



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:28 PM GMT

PDB ID : 4OZH
Title : S16 protein complex
Authors : Petersen, J.; Reid, H.H.; Rossjohn, J.
Deposited on : 2014-02-16
Resolution : 2.80 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
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

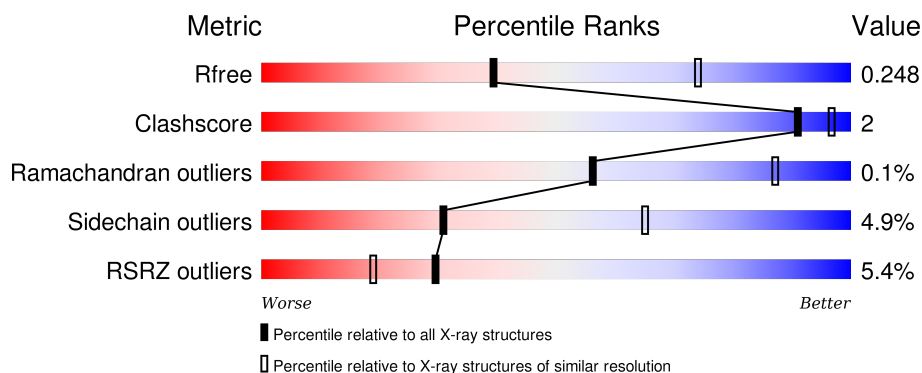
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	191	<div> <div>6%</div> <div>86% 8% 5%</div> </div>
1	C	191	<div> <div>12%</div> <div>88% 6% 5%</div> </div>
2	B	213	<div> <div>%</div> <div>75% 10% 15%</div> </div>
2	D	213	<div> <div>9%</div> <div>73% 12% 15%</div> </div>
3	E	203	<div> <div>3%</div> <div>88% 7% . .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	203	
4	F	241	
4	H	241	
5	I	13	
5	J	13	

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 12818 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	181	Total	C	N	O	S	0	0	0
			1445	931	236	275	3			
1	C	181	Total	C	N	O	S	0	0	0
			1438	926	236	273	3			

- 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	0	0
			1480	935	264	274	7			
2	D	181	Total	C	N	O	S	0	0	0
			1463	927	259	270	7			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	GLY	-	expression tag	UNP Q5Y7D3
B	-11	GLY	-	expression tag	UNP Q5Y7D3
B	-10	SER	-	expression tag	UNP Q5Y7D3
B	-9	ILE	-	expression tag	UNP Q5Y7D3
B	-8	GLU	-	expression tag	UNP Q5Y7D3
B	-7	GLY	-	expression tag	UNP Q5Y7D3
B	-6	ARG	-	expression tag	UNP Q5Y7D3
B	-5	GLY	-	expression tag	UNP Q5Y7D3
B	-4	GLY	-	expression tag	UNP Q5Y7D3
B	-3	SER	-	expression tag	UNP Q5Y7D3
B	-2	GLY	-	expression tag	UNP Q5Y7D3
B	-1	ALA	-	expression tag	UNP Q5Y7D3
B	0	SER	-	expression tag	UNP Q5Y7D3
B	193	THR	-	expression tag	UNP Q5Y7D3
B	194	GLY	-	expression tag	UNP Q5Y7D3
B	195	GLY	-	expression tag	UNP Q5Y7D3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	196	ASP	-	expression tag	UNP Q5Y7D3
B	197	ASP	-	expression tag	UNP Q5Y7D3
B	198	ASP	-	expression tag	UNP Q5Y7D3
B	199	ASP	-	expression tag	UNP Q5Y7D3
B	200	LYS	-	expression tag	UNP Q5Y7D3
D	-12	GLY	-	expression tag	UNP Q5Y7D3
D	-11	GLY	-	expression tag	UNP Q5Y7D3
D	-10	SER	-	expression tag	UNP Q5Y7D3
D	-9	ILE	-	expression tag	UNP Q5Y7D3
D	-8	GLU	-	expression tag	UNP Q5Y7D3
D	-7	GLY	-	expression tag	UNP Q5Y7D3
D	-6	ARG	-	expression tag	UNP Q5Y7D3
D	-5	GLY	-	expression tag	UNP Q5Y7D3
D	-4	GLY	-	expression tag	UNP Q5Y7D3
D	-3	SER	-	expression tag	UNP Q5Y7D3
D	-2	GLY	-	expression tag	UNP Q5Y7D3
D	-1	ALA	-	expression tag	UNP Q5Y7D3
D	0	SER	-	expression tag	UNP Q5Y7D3
D	193	THR	-	expression tag	UNP Q5Y7D3
D	194	GLY	-	expression tag	UNP Q5Y7D3
D	195	GLY	-	expression tag	UNP Q5Y7D3
D	196	ASP	-	expression tag	UNP Q5Y7D3
D	197	ASP	-	expression tag	UNP Q5Y7D3
D	198	ASP	-	expression tag	UNP Q5Y7D3
D	199	ASP	-	expression tag	UNP Q5Y7D3
D	200	LYS	-	expression tag	UNP Q5Y7D3

- Molecule 3 is a protein called T-cell receptor, s16, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	195	Total	C	N	O	S	0	0	0
			1492	935	252	295	10			
3	G	195	Total	C	N	O	S	0	0	0
			1489	934	251	294	10			

- Molecule 4 is a protein called T-cell receptor, s16, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	240	Total	C	N	O	S	0	0	0
			1872	1174	330	363	5			
4	H	240	Total	C	N	O	S	0	0	0
			1869	1174	327	363	5			

- Molecule 5 is a protein called Gliadin-alpha2 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	12	Total	C	N	O	0	0	0
			91	60	14	17			
5	J	13	Total	C	N	O	0	0	0
			96	63	15	18			

- Molecule 6 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (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	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	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		

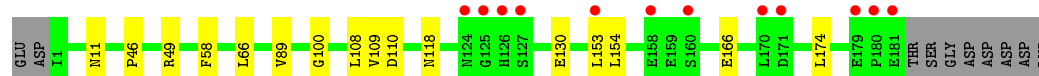
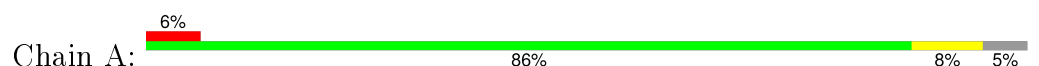
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	3	Total O 3 3	0	0
8	B	4	Total O 4 4	0	0
8	C	1	Total O 1 1	0	0
8	D	4	Total O 4 4	0	0
8	E	7	Total O 7 7	0	0
8	F	10	Total O 10 10	0	0
8	G	3	Total O 3 3	0	0
8	H	8	Total O 8 8	0	0

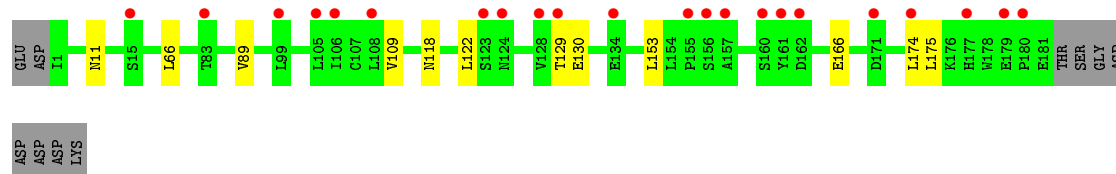
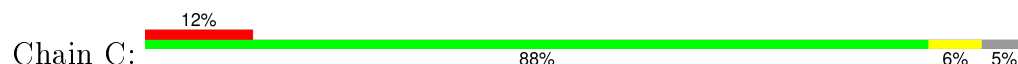
3 Residue-property plots

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.

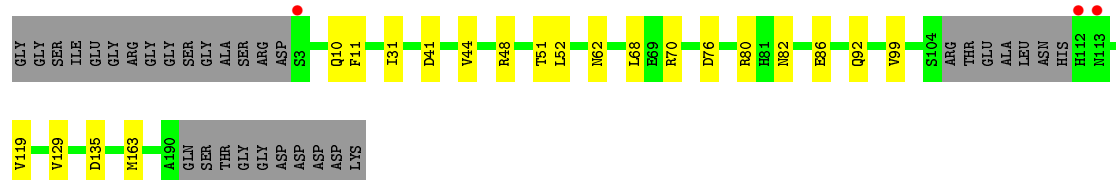
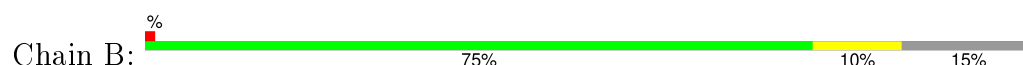
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain



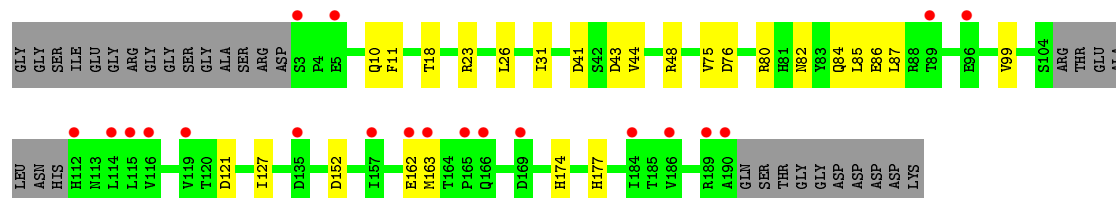
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain



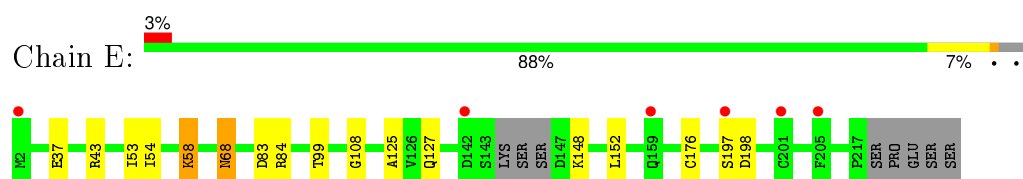
- Molecule 2: HLA class II histocompatibility antigen, DQ beta 1 chain



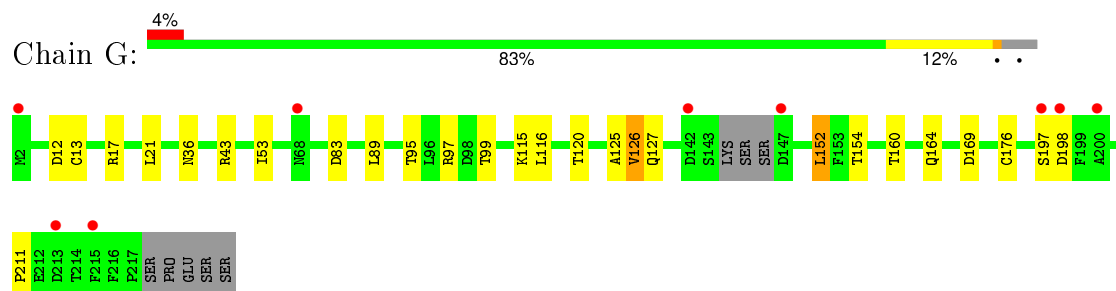
- Molecule 2: HLA class II histocompatibility antigen, DQ beta 1 chain



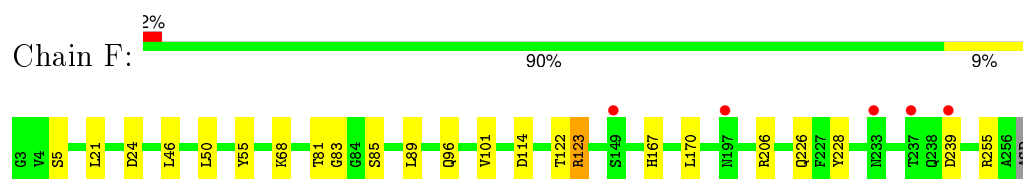
- Molecule 3: T-cell receptor, s16, alpha chain



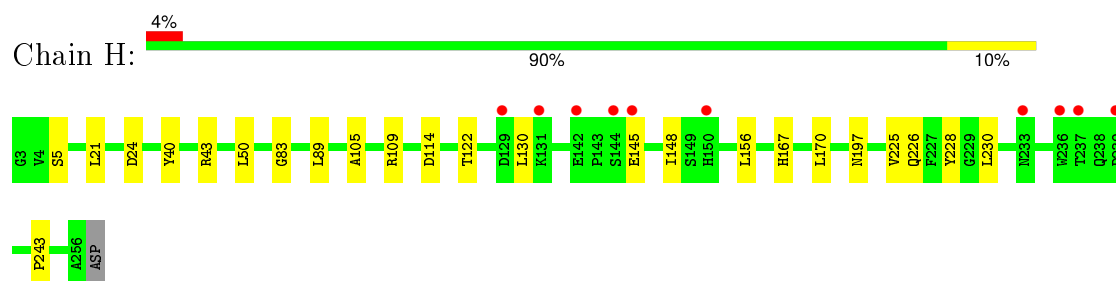
- Molecule 3: T-cell receptor, s16, alpha chain



- Molecule 4: T-cell receptor, s16, beta chain



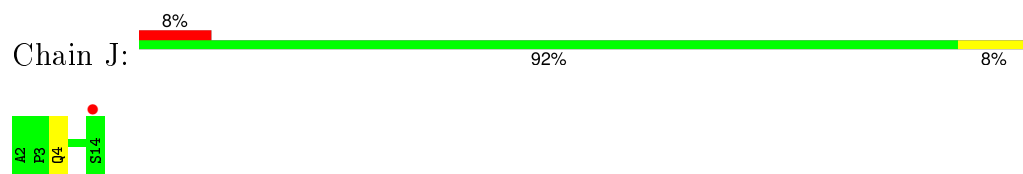
- Molecule 4: T-cell receptor, s16, beta chain



- Molecule 5: Gliadin-alpha2 peptide



- Molecule 5: Gliadin-alpha2 peptide



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	261.57Å 57.87Å 137.41Å 90.00° 114.05° 90.00°	Depositor
Resolution (Å)	40.99 – 2.80 47.00 – 2.80	Depositor EDS
% Data completeness (in resolution range)	93.1 (40.99-2.80) 93.2 (47.00-2.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 2.81Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.226 , 0.243 0.235 , 0.248	Depositor DCC
R_{free} test set	2204 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	58.3	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 43700 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12818	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.26% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.37	0/1487	0.54	0/2031
1	C	0.38	0/1480	0.54	0/2022
2	B	0.37	0/1513	0.55	0/2056
2	D	0.37	0/1496	0.55	0/2035
3	E	0.51	0/1526	0.60	1/2082 (0.0%)
3	G	0.54	0/1523	0.60	1/2078 (0.0%)
4	F	0.34	0/1920	0.53	0/2615
4	H	0.35	0/1917	0.52	0/2612
5	I	0.38	0/96	0.51	0/134
5	J	0.39	0/101	0.51	0/141
All	All	0.41	0/13059	0.55	2/17806 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	G	127	GLN	C-N-CD	5.83	140.65	128.40
3	E	127	GLN	C-N-CD	5.82	140.61	128.40

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	1445	0	1396	4	1
1	C	1438	0	1384	4	0
2	B	1480	0	1443	7	0
2	D	1463	0	1418	9	0
3	E	1492	0	1384	6	0
3	G	1489	0	1380	9	0
4	F	1872	0	1777	7	0
4	H	1869	0	1775	10	0
5	I	91	0	84	1	0
5	J	96	0	86	1	0
6	A	28	0	26	0	0
6	C	14	0	13	0	0
7	G	1	0	0	0	0
8	A	3	0	0	0	0
8	B	4	0	0	0	0
8	C	1	0	0	0	0
8	D	4	0	0	0	0
8	E	7	0	0	0	0
8	F	10	0	0	0	0
8	G	3	0	0	0	0
8	H	8	0	0	0	0
All	All	12818	0	12166	54	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 2.

All (54) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:11:ASN:HB3	1:C:66:LEU:HD11	1.64	0.79
4:F:21:LEU:HD22	4:F:122:THR:HG21	1.86	0.58
4:F:83:GLY:HA3	4:F:85:SER:H	1.69	0.57
1:A:118:ASN:HB2	1:A:166:GLU:HB2	1.86	0.57
2:D:10:GLN:HB2	2:D:31:ILE:HB	1.89	0.54
2:B:10:GLN:HB2	2:B:31:ILE:HB	1.89	0.53
1:C:118:ASN:HB2	1:C:166:GLU:HB2	1.92	0.52
1:A:11:ASN:HB2	2:B:11:PHE:HB3	1.91	0.52
2:B:76:ASP:HA	2:B:80:ARG:HB2	1.93	0.50
3:E:43:ARG:HB3	3:E:53:ILE:HD11	1.94	0.50
4:F:96:GLN:HG2	4:H:197:ASN:HB2	1.92	0.50
3:G:160:THR:HG21	3:G:211:PRO:HG3	1.94	0.50
2:D:76:ASP:HA	2:D:80:ARG:HB2	1.94	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:21:LEU:HD12	4:H:89:LEU:HD23	1.94	0.49
3:E:58:LYS:HG3	3:E:84:ARG:HD3	1.93	0.49
3:G:43:ARG:HB3	3:G:53:ILE:HD11	1.94	0.49
2:D:84:GLN:HA	2:D:87:LEU:HD12	1.94	0.49
3:G:21:LEU:HD12	3:G:89:LEU:HD23	1.96	0.48
4:H:21:LEU:HD22	4:H:122:THR:HG21	1.95	0.47
2:B:41:ASP:HB3	2:B:44:VAL:HG22	1.96	0.47
1:C:11:ASN:HB2	2:D:11:PHE:HB3	1.97	0.47
5:I:11:GLN:HG3	5:I:12:PRO:HD2	1.96	0.47
2:D:127:ILE:HD12	2:D:177:HIS:HD2	1.80	0.47
2:D:41:ASP:HB3	2:D:44:VAL:HG22	1.96	0.47
3:G:197:SER:HA	3:G:198:ASP:HA	1.69	0.46
3:G:95:THR:O	3:G:126:VAL:HG21	2.14	0.46
3:G:13:CYS:O	3:G:126:VAL:HA	2.15	0.46
4:H:230:LEU:HD12	4:H:243:PRO:HD2	1.97	0.46
1:A:46:PRO:HA	1:A:49:ARG:HD2	1.98	0.45
3:E:54:ILE:HD12	3:E:68:ASN:HB2	1.98	0.45
2:B:99:VAL:HG12	2:B:119:VAL:HG22	1.97	0.45
4:H:5:SER:HB2	4:H:24:ASP:HB3	1.98	0.45
1:C:89:VAL:HG22	1:C:109:VAL:HG13	1.99	0.45
4:F:21:LEU:HD12	4:F:89:LEU:HD23	1.97	0.45
2:D:82:ASN:O	2:D:86:GLU:HG2	2.17	0.45
3:E:197:SER:HA	3:E:198:ASP:HA	1.74	0.45
4:H:40:TYR:HB2	4:H:105:ALA:HB3	1.98	0.45
4:H:145:GLU:HA	4:H:148:ILE:HD12	1.98	0.45
4:F:101:VAL:HG22	4:F:123:ARG:HG2	1.99	0.44
1:A:89:VAL:HG22	1:A:109:VAL:HG13	1.99	0.44
4:H:167:HIS:HB3	4:H:228:TYR:HB2	1.99	0.44
3:G:152:LEU:HD22	3:G:154:THR:HB	1.98	0.44
2:D:82:ASN:HA	2:D:85:LEU:HD12	1.98	0.44
4:F:5:SER:HB2	4:F:24:ASP:HB2	1.99	0.44
3:E:99:THR:HG23	3:E:125:ALA:HA	2.00	0.44
2:D:18:THR:HB	2:D:23:ARG:HB2	1.98	0.44
4:F:167:HIS:HB3	4:F:228:TYR:HB2	2.00	0.44
4:H:170:LEU:HG	4:H:225:VAL:HG22	2.00	0.43
2:B:62:ASN:HA	2:B:68:LEU:HD13	2.01	0.43
3:G:99:THR:HG23	3:G:125:ALA:HA	2.01	0.42
3:G:36:ASN:HD22	5:J:4:GLN:HB2	1.86	0.41
2:B:82:ASN:O	2:B:86:GLU:HG2	2.20	0.41
3:E:37:GLU:HG2	3:E:108:GLY:HA3	2.02	0.41
4:H:130:LEU:HD22	4:H:230: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)
1:A:100:GLY:CA	1:A:130:GLU:OE2[2_659]	2.15	0.05

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	179/191 (94%)	175 (98%)	4 (2%)	0	100	100
1	C	179/191 (94%)	174 (97%)	5 (3%)	0	100	100
2	B	177/213 (83%)	167 (94%)	10 (6%)	0	100	100
2	D	177/213 (83%)	168 (95%)	8 (4%)	1 (1%)	30	65
3	E	191/203 (94%)	185 (97%)	6 (3%)	0	100	100
3	G	191/203 (94%)	178 (93%)	13 (7%)	0	100	100
4	F	238/241 (99%)	228 (96%)	10 (4%)	0	100	100
4	H	238/241 (99%)	228 (96%)	9 (4%)	1 (0%)	39	74
5	I	10/13 (77%)	9 (90%)	1 (10%)	0	100	100
5	J	11/13 (85%)	10 (91%)	1 (9%)	0	100	100
All	All	1591/1722 (92%)	1522 (96%)	67 (4%)	2 (0%)	56	87

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	H	83	GLY
2	D	121	ASP

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	165/174 (95%)	158 (96%)	7 (4%)	36	71
1	C	163/174 (94%)	157 (96%)	6 (4%)	41	76
2	B	164/188 (87%)	156 (95%)	8 (5%)	31	65
2	D	160/188 (85%)	151 (94%)	9 (6%)	26	59
3	E	162/181 (90%)	156 (96%)	6 (4%)	41	76
3	G	161/181 (89%)	149 (92%)	12 (8%)	17	43
4	F	203/209 (97%)	191 (94%)	12 (6%)	24	57
4	H	203/209 (97%)	197 (97%)	6 (3%)	48	82
5	I	10/11 (91%)	8 (80%)	2 (20%)	1	5
5	J	10/11 (91%)	10 (100%)	0	100	100
All	All	1401/1526 (92%)	1333 (95%)	68 (5%)	31	65

All (68) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	PHE
1	A	66	LEU
1	A	108	LEU
1	A	110	ASP
1	A	153	LEU
1	A	154	LEU
1	A	174	LEU
2	B	48	ARG
2	B	51	THR
2	B	52	LEU
2	B	70	ARG
2	B	92	GLN
2	B	129	VAL
2	B	135	ASP
2	B	163	MET
1	C	122	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	129	THR
1	C	130	GLU
1	C	153	LEU
1	C	174	LEU
1	C	175	LEU
2	D	26	LEU
2	D	43	ASP
2	D	48	ARG
2	D	75	VAL
2	D	99	VAL
2	D	152	ASP
2	D	162	GLU
2	D	163	MET
2	D	174	HIS
3	E	58	LYS
3	E	68	ASN
3	E	83	ASP
3	E	148	LYS
3	E	152	LEU
3	E	176	CYS
4	F	46	LEU
4	F	50	LEU
4	F	55	TYR
4	F	68	LYS
4	F	81	THR
4	F	114	ASP
4	F	123	ARG
4	F	170	LEU
4	F	206	ARG
4	F	226	GLN
4	F	239	ASP
4	F	255	ARG
3	G	12	ASP
3	G	17	ARG
3	G	83	ASP
3	G	97	ARG
3	G	115	LYS
3	G	116	LEU
3	G	120	THR
3	G	126	VAL
3	G	152	LEU
3	G	164	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	169	ASP
3	G	176	CYS
4	H	43	ARG
4	H	50	LEU
4	H	109	ARG
4	H	114	ASP
4	H	156	LEU
4	H	226	GLN
5	I	4	GLN
5	I	7	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	134	ASN
4	F	92	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](#)

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
6	NAG	A	301	1	14,14,15	0.25	0	15,19,21	0.57	0
6	NAG	A	302	1	14,14,15	0.26	0	15,19,21	0.59	1 (6%)
6	NAG	C	301	1	14,14,15	0.27	0	15,19,21	0.60	1 (6%)

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	A	302	1	-	0/6/23/26	0/1/1/1
6	NAG	C	301	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	302	NAG	C1-O5-C5	2.05	114.85	112.25
6	C	301	NAG	C1-O5-C5	2.09	114.90	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	181/191 (94%)	0.37	12 (6%) 22 13	43, 65, 93, 124	0
1	C	181/191 (94%)	0.77	22 (12%) 5 3	48, 81, 150, 194	0
2	B	181/213 (84%)	0.28	3 (1%) 73 63	40, 63, 85, 109	0
2	D	181/213 (84%)	0.66	20 (11%) 7 3	38, 78, 144, 187	0
3	E	195/203 (96%)	0.28	6 (3%) 52 40	27, 52, 91, 108	0
3	G	195/203 (96%)	0.43	9 (4%) 36 25	40, 63, 81, 92	0
4	F	240/241 (99%)	0.10	5 (2%) 67 56	31, 47, 78, 100	0
4	H	240/241 (99%)	0.20	10 (4%) 40 28	39, 55, 76, 91	0
5	I	12/13 (92%)	0.20	0 100 100	43, 50, 61, 75	0
5	J	13/13 (100%)	0.71	1 (7%) 16 8	55, 63, 84, 101	0
All	All	1619/1722 (94%)	0.37	88 (5%) 29 19	27, 60, 120, 194	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	2	MET	6.2
1	C	162	ASP	4.7
3	E	197	SER	4.7
3	G	215	PHE	3.9
2	B	112	HIS	3.9
2	D	3	SER	3.6
1	C	179	GLU	3.6
1	C	157	ALA	3.4
2	D	135	ASP	3.3
1	A	158	GLU	3.3
4	F	239	ASP	3.2
1	A	126	HIS	3.2
4	H	131	LYS	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	3	SER	3.1
3	G	197	SER	3.1
1	C	105	LEU	3.1
1	C	161	TYR	3.1
4	H	129	ASP	3.1
3	G	147	ASP	3.0
2	D	165	PRO	3.0
4	H	233	ASN	3.0
1	C	128	VAL	2.9
1	C	124	ASN	2.9
2	B	113	ASN	2.9
1	A	171	ASP	2.8
2	D	112	HIS	2.8
1	A	179	GLU	2.8
3	G	198	ASP	2.8
1	C	156	SER	2.8
2	D	5	GLU	2.8
1	C	171	ASP	2.8
3	E	201	CYS	2.7
2	D	189	ARG	2.7
4	F	233	ASN	2.7
2	D	184	ILE	2.7
2	D	169	ASP	2.7
1	C	99	LEU	2.6
1	C	15	SER	2.6
2	D	163	MET	2.6
4	F	149	SER	2.6
3	G	213	ASP	2.6
4	F	197	ASN	2.6
2	D	119	VAL	2.6
1	C	160	SER	2.5
2	D	190	ALA	2.5
1	C	83	THR	2.5
1	A	181	GLU	2.5
1	C	106	ILE	2.5
3	E	205	PHE	2.4
2	D	157	ILE	2.4
4	H	145	GLU	2.4
5	J	14	SER	2.4
3	E	159	GLN	2.3
1	A	170	LEU	2.3
1	C	177	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	116	VAL	2.3
2	D	162	GLU	2.3
2	D	166	GLN	2.3
1	A	125	GLY	2.3
2	D	115	LEU	2.3
1	C	174	LEU	2.2
4	H	142	GLU	2.2
4	H	236	TRP	2.2
3	E	2	MET	2.2
1	C	129	THR	2.2
4	F	237	THR	2.2
1	A	127	SER	2.2
1	A	160	SER	2.2
1	C	123	SER	2.2
4	H	239	ASP	2.1
1	C	134	GLU	2.1
4	H	237	THR	2.1
1	A	180	PRO	2.1
1	A	153	LEU	2.1
3	E	142	ASP	2.1
3	G	142	ASP	2.1
4	H	144	SER	2.1
1	C	108	LEU	2.1
1	C	180	PRO	2.1
2	D	186	VAL	2.1
3	G	200	ALA	2.1
2	D	89	THR	2.1
1	C	155	PRO	2.1
4	H	150	HIS	2.1
1	A	124	ASN	2.1
2	D	114	LEU	2.0
2	D	96	GLU	2.0
3	G	68	ASN	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 ⓘ

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	NAG	A	301	14/15	0.82	0.26	0.32	95,98,102,103	0
6	NAG	C	301	14/15	0.65	0.23	-0.72	113,120,127,128	0
7	CA	G	301	1/1	0.82	0.08	-5.36	71,71,71,71	0
6	NAG	A	302	14/15	0.69	0.24	-	124,127,136,137	0

6.5 Other polymers [i](#)

There are no such residues in this entry.