Full wwPDB X-ray Structure Validation Report ⓘ

May 12, 2020 – 11:39 pm BST

PDB ID	:	6EJG
Title	:	CRYSTAL STRUCTURE OF HUMAN CD81 LARGE EXTRACELLULAR LOOP IN COMPLEX WITH SINGLE CHAIN FV FRAGMENT 4
Authors	:	Harris, S.F.; Wong, A.; Kuglstatter, A.
Deposited on	:	2017-09-21
Resolution	:	2.82 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
https://www.wwpdb.org/validation/2017/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:

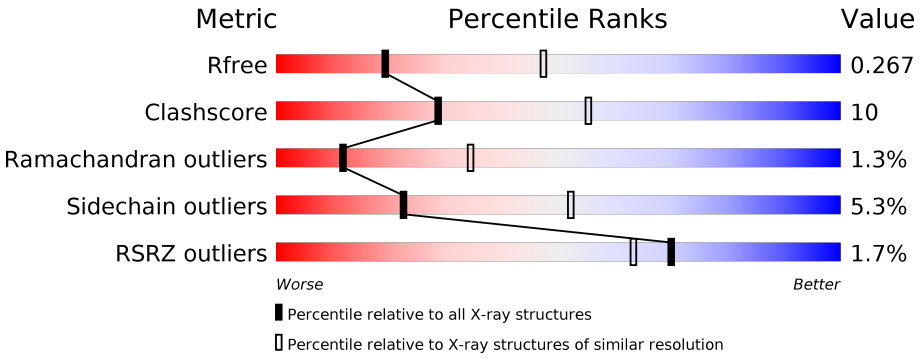
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.82 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}	130704	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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 respectively. 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	99	4% 67% 21% •• 10%
1	B	99	5% 66% 22% • 10%
2	C	251	76% 15% • 8%
2	D	251	75% 16% • 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4971 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 CD81 antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	89	Total	C	N	O	S	0	0	0
			683	424	115	140	4			
1	B	89	Total	C	N	O	S	0	0	0
			683	424	115	140	4			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	GLY	-	expression tag	UNP P60033
A	111	SER	-	expression tag	UNP P60033
A	203	HIS	-	expression tag	UNP P60033
A	204	HIS	-	expression tag	UNP P60033
A	205	HIS	-	expression tag	UNP P60033
A	206	HIS	-	expression tag	UNP P60033
A	207	HIS	-	expression tag	UNP P60033
A	208	HIS	-	expression tag	UNP P60033
B	110	GLY	-	expression tag	UNP P60033
B	111	SER	-	expression tag	UNP P60033
B	203	HIS	-	expression tag	UNP P60033
B	204	HIS	-	expression tag	UNP P60033
B	205	HIS	-	expression tag	UNP P60033
B	206	HIS	-	expression tag	UNP P60033
B	207	HIS	-	expression tag	UNP P60033
B	208	HIS	-	expression tag	UNP P60033

- Molecule 2 is a protein called SINGLE CHAIN FV FRAGMENT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	231	Total	C	N	O	S	0	1	0
			1774	1120	298	348	8			
2	D	231	Total	C	N	O	S	0	1	0
			1774	1120	298	348	8			

- Molecule 3 is water.

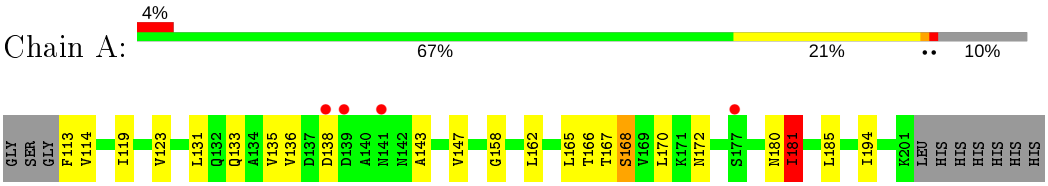
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	5	Total O 5 5	0	0
3	B	2	Total O 2 2	0	0
3	C	23	Total O 23 23	0	0
3	D	27	Total O 27 27	0	0

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 the various outlier classes 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.

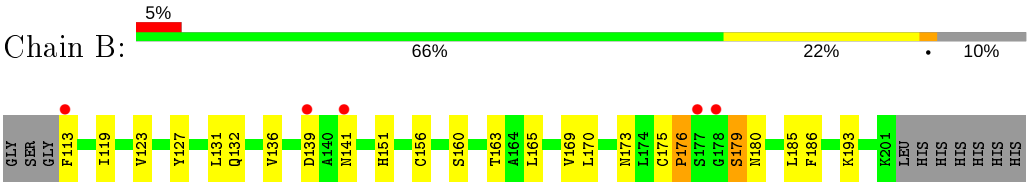
- Molecule 1: CD81 antigen

Chain A:



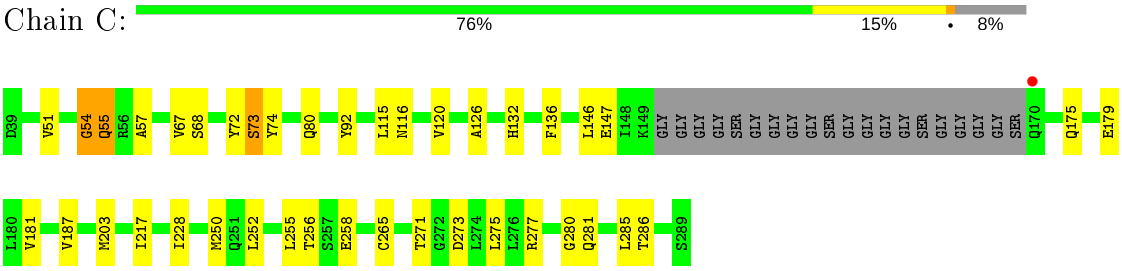
- Molecule 1: CD81 antigen

Chain B:



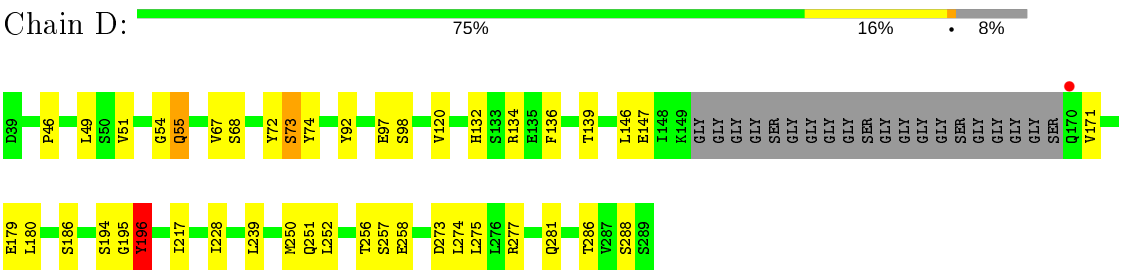
- Molecule 2: SINGLE CHAIN FV FRAGMENT

Chain C:



- Molecule 2: SINGLE CHAIN FV FRAGMENT

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	78.64Å 170.91Å 185.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.83 – 2.82 41.33 – 2.82	Depositor EDS
% Data completeness (in resolution range)	88.2 (46.83-2.82) 88.2 (41.33-2.82)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.17	Depositor
$< I/\sigma(I) >$ [1]	1.77 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.227 , 0.263 0.232 , 0.267	Depositor DCC
R_{free} test set	1348 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	47.8	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 19.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	4971	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.34	0/691	0.42	0/933
1	B	0.33	0/691	0.40	0/933
2	C	0.31	0/1815	0.53	0/2453
2	D	0.32	0/1815	0.55	0/2453
All	All	0.32	0/5012	0.51	0/6772

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	683	0	665	25	0
1	B	683	0	665	16	0
2	C	1774	0	1731	29	0
2	D	1774	0	1731	31	0
3	A	5	0	0	1	0
3	B	2	0	0	1	0
3	C	23	0	0	0	0
3	D	27	0	0	2	0
All	All	4971	0	4792	92	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 10.

All (92) 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:B:170:LEU:HD22	1:B:185:LEU:HD11	1.44	0.96
2:C:187:VAL:HG23	2:C:255:LEU:HD11	1.56	0.86
1:A:135:VAL:HG21	1:A:165:LEU:HD22	1.56	0.85
2:D:217:ILE:HD13	2:D:250:MET:CE	2.11	0.81
2:C:217:ILE:HG21	2:C:250:MET:HE3	1.63	0.81
2:C:256:THR:HG22	2:C:258:GLU:H	1.49	0.77
2:C:54:GLY:HA3	2:C:55:GLN:HB2	1.67	0.77
2:D:54:GLY:HA3	2:D:55:GLN:HB2	1.68	0.76
2:C:54:GLY:CA	2:C:55:GLN:HB2	2.18	0.73
2:D:217:ILE:HG21	2:D:250:MET:HE3	1.72	0.72
2:C:67:VAL:HG21	2:C:132:HIS:HB2	1.72	0.71
1:A:123:VAL:HG12	1:A:194:ILE:HD13	1.73	0.69
2:D:54:GLY:CA	2:D:55:GLN:HB2	2.22	0.69
1:A:133:GLN:HA	1:A:136:VAL:HG12	1.74	0.69
2:C:217:ILE:HD13	2:C:250:MET:CE	2.23	0.68
2:C:187:VAL:HG23	2:C:255:LEU:CD1	2.24	0.67
1:A:136:VAL:HG23	2:C:92:TYR:CZ	2.30	0.67
1:B:132:GLN:O	1:B:136:VAL:HG12	1.94	0.66
1:B:136:VAL:HG23	2:D:92:TYR:CZ	2.31	0.66
2:C:217:ILE:HD13	2:C:250:MET:HE1	1.76	0.65
1:B:170:LEU:HD11	1:B:180:ASN:HA	1.78	0.65
2:D:256:THR:HG22	2:D:258:GLU:H	1.61	0.65
2:D:217:ILE:HD13	2:D:250:MET:HE3	1.78	0.64
2:D:51:VAL:HG21	2:D:120:VAL:HG21	1.79	0.63
2:D:73:SER:HA	2:D:74:TYR:HB2	1.81	0.63
1:A:180:ASN:O	1:A:181:ILE:CD1	2.47	0.62
1:B:136:VAL:HG22	1:B:136:VAL:O	2.02	0.60
2:D:217:ILE:HD13	2:D:250:MET:HE1	1.84	0.60
1:A:114:VAL:CG2	1:A:119:ILE:HD11	2.32	0.59
1:A:158:GLY:HA3	1:A:166:THR:HG23	1.84	0.59
2:C:73:SER:HA	2:C:74:TYR:HB2	1.84	0.58
2:D:67:VAL:HG21	2:D:132:HIS:HB2	1.85	0.58
1:A:170:LEU:HD11	1:A:180:ASN:HA	1.86	0.57
1:A:180:ASN:O	1:A:181:ILE:HD13	2.05	0.57
2:D:67:VAL:O	2:D:74:TYR:HB2	2.06	0.55
1:A:135:VAL:CG2	1:A:165:LEU:HD22	2.32	0.54
2:D:46:PRO:HG3	2:D:49:LEU:HD13	1.89	0.54
1:A:138:ASP:HA	2:C:271:THR:HG21	1.91	0.53
1:A:136:VAL:HG11	2:C:72:TYR:CE1	2.43	0.53
2:D:136:PHE:CE1	2:D:228:ILE:HD13	2.44	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:136:PHE:CE1	2:C:228:ILE:HD13	2.45	0.52
2:C:175:GLN:HE21	2:C:280:GLY:HA3	1.76	0.51
1:B:165:LEU:HD11	2:D:74:TYR:OH	2.11	0.50
2:D:54:GLY:HA3	2:D:55:GLN:CB	2.37	0.50
1:A:136:VAL:HG11	2:C:72:TYR:CZ	2.47	0.50
2:C:54:GLY:N	2:C:55:GLN:HB2	2.27	0.49
1:A:114:VAL:HG22	1:A:119:ILE:HD11	1.94	0.49
1:B:136:VAL:HG11	2:D:72:TYR:CE1	2.48	0.49
2:D:273:ASP:O	2:D:275:LEU:N	2.38	0.48
1:A:166:THR:HG22	1:A:185:LEU:HD13	1.97	0.47
2:D:146:LEU:HD23	2:D:147:GLU:N	2.29	0.47
2:D:67:VAL:HG23	2:D:134:ARG:HB2	1.97	0.46
2:C:250:MET:HE2	2:C:252:LEU:HD21	1.96	0.46
2:C:181:VAL:HG21	2:C:255:LEU:HD13	1.96	0.46
1:A:114:VAL:HG21	1:A:119:ILE:HD11	1.98	0.46
1:B:127:TYR:OH	1:B:151:HIS:NE2	2.42	0.46
2:C:57:ALA:HB2	2:C:120:VAL:HG21	1.99	0.45
2:D:92:TYR:CE2	2:D:275:LEU:HD21	2.51	0.45
1:A:143:ALA:O	1:A:147:VAL:HG23	2.17	0.45
1:B:160:SER:O	1:B:163:THR:HG23	2.17	0.45
1:A:167:THR:HG21	2:C:228:ILE:HD11	1.98	0.45
2:C:67:VAL:O	2:C:74:TYR:HB2	2.17	0.45
2:D:139:THR:HG21	3:D:325:HOH:O	2.18	0.44
1:B:163:THR:HG22	3:B:302:HOH:O	2.17	0.44
1:B:175:CYS:HB3	1:B:176:PRO:HD2	2.00	0.44
2:C:273:ASP:OD1	2:C:275:LEU:HD22	2.18	0.43
2:D:250:MET:HE2	2:D:252:LEU:HD21	2.00	0.43
1:B:179:SER:OG	1:B:180:ASN:N	2.51	0.43
2:C:80:GLN:O	2:C:126:ALA:HB1	2.18	0.43
2:D:171:VAL:HG13	3:D:327:HOH:O	2.19	0.43
1:B:156:CYS:HB3	1:B:193:LYS:HG3	2.01	0.43
1:A:113:PHE:N	3:A:301:HOH:O	2.52	0.43
2:C:73:SER:HA	2:C:74:TYR:CB	2.49	0.42
2:D:195:GLY:O	2:D:196:TYR:HB3	2.19	0.42
2:C:115:LEU:HD12	2:C:116:ASN:N	2.35	0.42
2:C:187:VAL:HG21	2:C:285:LEU:HD21	2.01	0.42
1:A:170:LEU:HD22	1:A:185:LEU:HD21	2.01	0.42
1:B:131:LEU:HD12	1:B:169:VAL:HG11	2.02	0.41
2:D:180:LEU:HD11	2:D:288:SER:HB3	2.02	0.41
1:A:168:SER:O	1:A:172:ASN:ND2	2.53	0.41
1:A:180:ASN:O	1:A:181:ILE:HD12	2.17	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:256:THR:CG2	2:D:257:SER:N	2.82	0.41
2:D:73:SER:HA	2:D:74:TYR:CB	2.49	0.41
2:D:73:SER:O	2:D:92:TYR:HA	2.20	0.41
1:A:172:ASN:HD22	1:A:172:ASN:N	2.18	0.41
2:D:54:GLY:N	2:D:55:GLN:HB2	2.36	0.41
1:A:158:GLY:HA3	1:A:166:THR:CG2	2.49	0.41
1:B:136:VAL:HG11	2:D:72:TYR:CZ	2.55	0.41
1:A:131:LEU:HD23	1:A:162:LEU:HD12	2.02	0.40
1:B:119:ILE:O	1:B:123:VAL:HG23	2.21	0.40
2:C:51:VAL:HG21	2:C:120:VAL:HG21	2.04	0.40
2:C:203:MET:HB2	2:C:203:MET:HE2	1.88	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	87/99 (88%)	83 (95%)	3 (3%)	1 (1%)	14 39
1	B	87/99 (88%)	77 (88%)	8 (9%)	2 (2%)	6 20
2	C	228/251 (91%)	215 (94%)	11 (5%)	2 (1%)	17 44
2	D	228/251 (91%)	213 (93%)	12 (5%)	3 (1%)	12 34
All	All	630/700 (90%)	588 (93%)	34 (5%)	8 (1%)	12 34

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	181	ILE
2	C	55	GLN
2	D	55	GLN
2	D	196	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	176	PRO
1	B	179	SER
2	C	54	GLY
2	D	274	LEU

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	79/87 (91%)	77 (98%)	2 (2%)	47 78
1	B	79/87 (91%)	74 (94%)	5 (6%)	18 44
2	C	194/197 (98%)	184 (95%)	10 (5%)	23 53
2	D	194/197 (98%)	181 (93%)	13 (7%)	16 41
All	All	546/568 (96%)	516 (94%)	30 (6%)	22 50

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	168	SER
1	A	181	ILE
1	B	113	PHE
1	B	139	ASP
1	B	141	ASN
1	B	173	ASN
1	B	186	PHE
2	C	68	SER
2	C	73	SER
2	C	146	LEU
2	C	147	GLU
2	C	179	GLU
2	C	265[A]	CYS
2	C	265[B]	CYS
2	C	277	ARG
2	C	281	GLN
2	C	286	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	68	SER
2	D	73	SER
2	D	97	GLU
2	D	98	SER
2	D	179	GLU
2	D	186	SER
2	D	194	SER
2	D	196	TYR
2	D	239	LEU
2	D	251	GLN
2	D	277	ARG
2	D	281	GLN
2	D	286	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	125	GLN
1	A	141	ASN
1	A	142	ASN
1	A	172	ASN
1	B	129	GLN
1	B	172	ASN
1	B	180	ASN
1	B	184	ASN
2	C	55	GLN
2	C	175	GLN
2	D	84	GLN
2	D	174	GLN
2	D	175	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

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	89/99 (89%)	0.30	4 (4%)	33	23	48, 56, 64, 65	0
1	B	89/99 (89%)	0.33	5 (5%)	24	16	47, 57, 65, 68	0
2	C	231/251 (92%)	-0.11	1 (0%)	92	91	37, 45, 54, 57	0
2	D	231/251 (92%)	-0.20	1 (0%)	92	91	38, 43, 47, 58	0
All	All	640/700 (91%)	-0.03	11 (1%)	70	63	37, 46, 62, 68	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	178	GLY	4.9
1	A	138	ASP	4.3
1	A	141	ASN	3.1
2	D	170	GLN	3.0
1	A	139	ASP	2.8
1	B	141	ASN	2.8
1	B	113	PHE	2.7
1	B	177	SER	2.5
2	C	170	GLN	2.1
1	A	177	SER	2.1
1	B	139	ASP	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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.