Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 04:52 PM GMT

PDB ID	:	4GG6
Title	:	Protein complex
Authors	:	Broughton, S.E.; Theodossis, A.; Petersen, J.; Reid, H.H.; Rossjohn, J.
Deposited on	:	2012-08-06
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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

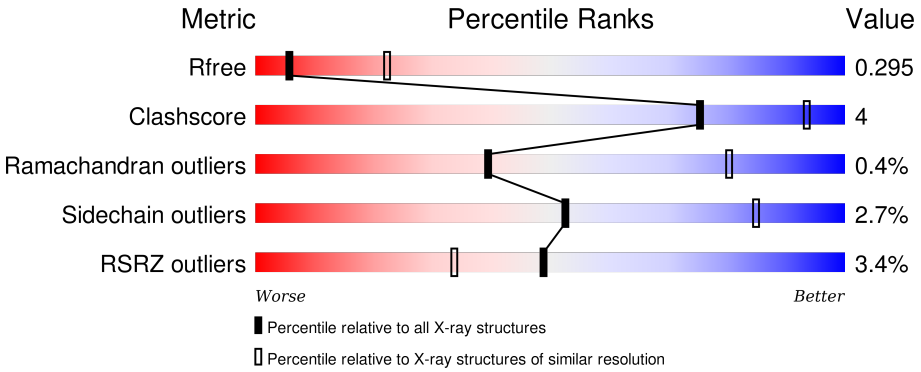
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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.



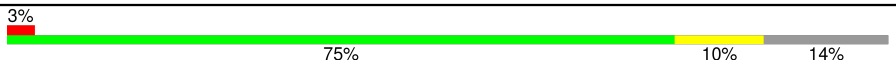
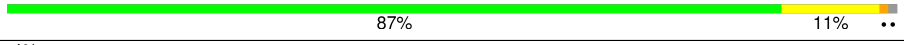
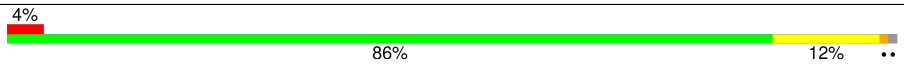
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1124 (3.24-3.16)
Clashscore	102246	1024 (3.22-3.18)
Ramachandran outliers	100387	1004 (3.22-3.18)
Sidechain outliers	100360	1003 (3.22-3.18)
RSRZ outliers	91569	1129 (3.24-3.16)

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 >=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 <=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	5% 85% 9% 6%
1	C	192	3% 81% 13% 6%
2	B	215	7% 65% 10% • 24%
2	D	215	% 69% 11% 20%
3	E	207	% 79% 12% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	207	
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	202	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11919 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	180	Total	C	N	O	S	0	0	0
			1376	890	223	261	2			
1	C	181	Total	C	N	O	S	0	0	0
			1398	903	227	266	2			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	SER	THR	SEE REMARK 999	UNP P01909
A	31	GLU	GLN	SEE REMARK 999	UNP P01909
A	37	GLU	GLY	SEE REMARK 999	UNP P01909
A	44	GLN	CYS	SEE REMARK 999	UNP P01909
A	47	LEU	VAL	SEE REMARK 999	UNP P01909
A	48	PHE	LEU	SEE REMARK 999	UNP P01909
A	50	ARG	GLN	SEE REMARK 999	UNP P01909
A	52	ARG	-	SEE REMARK 999	UNP P01909
A	72	ILE	SER	SEE REMARK 999	UNP P01909
A	73	VAL	LEU	SEE REMARK 999	UNP P01909
A	104	THR	ILE	SEE REMARK 999	UNP P01909
A	153	PHE	LEU	SEE REMARK 999	UNP P01909
A	158	ASP	GLU	SEE REMARK 999	UNP P01909
A	160	ILE	SER	SEE REMARK 999	UNP P01909
A	172	GLU	LYS	SEE REMARK 999	UNP P01909
A	182	SER	-	EXPRESSION TAG	UNP P01909
A	183	SER	-	EXPRESSION TAG	UNP P01909
A	184	ALA	-	EXPRESSION TAG	UNP P01909
A	185	ASP	-	EXPRESSION TAG	UNP P01909
A	186	LEU	-	EXPRESSION TAG	UNP P01909
A	187	VAL	-	EXPRESSION TAG	UNP P01909
A	188	PRO	-	EXPRESSION TAG	UNP P01909
A	189	ARG	-	EXPRESSION TAG	UNP P01909
C	23	SER	THR	SEE REMARK 999	UNP P01909
C	31	GLU	GLN	SEE REMARK 999	UNP P01909

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	37	GLU	GLY	SEE REMARK 999	UNP P01909
C	44	GLN	CYS	SEE REMARK 999	UNP P01909
C	47	LEU	VAL	SEE REMARK 999	UNP P01909
C	48	PHE	LEU	SEE REMARK 999	UNP P01909
C	50	ARG	GLN	SEE REMARK 999	UNP P01909
C	52	ARG	-	SEE REMARK 999	UNP P01909
C	72	ILE	SER	SEE REMARK 999	UNP P01909
C	73	VAL	LEU	SEE REMARK 999	UNP P01909
C	104	THR	ILE	SEE REMARK 999	UNP P01909
C	153	PHE	LEU	SEE REMARK 999	UNP P01909
C	158	ASP	GLU	SEE REMARK 999	UNP P01909
C	160	ILE	SER	SEE REMARK 999	UNP P01909
C	172	GLU	LYS	SEE REMARK 999	UNP P01909
C	182	SER	-	EXPRESSION TAG	UNP P01909
C	183	SER	-	EXPRESSION TAG	UNP P01909
C	184	ALA	-	EXPRESSION TAG	UNP P01909
C	185	ASP	-	EXPRESSION TAG	UNP P01909
C	186	LEU	-	EXPRESSION TAG	UNP P01909
C	187	VAL	-	EXPRESSION TAG	UNP P01909
C	188	PRO	-	EXPRESSION TAG	UNP P01909
C	189	ARG	-	EXPRESSION TAG	UNP P01909

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	163	Total	C	N	O	S	0	0	0
			1253	803	214	231	5			
2	D	173	Total	C	N	O	S	0	0	0
			1328	852	224	245	7			

There are 58 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	GLY	-	EXPRESSION TAG	UNP P01920
B	-13	GLY	-	EXPRESSION TAG	UNP P01920
B	-12	GLY	-	EXPRESSION TAG	UNP P01920
B	-11	GLY	-	EXPRESSION TAG	UNP P01920
B	-10	SER	-	EXPRESSION TAG	UNP P01920
B	-9	ILE	-	EXPRESSION TAG	UNP P01920
B	-8	GLU	-	EXPRESSION TAG	UNP P01920
B	-7	GLY	-	EXPRESSION TAG	UNP P01920
B	-6	ARG	-	EXPRESSION TAG	UNP P01920

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	GLY	-	EXPRESSION TAG	UNP P01920
B	-4	SER	-	EXPRESSION TAG	UNP P01920
B	-3	GLY	-	EXPRESSION TAG	UNP P01920
B	-2	GLY	-	EXPRESSION TAG	UNP P01920
B	-1	GLY	-	EXPRESSION TAG	UNP P01920
B	0	SER	-	EXPRESSION TAG	UNP P01920
B	13	GLY	ALA	SEE REMARK 999	UNP P01920
B	26	LEU	TYR	SEE REMARK 999	UNP P01920
B	45	GLY	GLU	SEE REMARK 999	UNP P01920
B	57	ALA	ASP	SEE REMARK 999	UNP P01920
B	167	ARG	HIS	SEE REMARK 999	UNP P01920
B	185	ILE	THR	SEE REMARK 999	UNP P01920
B	193	SER	-	EXPRESSION TAG	UNP P01920
B	194	SER	-	EXPRESSION TAG	UNP P01920
B	195	ALA	-	EXPRESSION TAG	UNP P01920
B	196	ASP	-	EXPRESSION TAG	UNP P01920
B	197	LEU	-	EXPRESSION TAG	UNP P01920
B	198	VAL	-	EXPRESSION TAG	UNP P01920
B	199	PRO	-	EXPRESSION TAG	UNP P01920
B	200	ARG	-	EXPRESSION TAG	UNP P01920
D	-14	GLY	-	EXPRESSION TAG	UNP P01920
D	-13	GLY	-	EXPRESSION TAG	UNP P01920
D	-12	GLY	-	EXPRESSION TAG	UNP P01920
D	-11	GLY	-	EXPRESSION TAG	UNP P01920
D	-10	SER	-	EXPRESSION TAG	UNP P01920
D	-9	ILE	-	EXPRESSION TAG	UNP P01920
D	-8	GLU	-	EXPRESSION TAG	UNP P01920
D	-7	GLY	-	EXPRESSION TAG	UNP P01920
D	-6	ARG	-	EXPRESSION TAG	UNP P01920
D	-5	GLY	-	EXPRESSION TAG	UNP P01920
D	-4	SER	-	EXPRESSION TAG	UNP P01920
D	-3	GLY	-	EXPRESSION TAG	UNP P01920
D	-2	GLY	-	EXPRESSION TAG	UNP P01920
D	-1	GLY	-	EXPRESSION TAG	UNP P01920
D	0	SER	-	EXPRESSION TAG	UNP P01920
D	13	GLY	ALA	SEE REMARK 999	UNP P01920
D	26	LEU	TYR	SEE REMARK 999	UNP P01920
D	45	GLY	GLU	SEE REMARK 999	UNP P01920
D	57	ALA	ASP	SEE REMARK 999	UNP P01920
D	167	ARG	HIS	SEE REMARK 999	UNP P01920
D	185	ILE	THR	SEE REMARK 999	UNP P01920
D	193	SER	-	EXPRESSION TAG	UNP P01920

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	194	SER	-	EXPRESSION TAG	UNP P01920
D	195	ALA	-	EXPRESSION TAG	UNP P01920
D	196	ASP	-	EXPRESSION TAG	UNP P01920
D	197	LEU	-	EXPRESSION TAG	UNP P01920
D	198	VAL	-	EXPRESSION TAG	UNP P01920
D	199	PRO	-	EXPRESSION TAG	UNP P01920
D	200	ARG	-	EXPRESSION TAG	UNP P01920

- Molecule 3 is a protein called T-CELL RECEPTOR, SP3.4 ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	187	Total	C	N	O	S	0	0	0
			1391	874	237	271	9			
3	G	177	Total	C	N	O	S	0	0	0
			1298	808	224	257	9			

- Molecule 4 is a protein called T-CELL RECEPTOR, SP3.4 BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	243	Total	C	N	O	S	0	0	0
			1827	1165	313	344	5			
4	H	243	Total	C	N	O	S	0	0	0
			1816	1158	311	342	5			

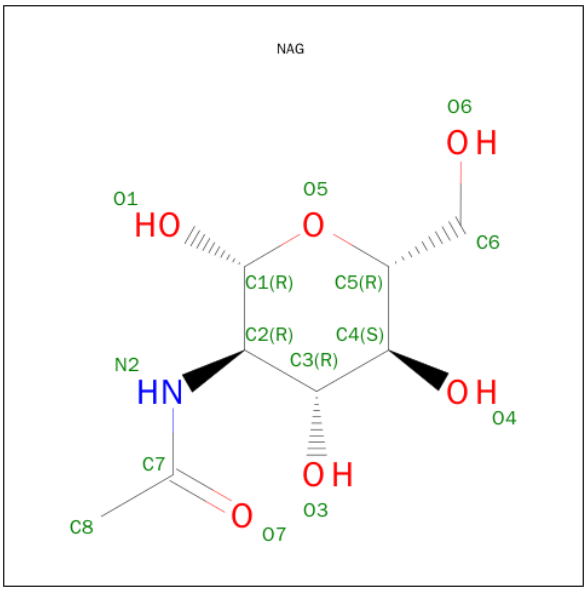
- Molecule 5 is a protein called Peptide from Alpha/beta-gliadin MM1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	13	Total	C	N	O	0	0	0
			95	56	16	23			
5	J	13	Total	C	N	O	0	0	0
			95	56	16	23			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	3	GLU	GLN	ENGINEERED MUTATION	UNP P18573
I	11	GLU	GLN	ENGINEERED MUTATION	UNP P18573
J	3	GLU	GLN	ENGINEERED MUTATION	UNP P18573
J	11	GLU	GLN	ENGINEERED MUTATION	UNP P18573

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



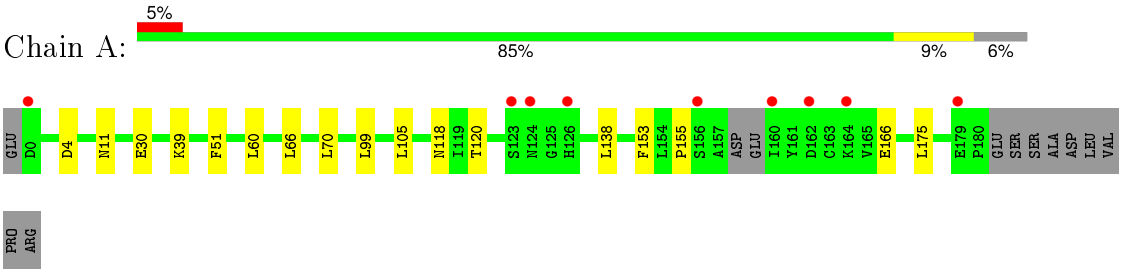
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O 14 8 1 5	0	0
6	A	1	Total C N O 14 8 1 5	0	0
6	C	1	Total C N O 14 8 1 5	0	0

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.

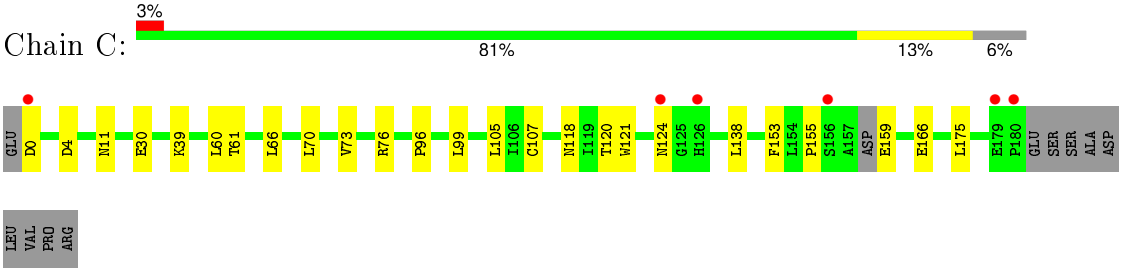
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain

Chain A:



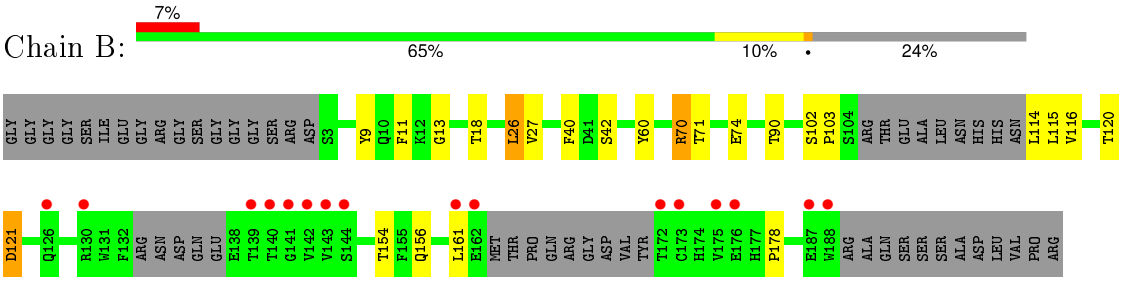
- Molecule 1: HLA class II histocompatibility antigen, DQ alpha 1 chain

Chain C:



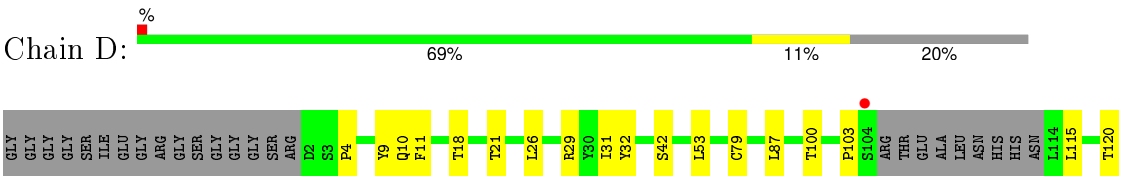
- Molecule 2: HLA class II histocompatibility antigen, DQ beta 1 chain

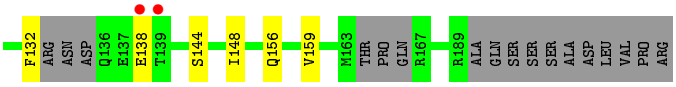
Chain B:



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

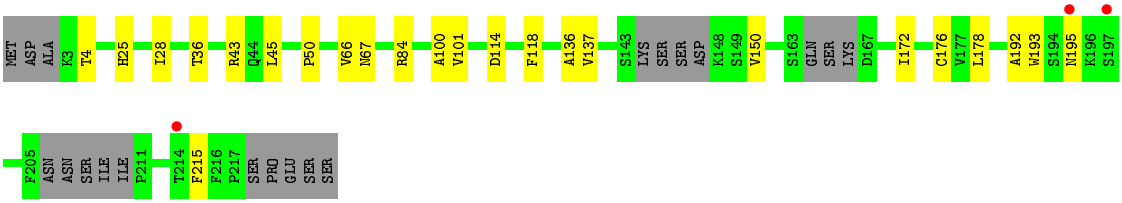
Chain D:





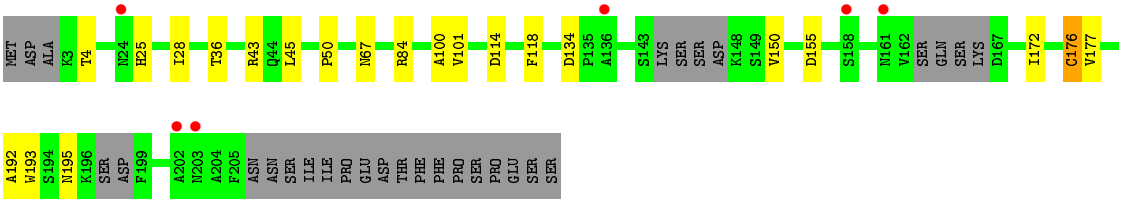
• Molecule 3: T-CELL RECEPTOR, SP3.4 ALPHA CHAIN

Chain E:



• Molecule 3: T-CELL RECEPTOR, SP3.4 ALPHA CHAIN

Chain G:

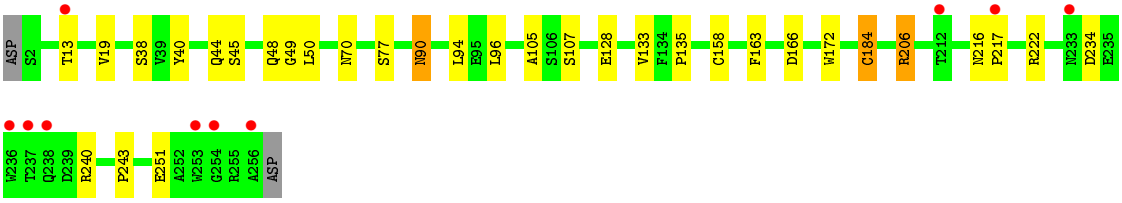


• Molecule 4: T-CELL RECEPTOR, SP3.4 BETA CHAIN

Chain F:

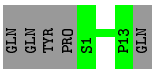
• Molecule 4: T-CELL RECEPTOR, SP3.4 BETA CHAIN

Chain H:



• Molecule 5: Peptide from Alpha/beta-gliadin MM1

Chain I:



• Molecule 5: Peptide from Alpha/beta-gliadin MM1

Chain J:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	111.71Å 134.32Å 140.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.88 – 3.20 35.88 – 3.19	Depositor EDS
% Data completeness (in resolution range)	91.0 (35.88-3.20) 90.8 (35.88-3.19)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.72 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.5.0088	Depositor
R, R_{free}	0.246 , 0.285 0.255 , 0.295	Depositor DCC
R_{free} test set	1646 reflections (5.37%)	DCC
Wilson B-factor (Å^2)	48.2	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.6	EDS
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	3 of 32385 reflections (0.009%)	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	11919	wwPDB-VP
Average B, all atoms (Å^2)	43.0	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 19.82 % 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 9.9707e-03.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.36	0/1417	0.50	0/1943
1	C	0.36	0/1441	0.50	0/1976
2	B	0.37	0/1284	0.52	0/1761
2	D	0.36	0/1361	0.53	0/1867
3	E	0.36	0/1422	0.52	0/1942
3	G	0.33	0/1324	0.50	0/1808
4	F	0.35	0/1878	0.50	0/2576
4	H	0.35	0/1867	0.49	0/2562
5	I	0.38	0/97	0.44	0/130
5	J	0.48	0/97	0.46	0/130
All	All	0.36	0/12188	0.51	0/16695

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	1376	0	1246	10	0
1	C	1398	0	1265	17	0
2	B	1253	0	1135	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1328	0	1208	15	0
3	E	1391	0	1249	10	0
3	G	1298	0	1149	10	0
4	F	1827	0	1658	15	0
4	H	1816	0	1643	16	0
5	I	95	0	80	0	0
5	J	95	0	80	2	0
6	A	28	0	26	0	0
6	C	14	0	13	0	0
All	All	11919	0	10752	89	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 4.

All (89) 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)
3:E:45:LEU:HD23	3:E:100:ALA:HB2	1.69	0.74
3:G:45:LEU:HD23	3:G:100:ALA:HB2	1.70	0.72
2:B:114:LEU:HA	2:B:161:LEU:O	1.94	0.67
2:D:26:LEU:HG	2:D:42:SER:HB3	1.77	0.66
3:G:118:PHE:HB2	4:H:50:LEU:HB3	1.79	0.65
2:B:103:PRO:HA	2:B:115:LEU:HA	1.80	0.64
1:A:118:ASN:HB2	1:A:166:GLU:HB2	1.81	0.62
1:C:70:LEU:HD13	2:D:9:TYR:HB2	1.83	0.60
1:C:118:ASN:HB2	1:C:166:GLU:HB2	1.83	0.60
4:F:133:VAL:O	4:F:240:ARG:NH2	2.35	0.60
1:A:70:LEU:HD13	2:B:9:TYR:HB2	1.84	0.60
3:E:50:PRO:HG3	4:F:50:LEU:HD21	1.85	0.59
1:A:11:ASN:HB2	2:B:11:PHE:HB3	1.84	0.58
4:H:45:SER:OG	4:H:48:GLN:HB2	2.03	0.58
2:D:132:PHE:HA	2:D:138:GLU:HA	1.84	0.58
3:E:118:PHE:HB2	4:F:50:LEU:HB3	1.84	0.57
1:C:39:LYS:HG2	1:C:60:LEU:HD11	1.86	0.57
1:C:11:ASN:HB2	2:D:11:PHE:HB3	1.86	0.56
4:F:40:TYR:HB2	4:F:105:ALA:HB3	1.86	0.56
4:F:45:SER:OG	4:F:48:GLN:HB2	2.07	0.55
3:G:50:PRO:HG3	4:H:50:LEU:HD21	1.88	0.55
2:B:26:LEU:HG	2:B:42:SER:HB3	1.89	0.55
1:C:30:GLU:HB2	1:C:138:LEU:HD21	1.89	0.55
2:D:144:SER:HB2	2:D:159:VAL:HG13	1.89	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:H:96:LEU:HD21	4:H:128:GLU:HG3	1.89	0.55
4:H:40:TYR:HB2	4:H:105:ALA:HB3	1.89	0.54
2:B:102:SER:O	2:B:116:VAL:HG12	2.07	0.54
3:E:136:ALA:HB2	3:E:215:PHE:HB3	1.90	0.53
4:F:184:CYS:O	4:F:206:ARG:HD2	2.07	0.53
1:C:124:ASN:HD21	1:C:159:GLU:HA	1.73	0.52
4:F:77:SER:HB2	4:F:90:ASN:HD22	1.74	0.52
4:H:184:CYS:O	4:H:206:ARG:HD2	2.10	0.52
4:H:133:VAL:O	4:H:240:ARG:NH2	2.41	0.51
1:A:39:LYS:HG2	1:A:60:LEU:HD11	1.92	0.50
4:H:77:SER:HB2	4:H:90:ASN:HD22	1.74	0.50
1:A:30:GLU:HB2	1:A:138:LEU:HD21	1.93	0.50
3:E:150:VAL:HG23	3:E:193:TRP:HB3	1.94	0.50
1:C:11:ASN:HB3	1:C:66:LEU:HD11	1.94	0.50
1:C:99:LEU:HA	1:C:155:PRO:HB2	1.94	0.49
4:H:222:ARG:HG3	4:H:251:GLU:HB3	1.95	0.49
1:A:11:ASN:HB3	1:A:66:LEU:HD11	1.94	0.49
2:B:27:VAL:HA	2:B:40:PHE:O	2.12	0.49
1:A:99:LEU:HA	1:A:155:PRO:HB2	1.95	0.49
1:C:96:PRO:HD3	2:D:120:THR:HG21	1.94	0.49
4:F:19:VAL:HG22	4:F:94:LEU:HD11	1.95	0.48
4:H:19:VAL:HG22	4:H:94:LEU:HD11	1.95	0.48
1:A:4:ASP:HB2	2:B:18:THR:HA	1.95	0.48
2:B:13:GLY:HA3	5:J:6:PHE:CD1	2.49	0.48
1:C:4:ASP:HB2	2:D:18:THR:HA	1.96	0.47
4:F:222:ARG:HG3	4:F:251:GLU:HB3	1.97	0.47
2:D:148:ILE:HB	2:D:156:GLN:HB3	1.96	0.47
3:G:172:ILE:HG23	3:G:192:ALA:HB2	1.97	0.46
2:B:121:ASP:HA	2:B:154:THR:HB	1.98	0.46
3:G:150:VAL:HG23	3:G:193:TRP:HB3	1.97	0.46
3:E:28:ILE:HG21	3:E:84:ARG:O	2.15	0.46
1:C:0:ASP:HA	2:D:29:ARG:HH22	1.80	0.46
4:F:158:CYS:HB2	4:F:172:TRP:CZ2	2.50	0.46
3:E:172:ILE:HG23	3:E:192:ALA:HB2	1.98	0.45
3:G:4:THR:HG22	3:G:25:HIS:HB3	1.98	0.45
2:B:60:TYR:CD2	5:J:12:ASN:HB2	2.51	0.45
2:D:10:GLN:HB2	2:D:31:ILE:HB	1.97	0.45
1:C:105:LEU:HG	1:C:153:PHE:CE2	2.51	0.45
4:H:158:CYS:HB2	4:H:172:TRP:CZ2	2.52	0.45
1:C:61:THR:HG23	4:F:110:VAL:HG21	1.99	0.45
1:C:73:VAL:HG11	2:D:32:TYR:HB2	1.99	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:105:LEU:HG	1:A:153:PHE:CE2	2.52	0.44
2:B:70:ARG:O	2:B:74:GLU:HG3	2.17	0.44
4:H:38:SER:HB2	4:H:107:SER:HB3	2.00	0.43
2:D:103:PRO:HA	2:D:115:LEU:HA	1.99	0.43
1:C:0:ASP:CA	2:D:29:ARG:HH22	2.32	0.43
3:E:4:THR:HG22	3:E:25:HIS:HB3	2.01	0.43
1:C:76:ARG:HD2	2:D:53:LEU:HD12	2.00	0.42
4:F:135:PRO:HD3	4:F:243:PRO:HB3	2.02	0.42
4:H:44:GLN:HA	4:H:49:GLY:O	2.19	0.42
4:F:163:PHE:HE2	4:F:166:ASP:HA	1.85	0.42
4:H:135:PRO:HD3	4:H:243:PRO:HB3	2.01	0.42
4:H:216:ASN:HA	4:H:217:PRO:HD3	1.90	0.42
2:B:120:THR:HG22	2:B:156:GLN:HG2	2.02	0.41
4:F:164:TYR:HA	4:F:165:PRO:HA	1.85	0.41
1:C:107:CYS:HB2	1:C:121:TRP:CZ2	2.56	0.41
4:H:163:PHE:HE2	4:H:166:ASP:HA	1.85	0.41
3:G:176:CYS:SG	3:G:177:VAL:N	2.91	0.41
3:G:28:ILE:HG21	3:G:84:ARG:O	2.20	0.41
3:G:43:ARG:HA	3:G:101:VAL:O	2.21	0.41
3:E:43:ARG:HA	3:E:101:VAL:O	2.22	0.40
4:F:38:SER:HB2	4:F:107:SER:HB3	2.03	0.40
3:E:178:LEU:H	3:E:178:LEU:HD12	1.86	0.40
1:A:51:PHE:HZ	2:D:100:THR:HG21	1.87	0.40
3:G:134:ASP:O	3:G:155:ASP:HB2	2.22	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	176/192 (92%)	169 (96%)	7 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	177/192 (92%)	170 (96%)	7 (4%)	0	100	100
2	B	155/215 (72%)	139 (90%)	14 (9%)	2 (1%)	15	59
2	D	165/215 (77%)	155 (94%)	8 (5%)	2 (1%)	16	60
3	E	179/207 (86%)	163 (91%)	15 (8%)	1 (1%)	30	75
3	G	169/207 (82%)	158 (94%)	10 (6%)	1 (1%)	30	75
4	F	241/245 (98%)	230 (95%)	11 (5%)	0	100	100
4	H	241/245 (98%)	229 (95%)	12 (5%)	0	100	100
5	I	11/18 (61%)	11 (100%)	0	0	100	100
5	J	11/18 (61%)	11 (100%)	0	0	100	100
All	All	1525/1754 (87%)	1435 (94%)	84 (6%)	6 (0%)	39	80

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	114	ASP
3	G	114	ASP
2	B	121	ASP
2	D	79	CYS
2	D	4	PRO
2	B	178	PRO

5.3.2 Protein sidechains (i)

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	143/176 (81%)	141 (99%)	2 (1%)	74	92
1	C	147/176 (84%)	145 (99%)	2 (1%)	74	92
2	B	125/190 (66%)	121 (97%)	4 (3%)	46	81
2	D	134/190 (70%)	132 (98%)	2 (2%)	72	91
3	E	142/184 (77%)	136 (96%)	6 (4%)	36	75
3	G	130/184 (71%)	126 (97%)	4 (3%)	47	82

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	F	181/212 (85%)	174 (96%)	7 (4%)	39 78
4	H	179/212 (84%)	173 (97%)	6 (3%)	44 80
5	I	11/16 (69%)	11 (100%)	0	100 100
5	J	11/16 (69%)	11 (100%)	0	100 100
All	All	1203/1556 (77%)	1170 (97%)	33 (3%)	52 85

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	120	THR
1	A	175	LEU
2	B	26	LEU
2	B	70	ARG
2	B	71	THR
2	B	90	THR
1	C	120	THR
1	C	175	LEU
2	D	21	THR
2	D	87	LEU
3	E	36	THR
3	E	66	VAL
3	E	67	ASN
3	E	137	VAL
3	E	176	CYS
3	E	195	ASN
4	F	13	THR
4	F	70	ASN
4	F	71	ILE
4	F	90	ASN
4	F	184	CYS
4	F	206	ARG
4	F	234	ASP
3	G	36	THR
3	G	67	ASN
3	G	176	CYS
3	G	195	ASN
4	H	13	THR
4	H	70	ASN
4	H	90	ASN
4	H	184	CYS
4	H	206	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	H	234	ASP

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

Mol	Chain	Res	Type
2	B	92	GLN
1	C	124	ASN
2	D	156	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

3 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	201	1	14,14,15	0.53	0	15,19,21	1.12	1 (6%)
6	NAG	A	202	1	14,14,15	0.45	0	15,19,21	0.76	0
6	NAG	C	201	1	14,14,15	0.47	0	15,19,21	1.01	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	201	1	-	0/6/23/26	0/1/1/1
6	NAG	A	202	1	-	0/6/23/26	0/1/1/1
6	NAG	C	201	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(^{o})	Ideal(^{o})
6	C	201	NAG	O5-C5-C6	2.27	112.26	107.35
6	A	201	NAG	C1-O5-C5	2.60	115.55	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^{th} 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	180/192 (93%)	0.23	9 (5%) 32 19	20, 39, 72, 82	0
1	C	181/192 (94%)	0.18	6 (3%) 50 35	24, 36, 62, 71	0
2	B	163/215 (75%)	0.40	16 (9%) 10 5	19, 48, 78, 85	0
2	D	173/215 (80%)	0.15	3 (1%) 73 60	25, 38, 66, 74	0
3	E	187/207 (90%)	0.12	3 (1%) 74 62	23, 35, 71, 83	0
3	G	177/207 (85%)	0.34	6 (3%) 49 34	27, 44, 83, 90	0
4	F	243/245 (99%)	0.09	1 (0%) 93 90	25, 36, 56, 75	0
4	H	243/245 (99%)	0.41	10 (4%) 41 27	27, 49, 81, 96	0
5	I	13/18 (72%)	-0.18	0 100 100	27, 29, 38, 41	0
5	J	13/18 (72%)	-0.11	0 100 100	22, 27, 43, 49	0
All	All	1573/1754 (89%)	0.23	54 (3%) 49 34	19, 40, 75, 96	0

All (54) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	139	THR	5.5
2	B	172	THR	4.6
1	C	180	PRO	4.6
4	H	236	TRP	4.0
1	C	156	SER	3.7
1	A	0	ASP	3.6
4	H	237	THR	3.6
4	F	256	ALA	3.4
2	B	161	LEU	3.2
1	C	179	GLU	3.1
1	A	124	ASN	3.1
1	C	0	ASP	3.0
2	B	140	THR	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	H	253	TRP	2.9
1	A	126	HIS	2.9
2	B	144	SER	2.7
2	D	104	SER	2.7
2	B	162	GLU	2.6
2	B	130	ARG	2.6
2	B	187	GLU	2.6
3	G	202	ALA	2.5
1	A	156	SER	2.5
2	B	141	GLY	2.5
3	E	197	SER	2.5
4	H	212	THR	2.5
3	E	214	THR	2.5
2	B	142	VAL	2.4
3	G	161	ASN	2.4
4	H	238	GLN	2.4
2	B	175	VAL	2.4
4	H	217	PRO	2.3
1	A	164	LYS	2.2
2	B	143	VAL	2.2
2	B	176	GLU	2.2
2	B	173	CYS	2.2
1	A	123	SER	2.2
2	D	139	THR	2.2
1	A	160	ILE	2.2
1	C	124	ASN	2.2
4	H	233	ASN	2.2
2	B	188	TRP	2.2
3	G	24	ASN	2.2
2	D	138	GLU	2.1
2	B	126	GLN	2.1
4	H	13	THR	2.1
1	A	162	ASP	2.1
3	G	158	SER	2.1
3	G	136	ALA	2.1
4	H	254	GLY	2.1
3	E	195	ASN	2.0
3	G	203	ASN	2.0
1	A	179	GLU	2.0
4	H	256	ALA	2.0
1	C	126	HIS	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^{th} 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	202	14/15	0.81	0.39	3.57	49,57,65,66	0
6	NAG	C	201	14/15	0.84	0.30	0.74	47,52,56,57	0
6	NAG	A	201	14/15	0.83	0.22	-	40,47,53,55	0

6.5 Other polymers (i)

There are no such residues in this entry.