wwPDB X-ray Structure Validation Summary Report (i)

Jan 8, 2024 – 02:05 PM EST

PDB ID	:	8G8N
Title	:	CTLA4 Fab with peptide
Authors	:	Williams, J.C.
Deposited on	:	2023-02-18
Resolution	:	3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

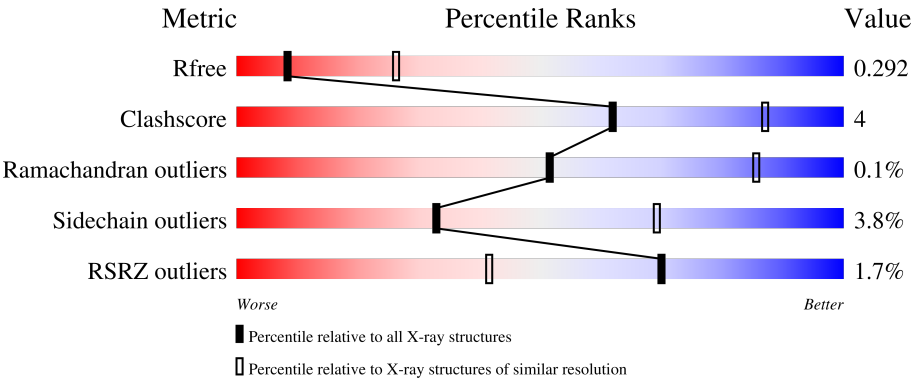
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.00 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.



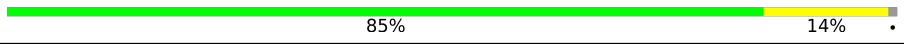
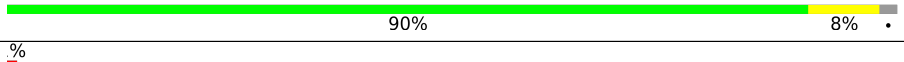
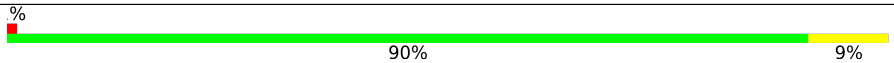
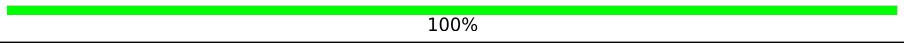
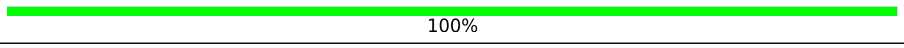
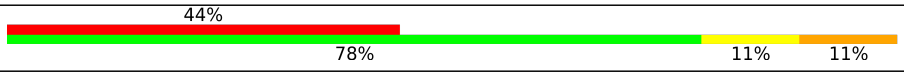
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 of 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	222	91% 5%
1	D	222	89% 7%
1	G	222	90% 6%
1	H	222	89% 6%
1	K	222	2% 74% 22%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	O	222	
2	B	213	
2	E	213	
2	I	213	
2	L	213	
2	M	213	
2	Q	213	
3	C	9	
3	F	9	
3	J	9	
3	P	9	
3	R	9	
3	Z	9	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 39271 atoms, of which 19311 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 Fab heavy chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	213	Total	C	H	N	O	S	0	0	0
			3200	1020	1585	272	315	8			
1	D	217	Total	C	H	N	O	S	0	0	0
			3255	1035	1612	277	322	9			
1	G	216	Total	C	H	N	O	S	0	0	0
			3244	1032	1607	276	320	9			
1	H	214	Total	C	H	N	O	S	0	0	0
			3222	1026	1598	274	316	8			
1	K	214	Total	C	H	N	O	S	0	0	0
			3206	1022	1587	273	316	8			
1	O	217	Total	C	H	N	O	S	0	0	0
			3251	1034	1610	277	321	9			

- Molecule 2 is a protein called Fab light chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	211	Total	C	H	N	O	S	0	0	0
			3198	1027	1562	273	331	5			
2	E	211	Total	C	H	N	O	S	0	0	0
			3193	1026	1555	273	333	6			
2	I	209	Total	C	H	N	O	S	0	0	0
			3157	1015	1536	271	329	6			
2	L	213	Total	C	H	N	O	S	0	0	0
			3219	1033	1571	275	334	6			
2	M	213	Total	C	H	N	O	S	0	0	0
			3225	1035	1573	275	336	6			
2	Q	211	Total	C	H	N	O	S	0	0	0
			3199	1028	1561	272	332	6			

- Molecule 3 is a protein called CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS.

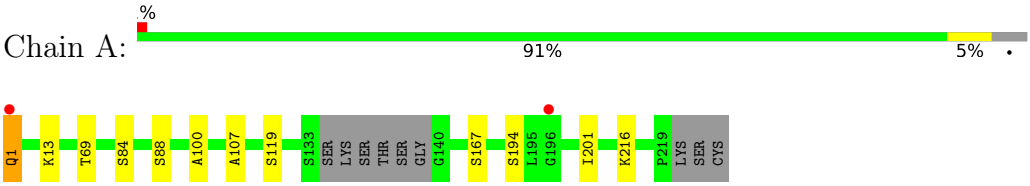
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			
3	F	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			
3	J	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			
3	P	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			
3	R	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			
3	Z	9	Total	C	H	N	O	S	0	0	0
			117	35	59	10	11	2			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

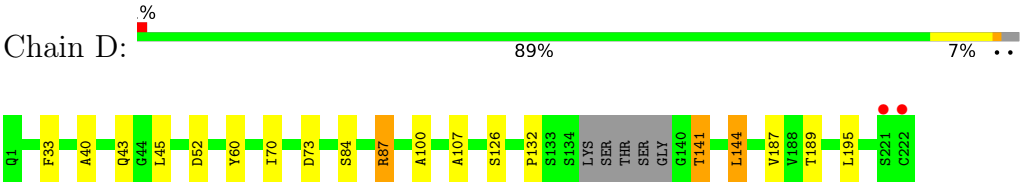
• Molecule 1: Fab heavy chain

Chain A:



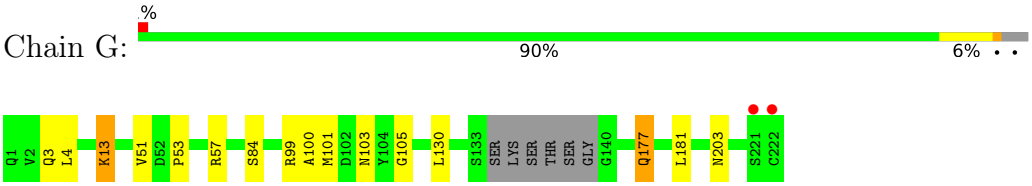
• Molecule 1: Fab heavy chain

Chain D:



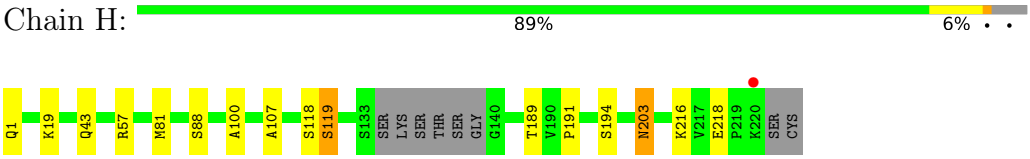
• Molecule 1: Fab heavy chain

Chain G:



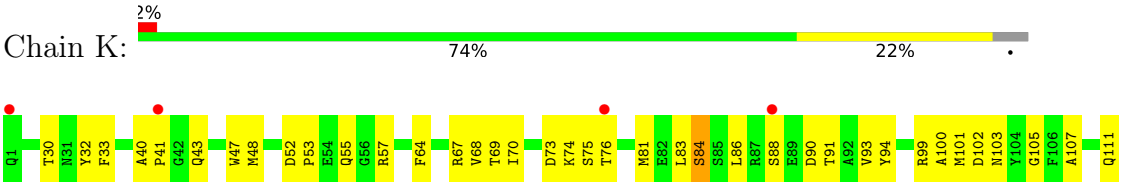
• Molecule 1: Fab heavy chain

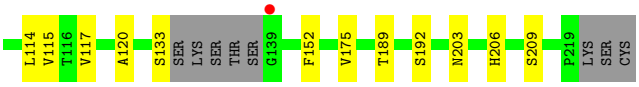
Chain H:



• Molecule 1: Fab heavy chain

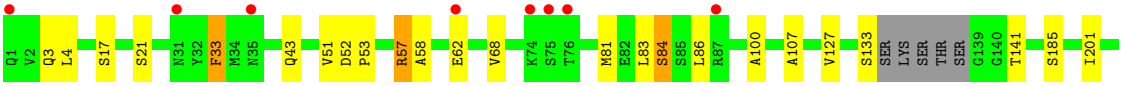
Chain K:

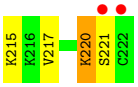




• Molecule 1: Fab heavy chain

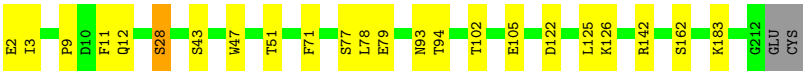
Chain O:





• Molecule 2: Fab light chain

Chain B:

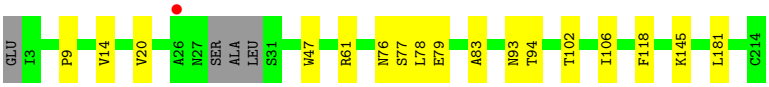


• Molecule 2: Fab light chain

Chain E:

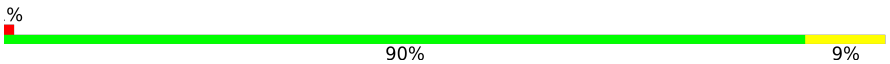
• Molecule 2: Fab light chain

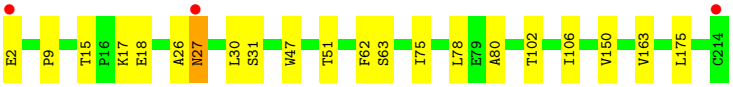
Chain I:



• Molecule 2: Fab light chain

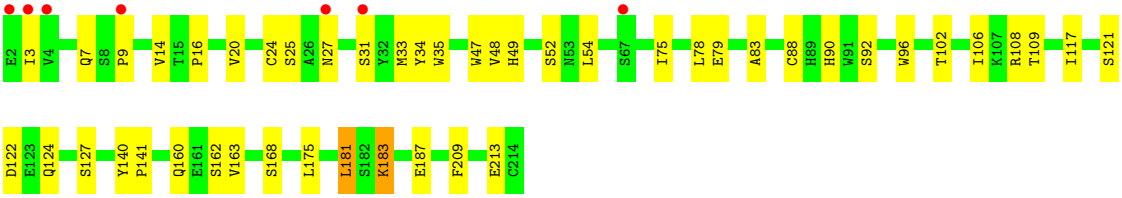
Chain L:



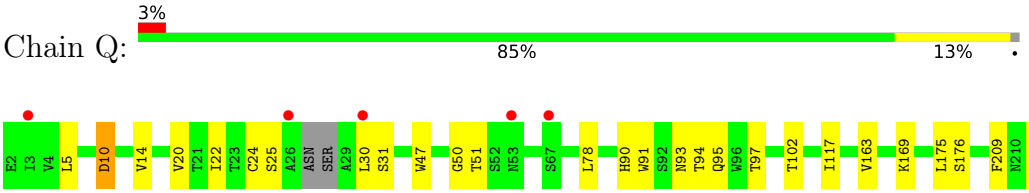


• Molecule 2: Fab light chain

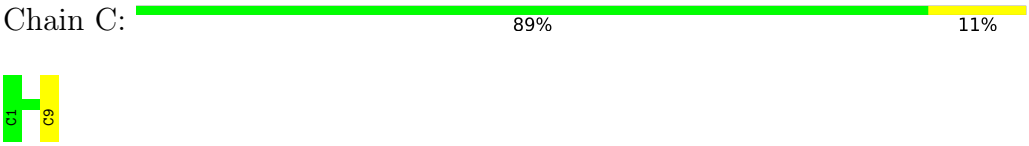
Chain M:



• Molecule 2: Fab light chain



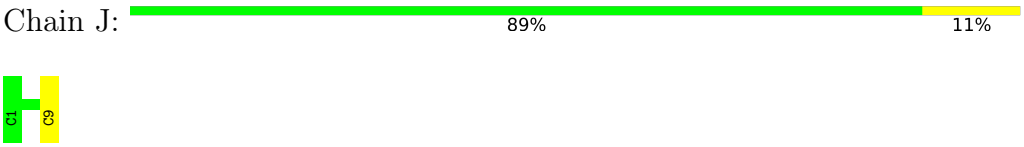
• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS



• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS

There are no outlier residues recorded for this chain.

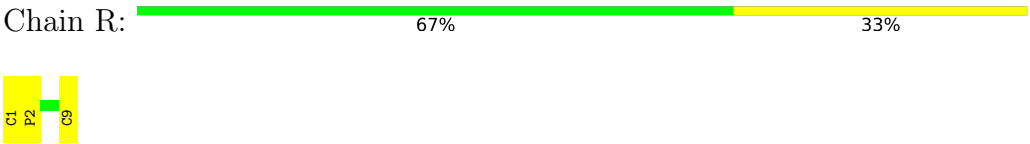
• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS



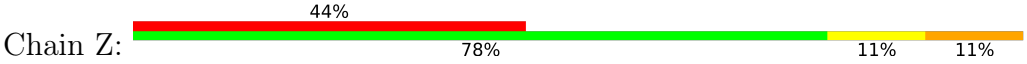
• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS

There are no outlier residues recorded for this chain.

• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS



• Molecule 3: CYS-PRO-GLY-LYS-GLY-LEU-PRO-SER-CYS




C1
K4
G5
L6
P7
S8
C9

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	72.43Å 203.51Å 225.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.03 – 3.00 20.03 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (20.03-3.00) 89.4 (20.03-3.00)	Depositor EDS
R_{merge}	0.27	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.98Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.239 , 0.288 0.242 , 0.292	Depositor DCC
R_{free} test set	3317 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	31.5	Xtriage
Anisotropy	0.586	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 27.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	39271	wwPDB-VP
Average B, all atoms (Å^2)	48.0	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 47.52 % 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 9.8040e-05. 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, 0.333 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	0.26	0/1653	0.50	0/2247
1	D	0.26	0/1681	0.50	0/2282
1	G	0.27	0/1675	0.51	0/2274
1	H	0.26	0/1662	0.50	0/2258
1	K	0.27	0/1657	0.52	0/2252
1	O	0.27	0/1679	0.51	0/2279
2	B	0.26	0/1677	0.48	0/2283
2	E	0.30	0/1678	0.49	0/2281
2	I	0.25	0/1661	0.48	0/2258
2	L	0.25	0/1689	0.48	0/2298
2	M	0.25	0/1693	0.49	0/2303
2	Q	0.25	0/1678	0.48	0/2281
3	C	0.41	0/59	0.43	0/77
3	F	0.34	0/59	0.55	0/77
3	J	0.33	0/59	0.52	0/77
3	P	0.37	0/59	0.48	0/77
3	R	0.44	0/59	0.59	0/77
3	Z	0.34	0/59	0.42	0/77
All	All	0.26	0/20437	0.50	0/27758

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	1615	1585	1585	4	0
1	D	1643	1612	1612	13	0
1	G	1637	1607	1607	7	0
1	H	1624	1598	1598	7	0
1	K	1619	1587	1588	38	0
1	O	1641	1610	1610	12	0
2	B	1636	1562	1562	8	0
2	E	1638	1555	1555	19	0
2	I	1621	1536	1535	15	0
2	L	1648	1571	1569	11	0
2	M	1652	1573	1573	29	0
2	Q	1638	1561	1560	19	0
3	C	58	59	59	0	0
3	F	58	59	59	0	0
3	J	58	59	59	0	0
3	P	58	59	59	0	0
3	R	58	59	59	1	0
3	Z	58	59	59	5	0
All	All	19960	19311	19308	174	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.

The worst 5 of 174 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)
2:E:103:LYS:HE2	2:E:105:GLU:OE2	1.65	0.95
1:H:1:GLN:OE1	1:H:1:GLN:N	2.05	0.88
1:K:99:ARG:NH1	3:Z:6:LEU:HD12	1.90	0.85
2:M:83:ALA:HB2	2:M:106:ILE:HD13	1.60	0.81
2:M:35:TRP:CZ3	2:M:88:CYS:HB3	2.17	0.80

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	209/222 (94%)	208 (100%)	1 (0%)	0	100	100
1	D	213/222 (96%)	208 (98%)	5 (2%)	0	100	100
1	G	212/222 (96%)	207 (98%)	5 (2%)	0	100	100
1	H	210/222 (95%)	208 (99%)	2 (1%)	0	100	100
1	K	210/222 (95%)	204 (97%)	6 (3%)	0	100	100
1	O	213/222 (96%)	205 (96%)	8 (4%)	0	100	100
2	B	209/213 (98%)	201 (96%)	7 (3%)	1 (0%)	29	68
2	E	207/213 (97%)	195 (94%)	11 (5%)	1 (0%)	29	68
2	I	205/213 (96%)	199 (97%)	6 (3%)	0	100	100
2	L	211/213 (99%)	203 (96%)	8 (4%)	0	100	100
2	M	211/213 (99%)	194 (92%)	16 (8%)	1 (0%)	29	68
2	Q	207/213 (97%)	194 (94%)	13 (6%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
3	F	7/9 (78%)	7 (100%)	0	0	100	100
3	J	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
3	P	7/9 (78%)	7 (100%)	0	0	100	100
3	R	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
3	Z	7/9 (78%)	7 (100%)	0	0	100	100
All	All	2559/2664 (96%)	2464 (96%)	92 (4%)	3 (0%)	51	85

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	3	ILE
2	M	3	ILE
2	E	204	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	179/187 (96%)	173 (97%)	6 (3%)	37 72
1	D	183/187 (98%)	177 (97%)	6 (3%)	38 73
1	G	182/187 (97%)	176 (97%)	6 (3%)	38 73
1	H	180/187 (96%)	173 (96%)	7 (4%)	32 69
1	K	179/187 (96%)	170 (95%)	9 (5%)	24 60
1	O	182/187 (97%)	171 (94%)	11 (6%)	19 53
2	B	186/188 (99%)	178 (96%)	8 (4%)	29 66
2	E	186/188 (99%)	183 (98%)	3 (2%)	62 86
2	I	184/188 (98%)	181 (98%)	3 (2%)	62 86
2	L	187/188 (100%)	182 (97%)	5 (3%)	44 77
2	M	188/188 (100%)	177 (94%)	11 (6%)	19 54
2	Q	186/188 (99%)	180 (97%)	6 (3%)	39 74
3	C	7/7 (100%)	6 (86%)	1 (14%)	3 15
3	F	7/7 (100%)	7 (100%)	0	100 100
3	J	7/7 (100%)	6 (86%)	1 (14%)	3 15
3	P	7/7 (100%)	7 (100%)	0	100 100
3	R	7/7 (100%)	5 (71%)	2 (29%)	0 2
3	Z	7/7 (100%)	6 (86%)	1 (14%)	3 15
All	All	2244/2292 (98%)	2158 (96%)	86 (4%)	33 69

5 of 86 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	31	SER
1	O	84	SER
2	M	52	SER
2	M	183	LYS
1	O	217	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	H	43	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 monosaccharides 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	213/222 (95%)	-0.38	2 (0%) 84 63	21, 33, 59, 78	0
1	D	217/222 (97%)	-0.17	2 (0%) 84 63	22, 41, 66, 101	0
1	G	216/222 (97%)	-0.26	2 (0%) 84 63	21, 40, 62, 95	0
1	H	214/222 (96%)	-0.40	1 (0%) 91 75	20, 32, 59, 94	0
1	K	214/222 (96%)	0.08	5 (2%) 60 31	24, 56, 80, 123	0
1	O	217/222 (97%)	0.11	10 (4%) 32 12	28, 55, 81, 95	0
2	B	211/213 (99%)	-0.29	0 100 100	21, 36, 52, 67	0
2	E	211/213 (99%)	-0.15	1 (0%) 91 75	26, 39, 67, 101	0
2	I	209/213 (98%)	-0.24	1 (0%) 91 75	23, 35, 66, 93	0
2	L	213/213 (100%)	-0.24	3 (1%) 75 49	22, 38, 62, 99	0
2	M	213/213 (100%)	0.05	7 (3%) 46 20	25, 54, 85, 109	0
2	Q	211/213 (99%)	0.16	6 (2%) 53 25	31, 53, 85, 115	0
3	C	9/9 (100%)	-0.56	0 100 100	24, 29, 36, 43	0
3	F	9/9 (100%)	-0.03	0 100 100	44, 47, 60, 80	0
3	J	9/9 (100%)	-0.32	0 100 100	43, 48, 56, 71	0
3	P	9/9 (100%)	-0.72	0 100 100	21, 26, 30, 35	0
3	R	9/9 (100%)	0.57	0 100 100	66, 70, 77, 94	0
3	Z	9/9 (100%)	1.53	4 (44%) 0 0	67, 74, 87, 92	0
All	All	2613/2664 (98%)	-0.14	44 (1%) 70 41	20, 40, 76, 123	0

The worst 5 of 44 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	214	CYS	4.4
1	O	221	SER	4.1
1	D	222	CYS	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	O	74	LYS	3.8
3	Z	9	CYS	3.6

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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.