N	2.27	0.50
1:C:45:LEU:HB2	1:C:48:PHE:CD2	2.47	0.50
3:E:111:ASN:HB3	3:E:115:MET:HE1	1.93	0.50
1:C:48:PHE:CD1	2:D:89:THR:HB	2.47	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:85:LEU:HD13	5:I:1:GLU:OE2	2.13	0.49
4:F:231:SER:N	4:F:234:ASP:OD1	2.46	0.49
1:C:70:LEU:HD13	2:D:9:TYR:HB2	1.96	0.48
1:C:59:ALA:O	1:C:63:ILE:HG12	2.15	0.47
3:E:141:ARG:HB2	4:F:142:GLU:HB2	1.95	0.47
4:H:167:HIS:HB3	4:H:228:TYR:HB2	1.97	0.47
1:A:85:GLU:HG3	2:B:34:ARG:NH2	2.29	0.47
2:D:97:PRO:HB3	2:D:122:PHE:HB3	1.96	0.47
1:C:48:PHE:HE1	2:D:89:THR:HB	1.80	0.47
4:F:5:THR:OG1	4:F:24:SER:HB2	2.15	0.47
1:A:59:ALA:O	1:A:63:ILE:HG12	2.16	0.46
2:B:121:ASP:HA	2:B:154:THR:HB	1.97	0.46
2:B:138:GLU:HG3	2:B:161:LEU:HD11	1.96	0.46
4:F:167:HIS:HB3	4:F:228:TYR:HB2	1.97	0.46
2:D:133:ARG:O	2:D:134:ASN:HB2	2.16	0.46
4:H:39:VAL:HG21	4:H:87:SER:HB2	1.98	0.46
1:A:105:LEU:HD21	1:A:178:TRP:CD2	2.50	0.46
2:D:10:GLN:HB2	2:D:31:ILE:HB	1.98	0.45
1:C:8:SER:C	1:C:9(A):GLY:HA2	2.36	0.45
4:F:166:ASP:HB2	4:F:189:PRO:HG2	1.98	0.45
4:F:8:PRO:HG3	4:F:11:LEU:HD13	1.99	0.45
3:G:33:VAL:CG1	3:G:34:SER:N	2.73	0.45
1:A:70:LEU:HD13	2:B:9:TYR:HB2	1.98	0.45
2:D:133:ARG:HG3	2:D:171:TYR:HE1	1.80	0.45
3:E:33:VAL:CG1	3:E:34:SER:N	2.79	0.45
3:E:149:SER:HB2	3:E:199:PHE:HD2	1.81	0.44
1:A:50:ARG:HD3	1:A:51:PHE:CZ	2.53	0.44
3:G:43:ARG:HB3	3:G:53:LEU:HD11	1.99	0.44
3:G:17:LEU:HB3	3:G:124:LEU:HD12	2.00	0.44
3:E:171:TYR:O	3:E:192:ALA:HA	2.17	0.44
1:A:95:SER:HB2	1:A:96:PRO:CD	2.46	0.44
1:C:113:PHE:CG	1:C:114:PRO:HA	2.52	0.44
3:E:49:GLY:HA2	4:F:103:PHE:CE1	2.53	0.44
4:H:10:HIS:HB3	4:H:167:HIS:HD1	1.83	0.43
3:E:96:PRO:HG2	3:E:186:LYS:HE2	2.00	0.43
4:H:133:VAL:O	4:H:243:PRO:HG3	2.18	0.43
3:G:33:VAL:CG1	3:G:34:SER:H	2.26	0.42
4:H:19:VAL:HG22	4:H:94:LEU:HD11	2.01	0.42
3:G:27:LEU:HD12	3:G:89:LEU:HD23	2.00	0.42
4:H:233:ASN:CG	4:H:234:ASP:N	2.73	0.42
2:D:177:HIS:CD2	2:D:178:PRO:HD2	2.54	0.42

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:F:21:LEU:HD22	4:F:122:THR:HG21	2.02	0.42
4:F:234:ASP:OD1	4:F:234:ASP:N	2.53	0.42
2:B:82:ASN:O	2:B:86:GLU:HG2	2.20	0.41
3:E:102:TYR:CE1	3:E:124:LEU:HD23	2.55	0.41
1:A:55:ASP:CG	3:G:112:ALA:HA	2.40	0.41
1:C:9:TYR:N	1:C:9(A):GLY:HA2	2.34	0.41
4:H:233:ASN:N	4:H:233:ASN:OD1	2.52	0.41
2:B:172:THR:HG23	2:B:185:ILE:HG23	2.02	0.41
2:D:133:ARG:HG3	2:D:171:TYR:CE1	2.55	0.41
4:F:9:LYS:HB3	4:F:10:HIS:CD2	2.55	0.41
1:A:113:PHE:CG	1:A:114:PRO:HA	2.56	0.40
3:E:43:ARG:HG2	3:E:53:LEU:HD21	2.03	0.40

All (1) 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)
2:B:167:ARG:NH2	4:F:238:GLN:O[1_546]	1.05	1.15

## 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	180/192 (94%)	178 (99%)	2 (1%)	0	100	100
1	C	180/192 (94%)	175 (97%)	5 (3%)	0	100	100
2	B	178/230 (77%)	170 (96%)	8 (4%)	0	100	100
2	D	167/230 (73%)	160 (96%)	7 (4%)	0	100	100
3	E	191/207 (92%)	185 (97%)	6 (3%)	0	100	100
3	G	190/207 (92%)	185 (97%)	5 (3%)	0	100	100
4	F	238/242 (98%)	230 (97%)	8 (3%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	H	238/242 (98%)	229 (96%)	9 (4%)	0	100	100
5	I	14/16 (88%)	14 (100%)	0	0	100	100
5	J	14/16 (88%)	14 (100%)	0	0	100	100
All	All	1590/1774 (90%)	1540 (97%)	50 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	165/176 (94%)	163 (99%)	2 (1%)	78	92
1	C	165/176 (94%)	161 (98%)	4 (2%)	57	81
2	B	163/203 (80%)	156 (96%)	7 (4%)	35	59
2	D	147/203 (72%)	140 (95%)	7 (5%)	31	53
3	E	163/183 (89%)	156 (96%)	7 (4%)	35	59
3	G	159/183 (87%)	148 (93%)	11 (7%)	19	34
4	F	204/210 (97%)	200 (98%)	4 (2%)	63	85
4	H	203/210 (97%)	196 (97%)	7 (3%)	44	70
5	I	12/13 (92%)	12 (100%)	0	100	100
5	J	12/13 (92%)	12 (100%)	0	100	100
All	All	1393/1570 (89%)	1344 (96%)	49 (4%)	43	68

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	ARG
1	A	158	ASP
2	B	26	LEU
2	B	66	GLU
2	B	126	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	135	ASP
2	B	136	GLN
2	B	138	GLU
2	B	189	ARG
1	C	0	ASP
1	C	11	ASN
1	C	101	GLN
1	C	158	ASP
2	D	66	GLU
2	D	92	GLN
2	D	94	ARG
2	D	100	THR
2	D	126	GLN
2	D	169	ASP
2	D	170	VAL
3	E	10	VAL
3	E	15	GLU
3	E	48	LYS
3	E	58	SER
3	E	66	LYS
3	E	115	MET
3	E	144	LYS
4	F	46	LEU
4	F	184	CYS
4	F	230	LEU
4	F	234	ASP
3	G	15	GLU
3	G	48	LYS
3	G	58	SER
3	G	67	GLU
3	G	86	GLU
3	G	115	MET
3	G	162	VAL
3	G	167	ASP
3	G	175	LYS
3	G	177	VAL
3	G	184	ASP
4	H	64	GLU
4	H	66	ARG
4	H	86	HIS
4	H	128	GLU
4	H	184	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	H	206	ARG
4	H	234	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	134	ASN
2	B	156	GLN
2	D	64	GLN
2	D	156	GLN
3	E	110	ASN
4	F	132	ASN
3	G	114	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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
6	NAG	A	301	1	14,14,15	0.29	0	15,19,21	0.51	0
6	NAG	B	1000	2	14,14,15	0.27	0	15,19,21	0.50	0
6	NAG	C	301	1	14,14,15	0.27	0	15,19,21	0.47	0
6	NAG	D	1000	2	14,14,15	0.29	0	15,19,21	0.50	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	A	301	1	-	0/6/23/26	0/1/1/1
6	NAG	B	1000	2	-	0/6/23/26	0/1/1/1
6	NAG	C	301	1	-	0/6/23/26	0/1/1/1
6	NAG	D	1000	2	-	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.

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, 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	182/192 (94%)	-0.06	4 (2%) 65 70	14, 30, 60, 70	0
1	C	182/192 (94%)	0.05	6 (3%) 50 56	8, 29, 70, 94	0
2	B	181/230 (78%)	-0.03	6 (3%) 50 56	13, 30, 75, 109	0
2	D	173/230 (75%)	0.10	10 (5%) 26 31	7, 26, 98, 128	0
3	E	193/207 (93%)	0.30	9 (4%) 35 41	10, 37, 72, 87	0
3	G	192/207 (92%)	0.24	12 (6%) 23 27	16, 42, 74, 96	0
4	F	240/242 (99%)	0.02	2 (0%) 87 89	8, 26, 53, 82	0
4	H	240/242 (99%)	0.13	8 (3%) 50 56	12, 35, 76, 102	0
5	I	16/16 (100%)	-0.01	1 (6%) 23 27	10, 18, 58, 65	0
5	J	16/16 (100%)	0.18	1 (6%) 23 27	16, 24, 64, 73	0
All	All	1615/1774 (91%)	0.09	59 (3%) 45 51	7, 32, 74, 128	0

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	J	-4	ALA	6.5
3	E	198	ASP	5.9
2	B	164	THR	4.4
3	G	146	SER	4.4
4	H	218	ARG	4.2
1	C	158	ASP	3.9
1	C	157	ALA	3.7
3	E	206	ASN	3.7
3	G	181	ARG	3.6
3	E	147	ASP	3.5
4	F	234	ASP	3.4
3	G	183	MET	3.3
3	E	146	SER	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	D	140	THR	3.3
3	E	181	ARG	3.3
3	G	32	THR	3.3
2	B	135	ASP	3.2
3	G	203	ASN	3.2
2	D	2	ASP	3.1
2	D	137	GLU	3.0
4	H	231	SER	3.0
1	A	175	LEU	3.0
1	C	177	HIS	2.8
2	D	103	PRO	2.8
1	A	158	ASP	2.7
3	G	145	SER	2.6
1	A	0	ASP	2.6
3	G	184	ASP	2.6
3	G	198	ASP	2.6
4	H	253	TRP	2.6
3	E	142	ASP	2.6
2	D	139	THR	2.5
2	B	190	ALA	2.5
1	C	175	LEU	2.5
4	H	177	LYS	2.5
2	D	171	TYR	2.5
4	F	177	LYS	2.4
5	I	-4	ALA	2.4
2	D	187	GLU	2.4
2	D	130	ARG	2.3
2	B	140	THR	2.3
3	G	206	ASN	2.3
1	C	180	PRO	2.3
2	B	188	TRP	2.3
3	E	168	SER	2.3
4	H	181	SER	2.2
3	G	46	PRO	2.2
3	G	207	ASN	2.2
1	A	160	ILE	2.2
2	B	191	GLN	2.2
4	H	239	ASP	2.2
3	E	136	ALA	2.1
1	C	179	GLU	2.1
4	H	174	VAL	2.1
3	E	144	LYS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	G	202	ALA	2.1
2	D	189	ARG	2.1
2	D	188	TRP	2.1
4	H	175	ASN	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( $\text{\AA}^2$ )	Q<0.9
6	NAG	C	301	14/15	0.77	0.23	1.92	66,76,82,82	0
6	NAG	A	301	14/15	0.70	0.32	-	81,87,92,94	0
7	CA	G	301	1/1	0.63	0.27	-	118,118,118,118	0
6	NAG	B	1000	14/15	0.74	0.17	-	82,85,88,88	0
7	CA	E	301	1/1	0.88	0.07	-	66,66,66,66	0
7	CA	C	302	1/1	0.85	0.21	-	65,65,65,65	0
7	CA	A	302	1/1	0.95	0.07	-	75,75,75,75	0
6	NAG	D	1000	14/15	0.79	0.19	-	70,73,75,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.