



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:29 PM GMT

PDB ID : 4OZI  
Title : S2 protein complex  
Authors : Petersen, J.; Reid, H.H.; Rossjohn, J.  
Deposited on : 2014-02-16  
Resolution : 3.20 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 (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

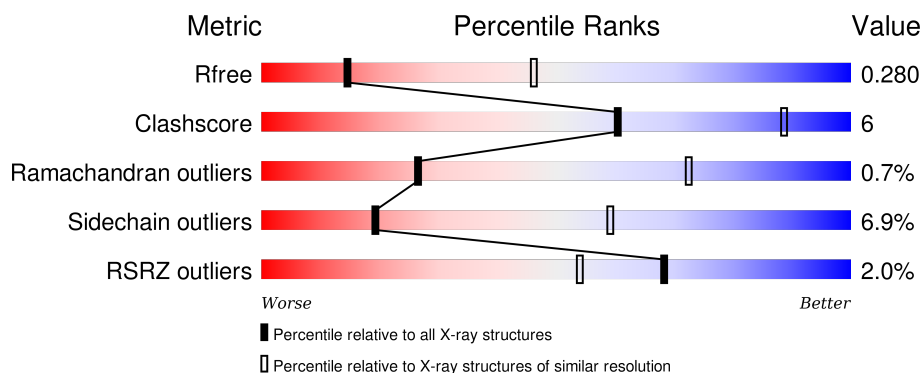
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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	191	<div> <div>83%</div> <div>12% • 5%</div> </div>
1	C	191	<div> <div>82%</div> <div>12% • 5%</div> </div>
2	B	213	<div> <div>7%</div> <div>73%</div> <div>12%</div> <div>15%</div> </div>
2	D	213	<div> <div>3%</div> <div>74%</div> <div>10%</div> <div>15%</div> </div>
3	E	207	<div> <div>%</div> <div>76%</div> <div>18%</div> <div>• •</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	207	<div><div></div><div>83%</div><div>12%</div><div></div><div></div></div>
4	F	244	<div><div>4%</div><div></div><div>79%</div><div>17%</div><div></div><div></div></div>
4	H	244	<div><div>%</div><div></div><div>78%</div><div>18%</div><div></div><div></div></div>
5	I	13	<div><div></div><div>85%</div><div>15%</div><div></div><div></div></div>
5	J	13	<div><div></div><div>69%</div><div>15%</div><div>15%</div><div></div></div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 12849 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
			1445	931	236	275	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	180	Total	C	N	O	S	0	0	0
			1462	926	257	272	7			
2	D	181	Total	C	N	O	S	0	0	0
			1464	926	257	274	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, s2, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	199	Total	C	N	O	S	0	0	0
			1537	970	255	304	8			
3	G	199	Total	C	N	O	S	0	0	0
			1524	962	254	300	8			

- Molecule 4 is a protein called T-cell receptor, s2, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	243	Total	C	N	O	S	0	0	0
			1882	1193	319	361	9			
4	H	243	Total	C	N	O	S	0	0	0
			1882	1193	318	362	9			

- Molecule 5 is a protein called deamidated Gliadin-alpha1 peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	11	Total	C	N	O	0	0	0
			89	62	12	15			
5	J	11	Total	C	N	O	0	0	0
			89	62	12	15			

- 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	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	B	1	Total	Ca	0	0
			1	1		
7	D	1	Total	Ca	0	0
			1	1		

### 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: HLA class II histocompatibility antigen, DQ alpha 1 chain

Chain A: 



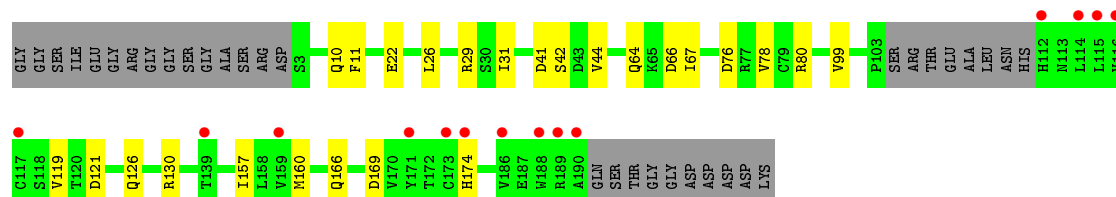
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain

Chain C: 



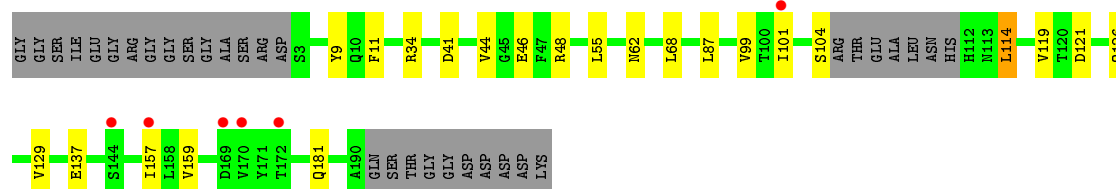
- Molecule 2: HLA class II histocompatibility antigen, DQ beta 1 chain

Chain B: 



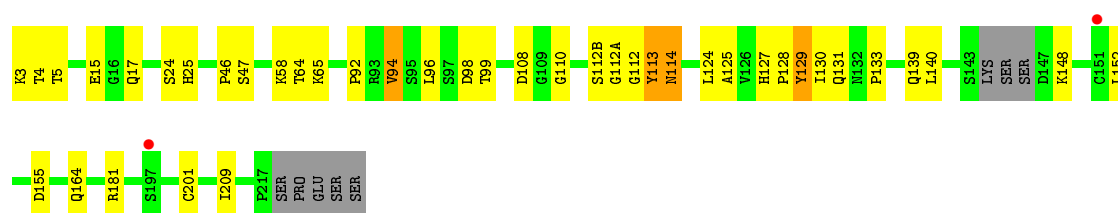
- Molecule 2: HLA class II histocompatibility antigen, DQ beta 1 chain

Chain D: 



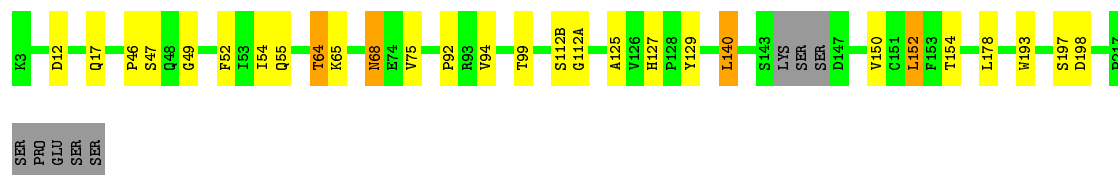
- Molecule 3: T-cell receptor, s2, alpha chain

Chain E: 



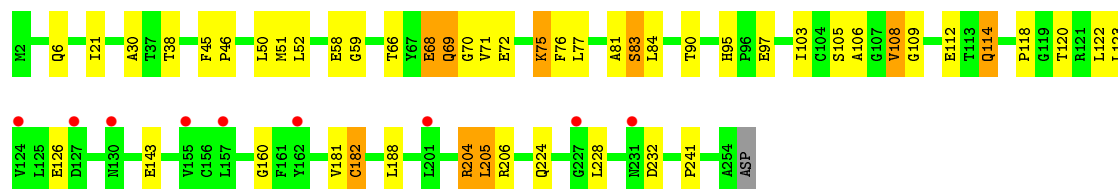
- Molecule 3: T-cell receptor, s2, alpha chain

Chain G: 83% 12%



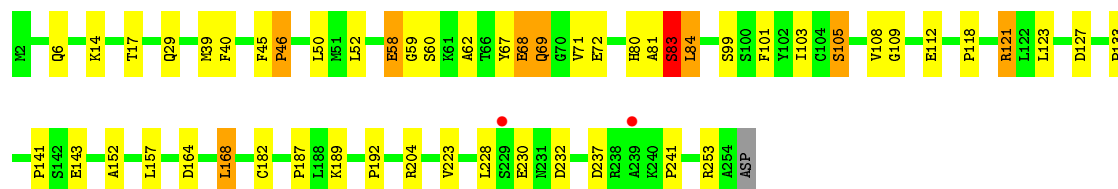
- Molecule 4: T-cell receptor, s2, beta chain

Chain F: 79% 17% 4%



- Molecule 4: T-cell receptor, s2, beta chain

Chain H: 78% 18% 4%



- Molecule 5: deamidated Gliadin-alpha1 peptide

Chain I: 85% 15%



- Molecule 5: deamidated Gliadin-alpha1 peptide

Chain J: 69% 15% 15%





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	233.03Å 142.24Å 101.09Å 90.00° 109.79° 90.00°	Depositor
Resolution (Å)	46.85 – 3.20 95.12 – 3.20	Depositor EDS
% Data completeness (in resolution range)	96.0 (46.85-3.20) 96.1 (95.12-3.20)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 3.19Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.212 , 0.243 0.241 , 0.280	Depositor DCC
$R_{free}$ test set	2542 reflections (5.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.8	Xtriage
Anisotropy	0.733	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 43.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.39$ , $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	0 of 49219 reflections	Xtriage
$F_o, F_c$ correlation	0.84	EDS
Total number of atoms	12849	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.58% 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, 0.375 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	0.40	0/1487	0.57	0/2031
1	C	0.39	0/1487	0.58	0/2031
2	B	0.38	0/1495	0.57	0/2034
2	D	0.39	0/1497	0.56	0/2038
3	E	0.57	1/1573 (0.1%)	0.66	1/2141 (0.0%)
3	G	0.51	0/1560	0.62	1/2125 (0.0%)
4	F	0.49	0/1934	0.59	0/2640
4	H	0.45	0/1934	0.56	0/2641
5	I	0.49	0/95	0.60	0/133
5	J	0.42	0/95	0.52	0/133
All	All	0.45	1/13157 (0.0%)	0.59	2/17947 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	58	LYS	C-N	-10.38	1.10	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	127	HIS	C-N-CD	-9.00	100.80	120.60
3	G	127	HIS	C-N-CD	5.90	140.78	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	1397	8	0
1	C	1445	0	1397	10	0
2	B	1462	0	1416	9	0
2	D	1464	0	1410	9	0
3	E	1537	0	1433	23	0
3	G	1524	0	1410	15	0
4	F	1882	0	1761	30	0
4	H	1882	0	1762	41	2
5	I	89	0	79	0	0
5	J	89	0	79	2	0
6	A	14	0	13	0	0
6	C	14	0	13	0	0
7	B	1	0	0	0	0
7	D	1	0	0	0	0
All	All	12849	0	12170	139	2

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 6.

All (139) 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:68:GLU:C	4:H:69:GLN:NE2	1.95	1.20
4:F:72:GLU:HB2	4:F:75:LYS:HB2	1.19	1.17
4:F:72:GLU:HG2	4:F:75:LYS:CD	1.76	1.15
4:F:72:GLU:CG	4:F:75:LYS:HD3	1.76	1.15
4:H:68:GLU:CA	4:H:69:GLN:NE2	2.18	1.06
4:F:72:GLU:CB	4:F:75:LYS:HB2	1.87	1.04
4:F:72:GLU:HG2	4:F:75:LYS:HD3	0.96	0.96
4:F:72:GLU:OE1	4:F:72:GLU:N	2.01	0.93
4:H:68:GLU:HA	4:H:69:GLN:HE22	1.33	0.92
4:H:68:GLU:C	4:H:69:GLN:CD	2.35	0.85
4:H:68:GLU:O	4:H:69:GLN:CD	2.16	0.83
3:E:15:GLU:O	3:E:129:TYR:CE2	2.32	0.81
4:H:68:GLU:HA	4:H:69:GLN:NE2	1.88	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:F:45:PHE:HB3	4:F:46:PRO:HD2	1.66	0.78
4:F:72:GLU:HB2	4:F:75:LYS:CB	2.08	0.78
4:H:29:GLN:HB3	4:H:84:LEU:HD22	1.65	0.78
3:E:108:ASP:OD1	3:E:112:GLY:HA3	1.85	0.76
3:E:15:GLU:O	3:E:129:TYR:HE2	1.69	0.74
4:F:108:VAL:N	4:F:109:GLY:HA2	2.04	0.71
4:F:50:LEU:HD21	4:F:103:ILE:HD12	1.73	0.70
4:H:68:GLU:HG2	4:H:69:GLN:HE21	1.56	0.70
4:H:108:VAL:N	4:H:109:GLY:HA2	2.07	0.69
4:H:50:LEU:HD21	4:H:103:ILE:HD12	1.74	0.69
4:H:68:GLU:O	4:H:69:GLN:NE2	2.26	0.69
3:G:99:THR:HG23	3:G:125:ALA:HA	1.75	0.69
3:G:112(B):SER:H	3:G:112(A):GLY:HA2	1.61	0.66
4:F:68:GLU:O	4:F:71:VAL:HG22	1.96	0.66
1:A:122:LEU:HD21	1:A:164:LYS:HB2	1.79	0.65
3:E:108:ASP:CG	3:E:112:GLY:HA3	2.17	0.65
4:H:52:LEU:HD23	4:H:68:GLU:HG3	1.78	0.64
4:F:68:GLU:O	4:F:68:GLU:HG3	1.95	0.64
3:E:108:ASP:OD1	3:E:112:GLY:CA	2.45	0.64
4:H:68:GLU:CG	4:H:69:GLN:HE21	2.11	0.64
1:C:45:LEU:HB3	1:C:48:LEU:HD12	1.80	0.63
4:H:68:GLU:HG2	4:H:69:GLN:NE2	2.12	0.63
1:A:63:ILE:HA	1:A:66:LEU:HD12	1.81	0.63
3:G:140:LEU:HD12	4:H:141:PRO:HA	1.80	0.62
3:E:4:THR:HG22	3:E:25:HIS:HB3	1.79	0.62
4:H:81:ALA:O	4:H:83:SER:OG	2.13	0.62
4:F:72:GLU:CB	4:F:75:LYS:HD3	2.28	0.62
4:F:81:ALA:O	4:F:83:SER:HB3	1.99	0.61
4:H:168:LEU:HG	4:H:223:VAL:HG22	1.82	0.61
3:G:150:VAL:HG12	3:G:193:TRP:HB3	1.83	0.60
3:E:130:ILE:HG22	3:E:133:PRO:HG3	1.82	0.60
4:H:68:GLU:CA	4:H:69:GLN:HE21	2.15	0.59
3:E:64:THR:HG22	3:E:65:LYS:H	1.66	0.59
3:E:64:THR:HG22	3:E:65:LYS:N	2.18	0.58
4:H:6:GLN:HB2	4:H:118:PRO:HD2	1.84	0.58
4:F:66:THR:HG22	4:F:66:THR:O	2.03	0.58
4:H:68:GLU:CB	4:H:69:GLN:HE21	2.17	0.58
2:D:62:ASN:HA	2:D:68:LEU:HD21	1.86	0.57
3:E:113:TYR:CD2	3:E:114:ASN:ND2	2.73	0.57
4:H:68:GLU:O	4:H:69:GLN:OE1	2.24	0.56
3:E:113:TYR:CE2	3:E:114:ASN:ND2	2.73	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:141:PRO:HG2	4:H:152:ALA:HB1	1.86	0.56
3:E:130:ILE:CG2	3:E:133:PRO:HG3	2.36	0.56
2:B:41:ASP:HB3	2:B:44:VAL:HG22	1.88	0.55
1:C:66:LEU:HD13	2:D:9:TYR:HD2	1.71	0.54
3:E:98:ASP:O	3:E:124:LEU:HD23	2.08	0.54
4:H:40:PHE:HB2	4:H:105:SER:HB2	1.88	0.54
2:D:129:VAL:HG11	2:D:159:VAL:HG21	1.90	0.54
1:C:99:LEU:HA	1:C:155:PRO:HB2	1.90	0.54
4:H:68:GLU:CB	4:H:69:GLN:NE2	2.71	0.54
1:C:118:ASN:HB2	1:C:166:GLU:HB2	1.90	0.54
4:F:77:LEU:HB2	4:F:90:THR:HB	1.92	0.51
2:D:104:SER:HB3	2:D:114:LEU:HB3	1.92	0.51
4:H:133:PRO:HD3	4:H:241:PRO:HB3	1.94	0.50
4:H:58:GLU:HG3	4:H:83:SER:H	1.77	0.49
1:A:118:ASN:HB2	1:A:166:GLU:HB2	1.94	0.49
3:G:129:TYR:C	3:G:129:TYR:CD2	2.86	0.49
2:B:76:ASP:HA	2:B:80:ARG:HB2	1.95	0.49
3:E:113:TYR:C	3:E:113:TYR:CD1	2.85	0.49
3:E:15:GLU:HG3	3:E:129:TYR:CE2	2.47	0.48
2:B:78:VAL:HG21	5:J:6:GLN:HB3	1.95	0.48
1:C:11:ASN:HB2	2:D:11:PHE:HB3	1.95	0.48
3:G:17:GLN:O	3:G:94:VAL:HG22	2.14	0.48
4:F:72:GLU:HB3	4:F:75:LYS:HB2	1.88	0.48
4:F:95:HIS:HD2	4:F:97:GLU:HB2	1.79	0.47
4:F:58:GLU:HA	4:F:59:GLY:HA2	1.46	0.47
4:H:71:VAL:O	4:H:71:VAL:HG23	2.14	0.47
3:E:17:GLN:O	3:E:94:VAL:HG22	2.15	0.46
3:E:99:THR:HG23	3:E:125:ALA:HA	1.97	0.46
3:G:64:THR:HG22	3:G:65:LYS:H	1.80	0.46
4:F:6:GLN:HB2	4:F:118:PRO:HD2	1.98	0.46
4:H:62:ALA:CB	4:H:67:TYR:CE1	2.99	0.46
4:H:69:GLN:NE2	4:H:69:GLN:N	2.60	0.46
1:C:57:GLN:HA	1:C:60:LEU:HD12	1.98	0.46
4:H:39:MET:HB2	4:H:80:HIS:CD2	2.51	0.46
1:A:104:ILE:HG12	1:A:152:THR:HG22	1.98	0.46
1:A:11:ASN:HB2	2:B:11:PHE:HB3	1.98	0.45
1:C:66:LEU:HD13	2:D:9:TYR:CD2	2.51	0.45
4:F:228:LEU:HD23	4:F:232:ASP:HB3	1.99	0.45
4:H:58:GLU:HG2	4:H:59:GLY:HA2	1.98	0.44
2:B:99:VAL:HG12	2:B:119:VAL:HG22	2.00	0.44
3:E:112(B):SER:H	3:E:112(A):GLY:HA2	1.81	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:F:21:ILE:HG12	4:F:120:THR:HG21	2.00	0.44
4:F:105:SER:HB2	4:F:114:GLN:HB2	2.00	0.44
3:G:46:PRO:HA	3:G:47:SER:HA	1.67	0.44
3:G:49:GLY:HA2	4:H:103:ILE:HD13	2.00	0.43
2:D:99:VAL:HG12	2:D:119:VAL:HG22	2.00	0.43
3:G:140:LEU:HD21	3:G:152:LEU:HB2	1.99	0.43
2:D:41:ASP:HB3	2:D:44:VAL:HG22	2.00	0.43
1:A:11:ASN:HB3	1:A:66:LEU:HD11	2.01	0.43
1:A:59:ALA:O	1:A:63:ILE:HG12	2.19	0.43
2:B:130:ARG:HB2	2:B:174:HIS:HB3	2.00	0.43
1:C:46:PRO:HA	1:C:49:ARG:HG3	2.00	0.43
3:E:5:THR:HG23	3:E:24:SER:HB3	2.01	0.43
3:G:197:SER:HA	3:G:198:ASP:HA	1.70	0.43
4:H:164:ASP:HB2	4:H:187:PRO:HG2	2.01	0.43
2:B:10:GLN:HB2	2:B:31:ILE:HB	2.01	0.43
3:G:152:LEU:HD22	3:G:154:THR:HB	2.00	0.42
2:D:44:VAL:HG23	2:D:46:GLU:H	1.83	0.42
1:A:98:THR:HB	1:A:101:GLN:HB2	2.00	0.42
3:E:110:GLY:O	3:E:112:GLY:N	2.40	0.42
4:H:69:GLN:N	4:H:69:GLN:CD	2.72	0.42
4:H:228:LEU:HD12	4:H:232:ASP:HB3	2.00	0.42
4:F:75:LYS:HB3	4:F:76:PHE:CD2	2.55	0.42
4:H:189:LYS:HB2	4:H:192:PRO:HG3	2.02	0.42
4:F:228:LEU:HD13	4:F:241:PRO:HD2	2.00	0.42
5:J:6:GLN:HA	5:J:7:PRO:HD3	1.95	0.42
2:B:26:LEU:HB3	2:B:42:SER:HB3	2.01	0.42
4:H:101:PHE:HD1	4:H:121:ARG:HG3	1.85	0.42
3:G:52:PHE:HE1	3:G:55:GLN:HB2	1.84	0.42
4:F:69:GLN:HA	4:F:70:GLY:HA2	1.81	0.41
4:F:30:ALA:HB2	4:F:106:ALA:HB1	2.03	0.41
3:G:178:LEU:HB3	4:H:182:CYS:HB2	2.03	0.41
4:F:182:CYS:HB3	4:F:204:ARG:HD2	2.02	0.41
4:H:45:PHE:HB3	4:H:46:PRO:HD2	2.02	0.41
3:E:130:ILE:HG22	3:E:133:PRO:CG	2.49	0.41
2:B:64:GLN:HB2	2:B:67:ILE:HD12	2.03	0.41
4:H:81:ALA:O	4:H:83:SER:CB	2.69	0.40
4:H:14:LYS:HE3	4:H:127:ASP:HA	2.02	0.40
3:E:46:PRO:HA	3:E:47:SER:HA	1.76	0.40
3:G:54:ILE:HD12	3:G:68:ASN:HB2	2.03	0.40
3:E:15:GLU:HG3	3:E:129:TYR:HE2	1.86	0.40
4:F:71:VAL:HG23	4:F:71:VAL:O	2.20	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:72:SER:HA	1:C:75:LYS:HE3	2.03	0.40
4:F:181:VAL:HG22	4:F:205:LEU:HD12	2.03	0.40
1:C:108:LEU:HD23	1:C:148:ILE:HG13	2.02	0.40

All (2) 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)
4:H:237:ASP:OD1	4:H:253:ARG:NH2[2_859]	1.95	0.25
4:H:237:ASP:OD2	4:H:253:ARG:NH2[2_859]	2.08	0.12

## 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%)	170 (95%)	9 (5%)	0	100	100
1	C	179/191 (94%)	172 (96%)	7 (4%)	0	100	100
2	B	176/213 (83%)	164 (93%)	11 (6%)	1 (1%)	30	75
2	D	177/213 (83%)	170 (96%)	6 (3%)	1 (1%)	30	75
3	E	195/207 (94%)	181 (93%)	11 (6%)	3 (2%)	13	55
3	G	195/207 (94%)	183 (94%)	11 (6%)	1 (0%)	34	78
4	F	241/244 (99%)	225 (93%)	14 (6%)	2 (1%)	24	69
4	H	241/244 (99%)	223 (92%)	14 (6%)	4 (2%)	11	52
5	I	9/13 (69%)	8 (89%)	1 (11%)	0	100	100
5	J	9/13 (69%)	8 (89%)	1 (11%)	0	100	100
All	All	1601/1736 (92%)	1504 (94%)	85 (5%)	12 (1%)	26	72

All (12) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	121	ASP
3	E	92	PRO
4	F	83	SER
4	F	160	GLY
3	G	92	PRO
4	H	83	SER
4	H	46	PRO
4	H	99	SER
3	E	128	PRO
2	D	121	ASP
4	H	230	GLU
3	E	94	VAL

### 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%)	153 (93%)	12 (7%)	17	57
1	C	165/174 (95%)	155 (94%)	10 (6%)	23	64
2	B	161/188 (86%)	153 (95%)	8 (5%)	30	71
2	D	161/188 (86%)	151 (94%)	10 (6%)	23	64
3	E	169/185 (91%)	154 (91%)	15 (9%)	12	44
3	G	165/185 (89%)	159 (96%)	6 (4%)	42	79
4	F	202/214 (94%)	182 (90%)	20 (10%)	10	38
4	H	203/214 (95%)	187 (92%)	16 (8%)	15	53
5	I	10/12 (83%)	10 (100%)	0	100	100
5	J	10/12 (83%)	10 (100%)	0	100	100
All	All	1411/1546 (91%)	1314 (93%)	97 (7%)	19	59

All (97) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	LEU
1	A	38	ARG

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	44	CYS
1	A	52	ARG
1	A	88	GLU
1	A	122	LEU
1	A	153	LEU
1	A	154	LEU
1	A	158	GLU
1	A	171	ASP
1	A	172	LYS
1	A	174	LEU
2	B	22	GLU
2	B	29	ARG
2	B	66	ASP
2	B	126	GLN
2	B	157	ILE
2	B	160	MET
2	B	166	GLN
2	B	169	ASP
1	C	8	SER
1	C	12	LEU
1	C	52	ARG
1	C	66	LEU
1	C	73	LEU
1	C	108	LEU
1	C	126	HIS
1	C	153	LEU
1	C	172	LYS
1	C	174	LEU
2	D	34	ARG
2	D	48	ARG
2	D	55	LEU
2	D	87	LEU
2	D	101	ILE
2	D	114	LEU
2	D	126	GLN
2	D	137	GLU
2	D	157	ILE
2	D	181	GLN
3	E	3	LYS
3	E	96	LEU
3	E	113	TYR
3	E	114	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	E	129	TYR
3	E	131	GLN
3	E	139	GLN
3	E	140	LEU
3	E	148	LYS
3	E	152	LEU
3	E	155	ASP
3	E	164	GLN
3	E	181	ARG
3	E	201	CYS
3	E	209	ILE
4	F	38	THR
4	F	51	MET
4	F	52	LEU
4	F	68	GLU
4	F	69	GLN
4	F	75	LYS
4	F	84	LEU
4	F	108	VAL
4	F	112	GLU
4	F	114	GLN
4	F	122	LEU
4	F	123	LEU
4	F	126	GLU
4	F	143	GLU
4	F	182	CYS
4	F	188	LEU
4	F	204	ARG
4	F	205	LEU
4	F	206	ARG
4	F	224	GLN
3	G	12	ASP
3	G	64	THR
3	G	68	ASN
3	G	75	VAL
3	G	140	LEU
3	G	152	LEU
4	H	17	THR
4	H	58	GLU
4	H	60	SER
4	H	68	GLU
4	H	69	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	H	72	GLU
4	H	83	SER
4	H	84	LEU
4	H	105	SER
4	H	112	GLU
4	H	121	ARG
4	H	123	LEU
4	H	143	GLU
4	H	157	LEU
4	H	168	LEU
4	H	204	ARG

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

Mol	Chain	Res	Type
1	C	14	GLN
3	E	114	ASN
3	E	131	GLN
4	F	95	HIS
4	F	217	ASN
4	H	69	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, 2 are monoatomic - leaving 2 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	1001	1	14,14,15	0.27	0	15,19,21	0.61	1 (6%)
6	NAG	C	1001	1	14,14,15	0.29	0	15,19,21	0.74	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	1001	1	-	0/6/23/26	0/1/1/1
6	NAG	C	1001	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( $^{\circ}$ )	Ideal( $^{\circ}$ )
6	A	1001	NAG	C1-O5-C5	2.06	114.87	112.25
6	C	1001	NAG	C1-O5-C5	2.64	115.60	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, 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	181/191 (94%)	0.02	0  	26, 49, 75, 106	0
1	C	181/191 (94%)	-0.02	0  	26, 44, 71, 108	0
2	B	180/213 (84%)	0.50	14 (7%)  	27, 56, 101, 129	0
2	D	181/213 (84%)	0.48	6 (3%)  	31, 54, 106, 134	0
3	E	199/207 (96%)	0.19	2 (1%)  	33, 57, 93, 109	0
3	G	199/207 (96%)	0.19	0  	26, 55, 84, 93	0
4	F	243/244 (99%)	0.41	9 (3%)  	32, 66, 91, 101	0
4	H	243/244 (99%)	0.26	2 (0%)  	33, 63, 88, 103	0
5	I	11/13 (84%)	0.66	0  	40, 44, 55, 58	0
5	J	11/13 (84%)	0.25	0  	34, 42, 45, 45	0
All	All	1629/1736 (93%)	0.26	33 (2%)  	26, 56, 92, 134	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	112	HIS	5.2
3	E	197	SER	3.7
2	D	170	VAL	3.5
2	B	188	TRP	3.5
2	B	139	THR	3.3
2	B	190	ALA	3.0
4	F	162	TYR	3.0
2	B	114	LEU	2.9
2	D	172	THR	2.7
4	H	229	SER	2.6
2	B	116	VAL	2.6
3	E	151	CYS	2.5
4	F	155	VAL	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	159	VAL	2.4
2	B	171	TYR	2.4
4	F	231	ASN	2.4
2	B	115	LEU	2.4
4	F	157	LEU	2.4
2	B	173	CYS	2.3
2	B	117	CYS	2.3
4	H	239	ALA	2.3
2	D	144	SER	2.3
4	F	127	ASP	2.2
2	D	169	ASP	2.2
4	F	124	VAL	2.2
2	D	101	ILE	2.2
2	B	186	VAL	2.1
4	F	201	LEU	2.1
4	F	130	ASN	2.0
4	F	227	GLY	2.0
2	B	174	HIS	2.0
2	D	157	ILE	2.0
2	B	189	ARG	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(Å <sup>2</sup> )	Q<0.9
6	NAG	C	1001	14/15	0.81	0.26	1.12	58,62,67,70	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NAG	A	1001	14/15	0.85	0.25	0.27	69,72,79,81	0
7	CA	D	301	1/1	0.95	0.11	-	21,21,21,21	0
7	CA	B	301	1/1	0.94	0.11	-	30,30,30,30	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.