



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:53 PM GMT

PDB ID : 4Z7U
Title : S13 complex
Authors : Petersen, J.; Rossjohn, J.; Reid, H.H.; Koning, F.
Deposited on : 2015-04-08
Resolution : 2.70 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

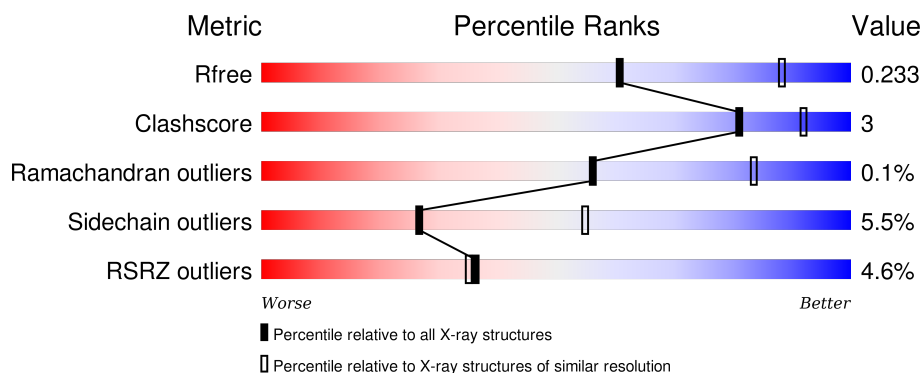
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	192	<div> <div></div> <div>83% 11% 5%</div> </div>
1	C	192	<div> <div></div> <div>84% 11% 5%</div> </div>
2	B	213	<div> <div>15%</div> <div>69% 10% 21%</div> </div>
2	D	213	<div> <div>13%</div> <div>68% 12% 19%</div> </div>
3	E	203	<div> <div></div> <div>86% 10% ..</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	203	
4	F	245	
4	H	245	
5	I	18	
5	J	18	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	NAG	A	1000	-	-	-	X
6	NAG	C	1000	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13076 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 MHC class II HLA-DQ-alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	182	Total	C	N	O	S	0	0	0
			1450	933	239	276	2			
1	C	182	Total	C	N	O	S	0	0	0
			1454	936	240	276	2			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	182	THR	-	expression tag	UNP Q30069
A	183	SER	-	expression tag	UNP Q30069
A	184	GLY	-	expression tag	UNP Q30069
A	185	ASP	-	expression tag	UNP Q30069
A	186	ASP	-	expression tag	UNP Q30069
A	187	ASP	-	expression tag	UNP Q30069
A	188	ASP	-	expression tag	UNP Q30069
A	189	LYS	-	expression tag	UNP Q30069
C	182	THR	-	expression tag	UNP Q30069
C	183	SER	-	expression tag	UNP Q30069
C	184	GLY	-	expression tag	UNP Q30069
C	185	ASP	-	expression tag	UNP Q30069
C	186	ASP	-	expression tag	UNP Q30069
C	187	ASP	-	expression tag	UNP Q30069
C	188	ASP	-	expression tag	UNP Q30069
C	189	LYS	-	expression tag	UNP Q30069

- Molecule 2 is a protein called MHC class II HLA-DQ-beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	169	Total	C	N	O	S	0	0	0
			1345	858	232	248	7			
2	D	172	Total	C	N	O	S	0	0	0
			1354	863	235	250	6			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-12	GLY	-	expression tag	UNP O19707
B	-11	GLY	-	expression tag	UNP O19707
B	-10	SER	-	expression tag	UNP O19707
B	-9	ILE	-	expression tag	UNP O19707
B	-8	GLU	-	expression tag	UNP O19707
B	-7	GLY	-	expression tag	UNP O19707
B	-6	ARG	-	expression tag	UNP O19707
B	-5	GLY	-	expression tag	UNP O19707
B	-4	GLY	-	expression tag	UNP O19707
B	-3	SER	-	expression tag	UNP O19707
B	-2	GLY	-	expression tag	UNP O19707
B	-1	ALA	-	expression tag	UNP O19707
B	0	SER	-	expression tag	UNP O19707
B	193	THR	-	expression tag	UNP O19707
B	194	GLY	-	expression tag	UNP O19707
B	195	GLY	-	expression tag	UNP O19707
B	196	ASP	-	expression tag	UNP O19707
B	197	ASP	-	expression tag	UNP O19707
B	198	ASP	-	expression tag	UNP O19707
B	199	ASP	-	expression tag	UNP O19707
B	200	LYS	-	expression tag	UNP O19707
D	-12	GLY	-	expression tag	UNP O19707
D	-11	GLY	-	expression tag	UNP O19707
D	-10	SER	-	expression tag	UNP O19707
D	-9	ILE	-	expression tag	UNP O19707
D	-8	GLU	-	expression tag	UNP O19707
D	-7	GLY	-	expression tag	UNP O19707
D	-6	ARG	-	expression tag	UNP O19707
D	-5	GLY	-	expression tag	UNP O19707
D	-4	GLY	-	expression tag	UNP O19707
D	-3	SER	-	expression tag	UNP O19707
D	-2	GLY	-	expression tag	UNP O19707
D	-1	ALA	-	expression tag	UNP O19707
D	0	SER	-	expression tag	UNP O19707
D	193	THR	-	expression tag	UNP O19707
D	194	GLY	-	expression tag	UNP O19707
D	195	GLY	-	expression tag	UNP O19707
D	196	ASP	-	expression tag	UNP O19707
D	197	ASP	-	expression tag	UNP O19707
D	198	ASP	-	expression tag	UNP O19707
D	199	ASP	-	expression tag	UNP O19707
D	200	LYS	-	expression tag	UNP O19707

- Molecule 3 is a protein called T-CELL RECEPTOR, S13 ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	198	Total	C	N	O	S	0	0	0
			1533	953	262	310	8			
3	G	198	Total	C	N	O	S	0	0	0
			1525	951	262	303	9			

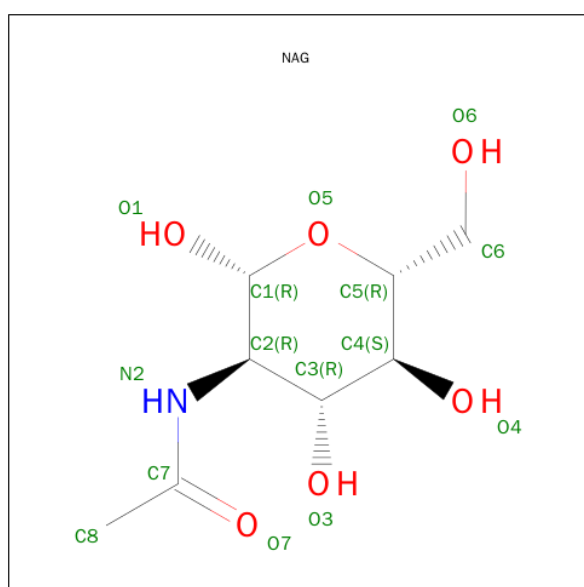
- Molecule 4 is a protein called T-CELL RECEPTOR, S13 BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	244	Total	C	N	O	S	0	0	0
			1904	1205	328	366	5			
4	H	244	Total	C	N	O	S	0	0	0
			1905	1205	329	366	5			

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

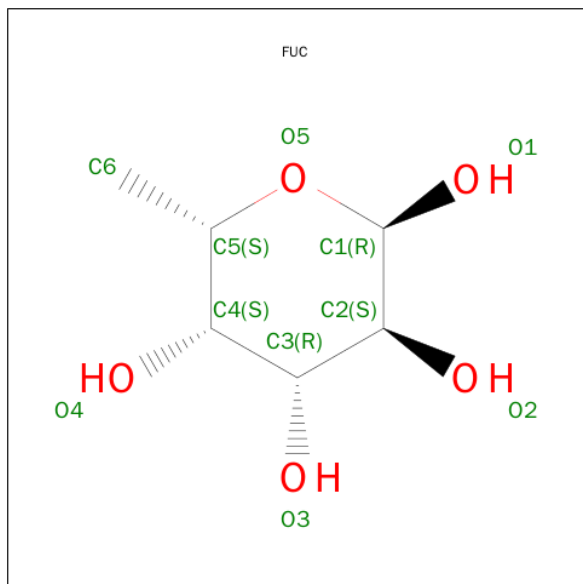
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	15	Total	C	N	O	0	0	0
			107	64	18	25			
5	J	16	Total	C	N	O	0	0	0
			112	67	19	26			

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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	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		
6	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 7 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C₆H₁₂O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			10	6	4		
7	B	1	Total	C	O	0	0
			10	6	4		
7	D	1	Total	C	O	0	0
			10	6	4		
7	D	1	Total	C	O	0	0
			10	6	4		

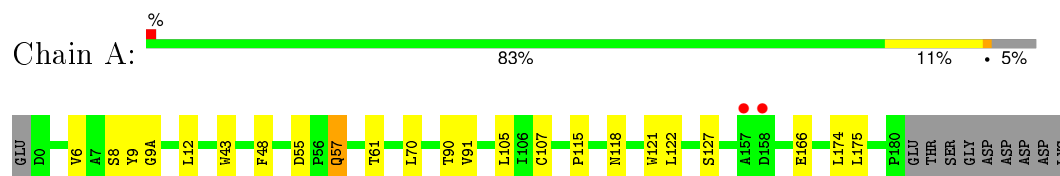
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	29	Total 29	O 29	0	0
8	B	22	Total 22	O 22	0	0
8	C	25	Total 25	O 25	0	0
8	D	20	Total 20	O 20	0	0
8	E	48	Total 48	O 48	0	0
8	F	28	Total 28	O 28	0	0
8	G	44	Total 44	O 44	0	0
8	H	43	Total 43	O 43	0	0
8	I	2	Total 2	O 2	0	0
8	J	2	Total 2	O 2	0	0

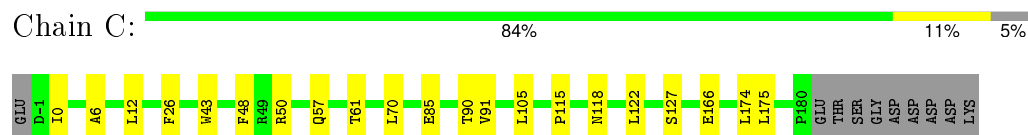
3 Residue-property plots [i](#)

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

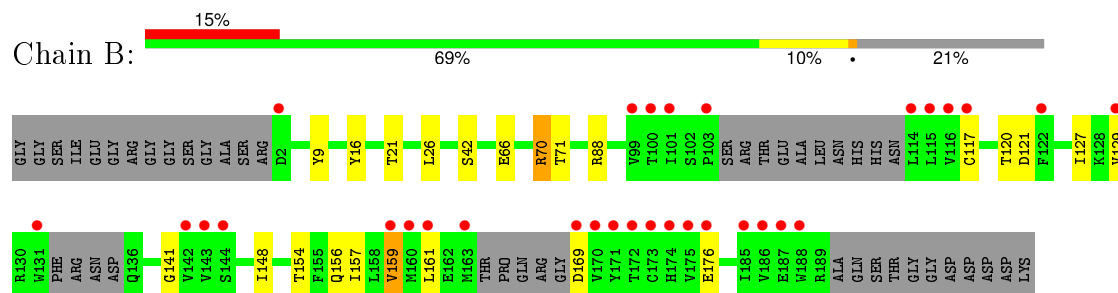
- Molecule 1: MHC class II HLA-DQ-alpha chain



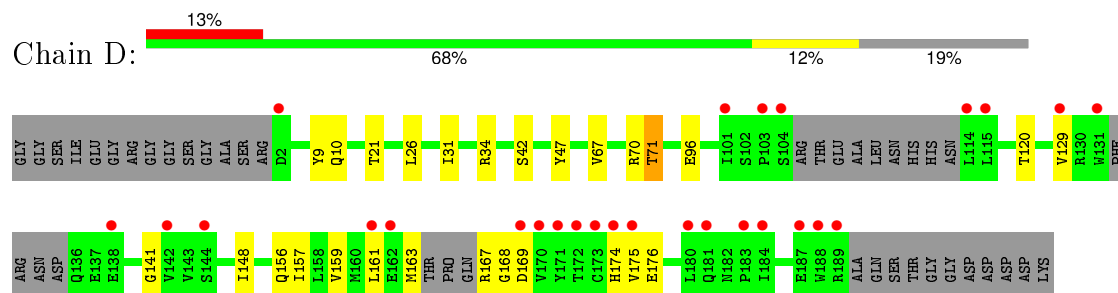
- Molecule 1: MHC class II HLA-DQ-alpha chain



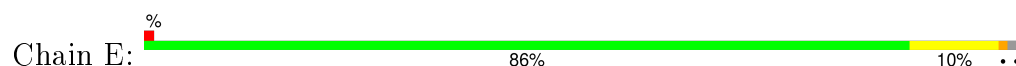
- Molecule 2: MHC class II HLA-DQ-beta-1



- Molecule 2: MHC class II HLA-DQ-beta-1

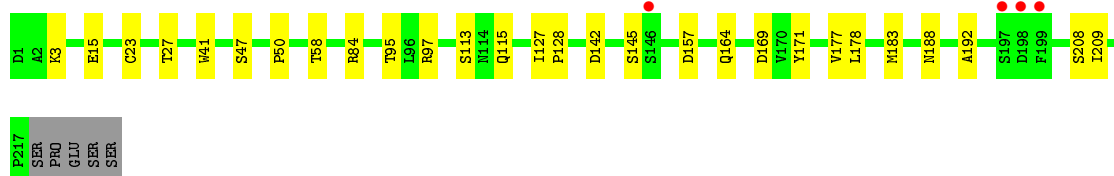
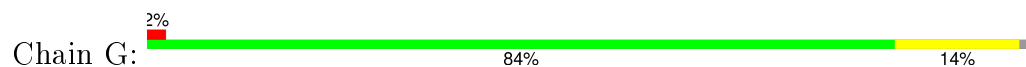


- Molecule 3: T-CELL RECEPTOR, S13 ALPHA CHAIN

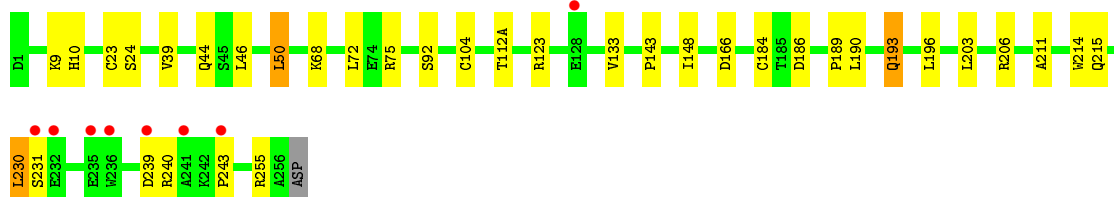
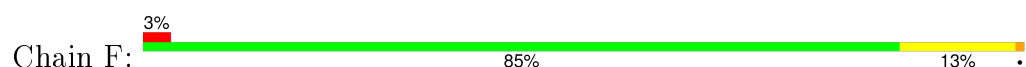




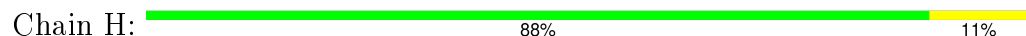
- Molecule 3: T-CELL RECEPTOR, S13 ALPHA CHAIN



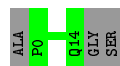
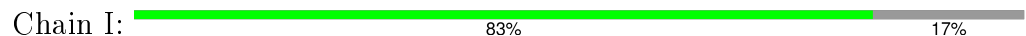
- Molecule 4: T-CELL RECEPTOR, S13 BETA CHAIN



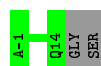
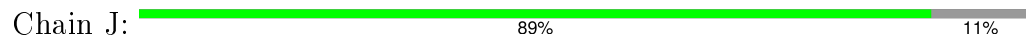
- Molecule 4: T-CELL RECEPTOR, S13 BETA CHAIN



- Molecule 5: deamidated DQ8-glia-alpha1 peptide



- Molecule 5: deamidated DQ8-glia-alpha1 peptide



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.45Å 123.81Å 223.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.53 – 2.70 38.93 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.4 (39.53-2.70) 94.4 (38.93-2.70)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.34 (at 2.69Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.178 , 0.220 0.191 , 0.233	Depositor DCC
R_{free} test set	3020 reflections (5.24%)	DCC
Wilson B-factor (Å ²)	51.7	Xtriage
Anisotropy	0.542	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 65.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 60705 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13076	wwPDB-VP
Average B, all atoms (Å ²)	67.0	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 36.50 % 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 4.9503e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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: NAG, FUC

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	0.46	0/1494	0.67	0/2044
1	C	0.46	0/1498	0.68	0/2048
2	B	0.47	0/1378	0.68	0/1881
2	D	0.47	0/1387	0.70	0/1895
3	E	0.57	0/1568	0.76	0/2139
3	G	0.57	0/1560	0.74	0/2129
4	F	0.47	0/1955	0.70	0/2669
4	H	0.50	0/1956	0.72	0/2670
5	I	0.45	0/110	0.61	0/148
5	J	0.57	0/115	0.61	0/156
All	All	0.50	0/13021	0.71	0/17779

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	1450	0	1362	13	0
1	C	1454	0	1373	9	0
2	B	1345	0	1245	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1354	0	1246	13	0
3	E	1533	0	1423	11	0
3	G	1525	0	1413	10	0
4	F	1904	0	1795	13	0
4	H	1905	0	1800	12	0
5	I	107	0	89	0	0
5	J	112	0	94	0	0
6	A	14	0	13	0	0
6	B	28	0	23	0	0
6	C	14	0	13	0	0
6	D	28	0	23	0	0
7	B	20	0	20	1	0
7	D	20	0	20	1	0
8	A	29	0	0	0	0
8	B	22	0	0	1	0
8	C	25	0	0	0	0
8	D	20	0	0	0	0
8	E	48	0	0	2	0
8	F	28	0	0	0	0
8	G	44	0	0	0	0
8	H	43	0	0	0	0
8	I	2	0	0	0	0
8	J	2	0	0	0	0
All	All	13076	0	11952	81	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 (81) 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)
4:H:109:THR:HG22	4:H:111:GLY:H	1.52	0.73
2:B:120:THR:HG22	2:B:156:GLN:HG3	1.76	0.67
2:B:66:GLU:O	2:B:70:ARG:HG2	1.94	0.67
2:B:117:CYS:HB3	2:B:159:VAL:HG23	1.77	0.66
3:G:128:PRO:HG3	3:G:177:VAL:HG11	1.75	0.66
4:F:133:VAL:HG11	4:F:230:LEU:HD13	1.77	0.65
2:B:26:LEU:HG	2:B:42:SER:HB3	1.82	0.60
2:D:129:VAL:HG22	2:D:175:VAL:HG13	1.84	0.60
4:H:109:THR:HG23	4:H:110:PRO:HD2	1.84	0.59
4:H:166:ASP:HB2	4:H:189:PRO:HG2	1.84	0.58

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:167:ARG:N	2:D:168:GLY:HA2	2.18	0.58
2:D:47:TYR:OH	2:D:71:THR:HG21	2.04	0.57
4:F:166:ASP:HB2	4:F:189:PRO:HG2	1.85	0.57
3:E:120:THR:CG2	8:E:303:HOH:O	2.53	0.57
4:H:21:LEU:HD22	4:H:122:THR:HG21	1.86	0.56
1:A:70:LEU:HD13	2:B:9:TYR:HB2	1.88	0.56
2:B:88:ARG:NH1	8:B:1101:HOH:O	2.38	0.56
1:A:43:TRP:CE3	1:A:48:PHE:HB3	2.42	0.55
3:G:178:LEU:HB3	4:H:184:CYS:HB2	1.87	0.55
3:E:12:GLU:HG3	3:E:125:THR:HG22	1.88	0.55
1:C:43:TRP:CE3	1:C:48:PHE:HB3	2.42	0.54
4:H:44:GLN:HB2	4:H:50:LEU:HD13	1.90	0.53
1:A:55:ASP:OD1	1:A:57:GLN:HB3	2.08	0.52
1:A:9:TYR:N	1:A:9(A):GLY:HA2	2.23	0.52
3:E:95:THR:HG22	3:E:97:ARG:H	1.74	0.52
2:D:129:VAL:HG13	2:D:175:VAL:HG22	1.91	0.52
3:G:95:THR:HG22	3:G:97:ARG:H	1.74	0.52
3:E:142:ASP:HB3	3:E:145:SER:O	2.10	0.52
3:G:50:PRO:HD2	4:H:118:PHE:CD2	2.45	0.52
2:B:21:THR:H	7:B:1003:FUC:H4	1.75	0.51
2:D:21:THR:H	7:D:1003:FUC:H4	1.76	0.51
1:A:8:SER:C	1:A:9(A):GLY:HA2	2.31	0.51
4:H:167:HIS:HB3	4:H:228:TYR:HB2	1.93	0.51
2:B:148:ILE:HB	2:B:156:GLN:HB3	1.93	0.51
4:F:133:VAL:O	4:F:240:ARG:NH2	2.41	0.50
4:F:44:GLN:HB2	4:F:50:LEU:HD13	1.93	0.50
3:G:142:ASP:HB3	3:G:145:SER:O	2.12	0.50
1:A:6:VAL:HG22	2:B:16:TYR:HD1	1.77	0.50
2:D:148:ILE:HB	2:D:156:GLN:HB3	1.93	0.49
1:C:70:LEU:HD13	2:D:9:TYR:HB2	1.93	0.49
4:H:186:ASP:HB2	4:H:203:LEU:HD12	1.94	0.49
2:B:129:VAL:HG21	2:B:159:VAL:HG21	1.95	0.49
2:D:26:LEU:HG	2:D:42:SER:HB3	1.93	0.48
4:F:230:LEU:HD22	4:F:243:PRO:HD2	1.95	0.48
3:G:177:VAL:HG22	3:G:188:ASN:OD1	2.14	0.48
4:H:133:VAL:O	4:H:243:PRO:HG3	2.15	0.47
3:E:178:LEU:HB3	4:F:184:CYS:HB2	1.98	0.46
1:A:122:LEU:HD23	1:A:127:SER:HA	1.98	0.46
4:F:143:PRO:HD2	4:F:214:TRP:CE2	2.50	0.46
2:D:10:GLN:HB2	2:D:31:ILE:HB	1.97	0.46
4:F:215:GLN:HA	4:F:255:ARG:O	2.17	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:G:58:THR:HG22	3:G:84:ARG:HD2	1.98	0.45
1:C:12:LEU:HD11	1:C:115:PRO:HG3	1.97	0.45
1:A:91:VAL:HG12	1:A:105:LEU:HD22	1.99	0.45
1:C:6:ALA:HB2	1:C:26:PHE:CD1	2.52	0.45
4:H:215:GLN:HA	4:H:255:ARG:O	2.17	0.44
2:D:67:VAL:O	2:D:71:THR:HG22	2.17	0.44
1:C:91:VAL:HG12	1:C:105:LEU:HD22	2.00	0.44
2:D:70:ARG:HD3	4:F:112(A):THR:HB	1.98	0.44
2:B:121:ASP:HA	2:B:154:THR:HB	1.98	0.44
4:H:55:GLN:HB3	4:H:67:ALA:HB3	2.00	0.43
1:A:6:VAL:HG22	2:B:16:TYR:CD1	2.53	0.43
3:E:23:CYS:HB2	3:E:41:TRP:CZ2	2.54	0.43
1:C:118:ASN:HB2	1:C:166:GLU:HB2	1.99	0.43
4:F:148:ILE:HG23	4:F:211:ALA:HB1	2.01	0.43
3:E:157:ASP:OD1	3:E:159:GLN:HG3	2.18	0.43
3:G:23:CYS:HB2	3:G:41:TRP:CZ2	2.54	0.43
1:A:12:LEU:HD11	1:A:115:PRO:HG3	2.01	0.42
4:F:193:GLN:HB3	4:F:196:LEU:HD13	2.00	0.42
4:F:186:ASP:HB2	4:F:203:LEU:HD12	2.00	0.42
1:A:118:ASN:HB2	1:A:166:GLU:HB2	2.01	0.42
3:E:176:CYS:C	4:F:184:CYS:SG	2.99	0.41
3:E:46:PRO:HA	3:E:47:SER:HA	1.83	0.41
1:A:107:CYS:HB2	1:A:121:TRP:CH2	2.55	0.41
2:D:120:THR:HG22	2:D:156:GLN:HG3	2.02	0.41
1:C:85:GLU:HG3	2:D:34:ARG:HH21	1.85	0.41
3:G:171:TYR:O	3:G:192:ALA:HA	2.21	0.40
1:C:122:LEU:HD23	1:C:127:SER:HA	2.03	0.40
3:E:120:THR:HG22	8:E:303:HOH:O	2.17	0.40
1:A:61:THR:HG21	3:G:113:SER:CB	2.51	0.40
1:C:61:THR:HG21	3:E:113:SER:CB	2.51	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	180/192 (94%)	175 (97%)	5 (3%)	0	100	100
1	C	180/192 (94%)	177 (98%)	3 (2%)	0	100	100
2	B	161/213 (76%)	152 (94%)	8 (5%)	1 (1%)	30	59
2	D	164/213 (77%)	153 (93%)	10 (6%)	1 (1%)	30	59
3	E	196/203 (97%)	181 (92%)	15 (8%)	0	100	100
3	G	196/203 (97%)	184 (94%)	12 (6%)	0	100	100
4	F	242/245 (99%)	232 (96%)	10 (4%)	0	100	100
4	H	242/245 (99%)	237 (98%)	5 (2%)	0	100	100
5	I	13/18 (72%)	13 (100%)	0	0	100	100
5	J	14/18 (78%)	14 (100%)	0	0	100	100
All	All	1588/1742 (91%)	1518 (96%)	68 (4%)	2 (0%)	56	83

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	141	GLY
2	D	141	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	161/176 (92%)	157 (98%)	4 (2%)	55	84
1	C	162/176 (92%)	156 (96%)	6 (4%)	41	72
2	B	140/189 (74%)	132 (94%)	8 (6%)	25	53
2	D	139/189 (74%)	130 (94%)	9 (6%)	21	46
3	E	172/185 (93%)	162 (94%)	10 (6%)	25	52
3	G	168/185 (91%)	156 (93%)	12 (7%)	18	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	F	203/213 (95%)	184 (91%)	19 (9%)	11	25
4	H	204/213 (96%)	196 (96%)	8 (4%)	39	70
5	I	12/14 (86%)	12 (100%)	0	100	100
5	J	12/14 (86%)	12 (100%)	0	100	100
All	All	1373/1554 (88%)	1297 (94%)	76 (6%)	27	55

All (76) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	GLN
1	A	90	THR
1	A	174	LEU
1	A	175	LEU
2	B	70	ARG
2	B	71	THR
2	B	127	ILE
2	B	157	ILE
2	B	159	VAL
2	B	161	LEU
2	B	169	ASP
2	B	176	GLU
1	C	0	ILE
1	C	50	ARG
1	C	57	GLN
1	C	90	THR
1	C	174	LEU
1	C	175	LEU
2	D	71	THR
2	D	96	GLU
2	D	157	ILE
2	D	159	VAL
2	D	161	LEU
2	D	163	MET
2	D	169	ASP
2	D	174	HIS
2	D	176	GLU
3	E	15	GLU
3	E	27	THR
3	E	120	THR
3	E	125	THR
3	E	144	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	158	SER
3	E	159	GLN
3	E	164	GLN
3	E	169	ASP
3	E	209	ILE
4	F	9	LYS
4	F	10	HIS
4	F	23	CYS
4	F	24	SER
4	F	39	VAL
4	F	46	LEU
4	F	50	LEU
4	F	68	LYS
4	F	72	LEU
4	F	75	ARG
4	F	92	SER
4	F	104	CYS
4	F	123	ARG
4	F	190	LEU
4	F	193	GLN
4	F	206	ARG
4	F	230	LEU
4	F	231	SER
4	F	239	ASP
3	G	3	LYS
3	G	15	GLU
3	G	27	THR
3	G	47	SER
3	G	115	GLN
3	G	127	ILE
3	G	157	ASP
3	G	164	GLN
3	G	169	ASP
3	G	183	MET
3	G	208	SER
3	G	209	ILE
4	H	23	CYS
4	H	37	LEU
4	H	46	LEU
4	H	50	LEU
4	H	68	LYS
4	H	104	CYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	206	ARG
4	H	231	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	174	HIS
3	E	67	ASN
3	G	67	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](#)

10 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	A	1000	1	14,14,15	0.31	0	15,19,21	1.22	1 (6%)
6	NAG	B	1000	2,7,6	14,14,15	0.27	0	15,19,21	1.02	1 (6%)
6	NAG	B	1001	6	14,14,15	0.30	0	15,19,21	0.53	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
7	FUC	B	1002	6	10,10,11	0.35	0	14,14,16	1.16	1 (7%)
7	FUC	B	1003	6	10,10,11	0.43	0	14,14,16	2.61	2 (14%)
6	NAG	C	1000	1	14,14,15	0.32	0	15,19,21	1.22	1 (6%)
6	NAG	D	1000	2,7,6	14,14,15	0.33	0	15,19,21	0.98	1 (6%)
6	NAG	D	1001	6	14,14,15	0.30	0	15,19,21	0.46	0
7	FUC	D	1002	6	10,10,11	0.37	0	14,14,16	1.19	1 (7%)
7	FUC	D	1003	6	10,10,11	0.34	0	14,14,16	1.26	1 (7%)

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	1000	1	-	0/6/23/26	0/1/1/1
6	NAG	B	1000	2,7,6	-	0/6/23/26	0/1/1/1
6	NAG	B	1001	6	-	0/6/23/26	0/1/1/1
7	FUC	B	1002	6	-	0/0/17/20	0/1/1/1
7	FUC	B	1003	6	-	0/0/17/20	0/1/1/1
6	NAG	C	1000	1	-	0/6/23/26	0/1/1/1
6	NAG	D	1000	2,7,6	-	0/6/23/26	0/1/1/1
6	NAG	D	1001	6	-	0/6/23/26	0/1/1/1
7	FUC	D	1002	6	-	0/0/17/20	0/1/1/1
7	FUC	D	1003	6	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1003	FUC	C1-C2-C3	3.38	113.53	109.54
6	D	1000	NAG	C1-O5-C5	3.48	116.67	112.25
6	B	1000	NAG	C1-O5-C5	3.56	116.76	112.25
7	B	1002	FUC	C1-O5-C5	3.78	118.22	112.38
7	D	1002	FUC	C1-O5-C5	3.98	118.52	112.38
6	C	1000	NAG	C1-O5-C5	4.18	117.55	112.25
7	D	1003	FUC	C1-O5-C5	4.19	118.86	112.38
6	A	1000	NAG	C1-O5-C5	4.40	117.83	112.25
7	B	1003	FUC	C1-O5-C5	8.51	125.53	112.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	1003	FUC	1	0
7	D	1003	FUC	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, 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	182/192 (94%)	-0.06	2 (1%) 82 83	35, 64, 109, 140	0
1	C	182/192 (94%)	-0.27	0 100 100	36, 65, 95, 116	0
2	B	169/213 (79%)	0.77	31 (18%) 2 1	36, 75, 156, 176	0
2	D	172/213 (80%)	0.42	27 (15%) 3 2	40, 73, 149, 164	0
3	E	198/203 (97%)	-0.19	2 (1%) 84 85	29, 54, 101, 114	0
3	G	198/203 (97%)	-0.24	4 (2%) 68 69	27, 54, 96, 110	0
4	F	244/245 (99%)	-0.04	8 (3%) 50 50	40, 70, 116, 145	0
4	H	244/245 (99%)	-0.26	1 (0%) 93 94	30, 57, 88, 111	0
5	I	15/18 (83%)	-0.17	0 100 100	37, 50, 76, 92	0
5	J	16/18 (88%)	0.03	0 100 100	39, 47, 87, 93	0
All	All	1620/1742 (92%)	-0.01	75 (4%) 36 35	27, 64, 119, 176	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	171	TYR	9.4
2	B	170	VAL	8.4
2	D	171	TYR	6.4
2	B	160	MET	5.9
2	D	188	TRP	5.3
2	B	188	TRP	5.2
2	D	170	VAL	5.0
2	B	174	HIS	5.0
2	B	173	CYS	4.8
2	B	142	VAL	4.7
2	B	187	GLU	4.7
2	B	129	VAL	4.6
2	B	116	VAL	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	131	TRP	4.4
2	D	161	LEU	4.2
2	B	186	VAL	4.2
2	B	114	LEU	4.1
2	D	184	ILE	4.1
2	D	115	LEU	4.0
4	F	239	ASP	3.9
1	A	157	ALA	3.9
2	B	159	VAL	3.8
2	B	172	THR	3.8
2	B	161	LEU	3.7
2	B	175	VAL	3.6
2	B	115	LEU	3.6
2	D	138	GLU	3.6
2	D	131	TRP	3.5
2	B	103	PRO	3.4
1	A	158	ASP	3.4
3	G	197	SER	3.4
2	D	172	THR	3.3
2	D	114	LEU	3.2
3	E	197	SER	3.2
2	B	169	ASP	3.2
2	D	187	GLU	3.0
2	D	181	GLN	3.0
2	D	101	ILE	2.9
4	F	236	TRP	2.9
2	D	103	PRO	2.9
2	B	144	SER	2.7
4	F	231	SER	2.7
2	B	163	MET	2.7
2	D	129	VAL	2.7
2	B	117	CYS	2.6
2	D	104	SER	2.6
4	F	232	GLU	2.6
2	D	173	CYS	2.5
2	D	144	SER	2.5
3	G	146	SER	2.5
3	G	198	ASP	2.5
4	F	243	PRO	2.5
2	D	2	ASP	2.4
2	B	143	VAL	2.4
2	D	180	LEU	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	H	1	ASP	2.3
2	D	142	VAL	2.3
2	D	174	HIS	2.3
2	B	101	ILE	2.3
2	B	122	PHE	2.2
2	D	162	GLU	2.2
4	F	235	GLU	2.2
2	B	185	ILE	2.2
2	D	175	VAL	2.2
3	E	182	SER	2.1
2	D	169	ASP	2.1
2	B	99	VAL	2.1
2	B	100	THR	2.1
4	F	241	ALA	2.1
3	G	199	PHE	2.1
2	D	183	PRO	2.1
2	D	189	ARG	2.0
4	F	128	GLU	2.0
2	B	176	GLU	2.0
2	B	2	ASP	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, 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	1000	14/15	0.94	0.21	2.49	88,93,97,97	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
6	NAG	C	1000	14/15	0.84	0.25	2.40	112,117,121,121	0
7	FUC	B	1003	10/11	0.83	0.36	-	122,125,126,127	0
6	NAG	B	1001	14/15	0.94	0.20	-	112,121,127,129	0
6	NAG	B	1000	14/15	0.83	0.14	-	108,113,119,125	0
6	NAG	D	1000	14/15	0.87	0.14	-	88,95,104,106	0
6	NAG	D	1001	14/15	0.92	0.17	-	100,107,115,118	0
7	FUC	D	1003	10/11	0.89	0.35	-	109,111,113,115	0
7	FUC	B	1002	10/11	0.92	0.32	-	119,123,131,131	0
7	FUC	D	1002	10/11	0.86	0.21	-	112,118,121,127	0

6.5 Other polymers [i](#)

There are no such residues in this entry.