wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2021 – 05:45 AM EDT

PDB ID	:	3C5Z
Title	:	Crystal structure of mouse MHC class II I-Ab/3K peptide complexed with mouse TCR B3K506
Authors	:	Dai, S.; Kappler, J.
Deposited on	:	2008-02-01
Resolution	:	2.55 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 following versions of software and data (see references (1)) were used in the production of this report:

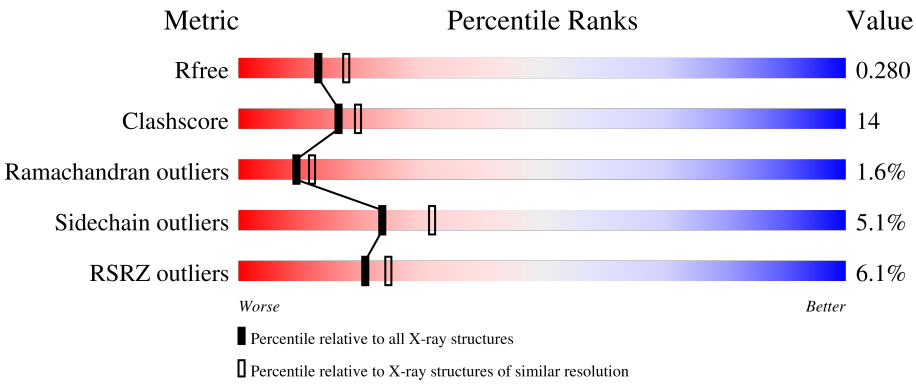
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.55 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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Ramachandran outliers	138981	1315 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

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	202	11% 74% 24% •
1	E	202	6% 68% 29% •
2	B	240	4% 67% 31% •
2	F	240	8% 68% 30% •
3	C	182	4% 66% 31% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	182	5% 72% 26%
4	D	217	6% 61% 29% 7%
4	H	217	2% 65% 26% 7%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13288 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 TCR B3K506 Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	202	1581	990	261	323	7	0	0	0
1	E	201	1573	985	260	321	7	0	0	0

- Molecule 2 is a protein called TCR B3K506 Beta Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	240	1902	1196	330	370	6	0	0	0
2	F	240	1902	1196	330	370	6	0	0	0

- Molecule 3 is a protein called H-2 class II histocompatibility antigen, A-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	182	1459	944	230	282	3	0	0	0
3	G	182	1459	944	230	282	3	0	0	0

- Molecule 4 is a protein called 3K peptide, Linker, and H-2 class II histocompatibility antigen (A beta chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	202	1673	1047	304	315	7	0	0	0
4	H	202	1673	1047	304	315	7	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	14	GLY	-	linker	UNP P14483
D	15	GLY	-	linker	UNP P14483
D	16	GLY	-	linker	UNP P14483
D	17	GLY	-	linker	UNP P14483
D	18	SER	-	linker	UNP P14483
D	19	LEU	-	linker	UNP P14483
D	20	VAL	-	linker	UNP P14483
D	21	PRO	-	linker	UNP P14483
D	22	ARG	-	linker	UNP P14483
D	23	GLY	-	linker	UNP P14483
D	24	SER	-	linker	UNP P14483
D	25	GLY	-	linker	UNP P14483
D	26	GLY	-	linker	UNP P14483
D	27	GLY	-	linker	UNP P14483
D	28	GLY	-	linker	UNP P14483
D	216	LYS	ARG	engineered mutation	UNP P14483
H	14	GLY	-	linker	UNP P14483
H	15	GLY	-	linker	UNP P14483
H	16	GLY	-	linker	UNP P14483
H	17	GLY	-	linker	UNP P14483
H	18	SER	-	linker	UNP P14483
H	19	LEU	-	linker	UNP P14483
H	20	VAL	-	linker	UNP P14483
H	21	PRO	-	linker	UNP P14483
H	22	ARG	-	linker	UNP P14483
H	23	GLY	-	linker	UNP P14483
H	24	SER	-	linker	UNP P14483
H	25	GLY	-	linker	UNP P14483
H	26	GLY	-	linker	UNP P14483
H	27	GLY	-	linker	UNP P14483
H	28	GLY	-	linker	UNP P14483
H	216	LYS	ARG	engineered mutation	UNP P14483

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	6	6	6	0	0
			Total	O		
5	B	6	6	6	0	0
			Total	O		
5	C	10	10	10	0	0
			Total	O		
5	D	10	10	10	0	0

Continued on next page...

Continued from previous page...

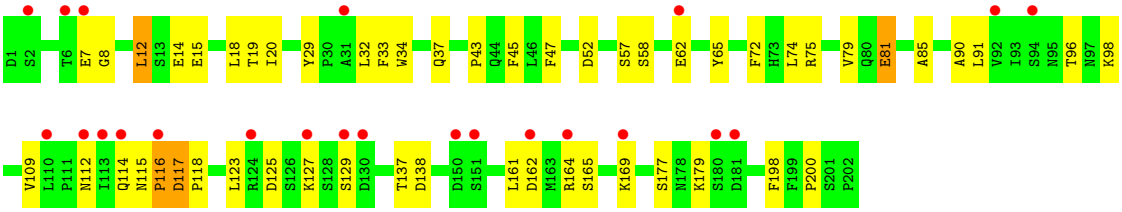
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	E	7	7	7	0	0
			Total	O		
5	F	4	4	4	0	0
			Total	O		
5	G	8	8	8	0	0
			Total	O		
5	H	15	15	15	0	0

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.

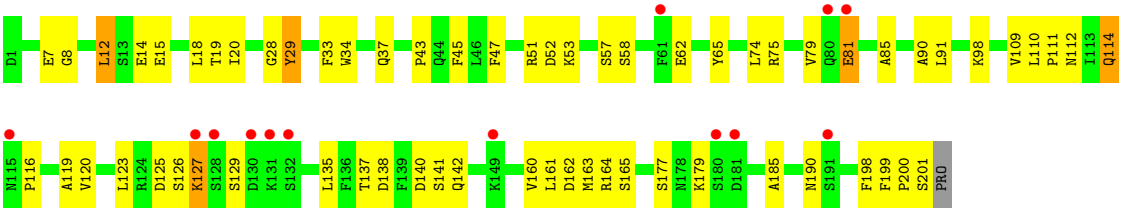
- Molecule 1: TCR B3K506 Alpha Chain

Chain A:



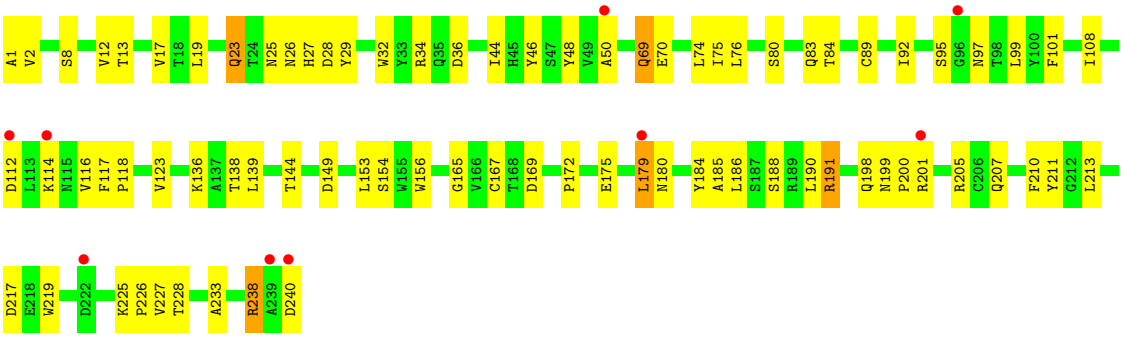
- Molecule 1: TCR B3K506 Alpha Chain

Chain E:



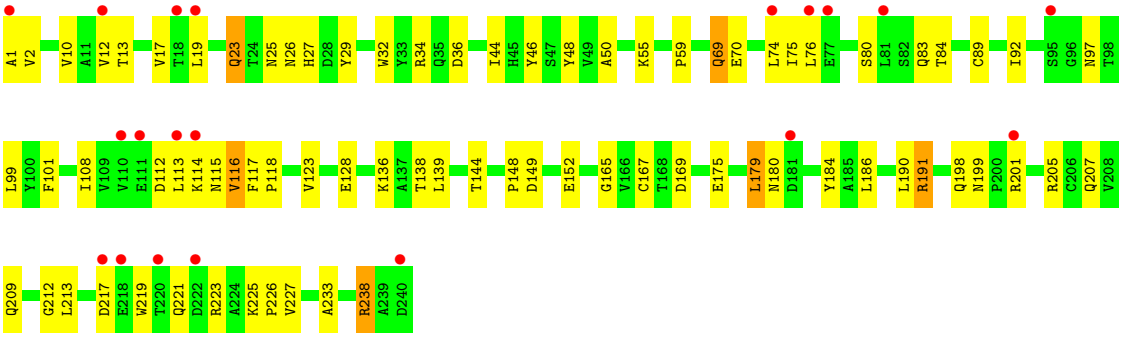
- Molecule 2: TCR B3K506 Beta Chain

Chain B:



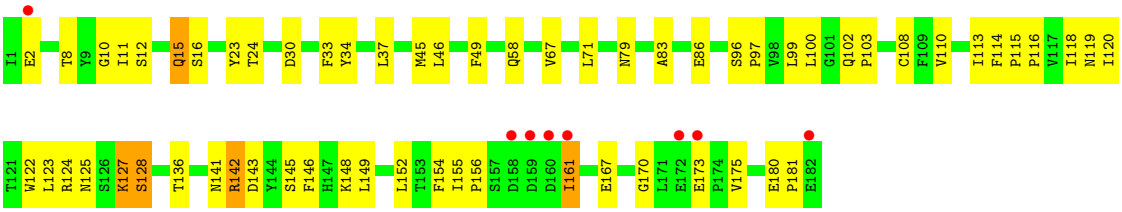
- Molecule 2: TCR B3K506 Beta Chain

Chain F:



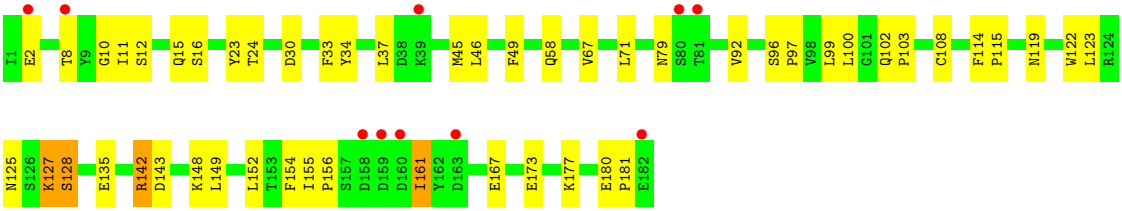
- Molecule 3: H-2 class II histocompatibility antigen, A-B alpha chain

Chain C:



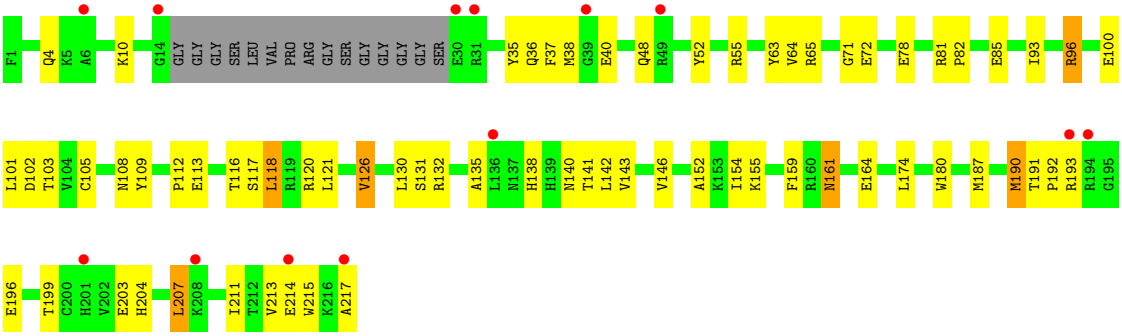
- Molecule 3: H-2 class II histocompatibility antigen, A-B alpha chain

Chain G:



- Molecule 4: 3K peptide, Linker, and H-2 class II histocompatibility antigen (A beta chain)

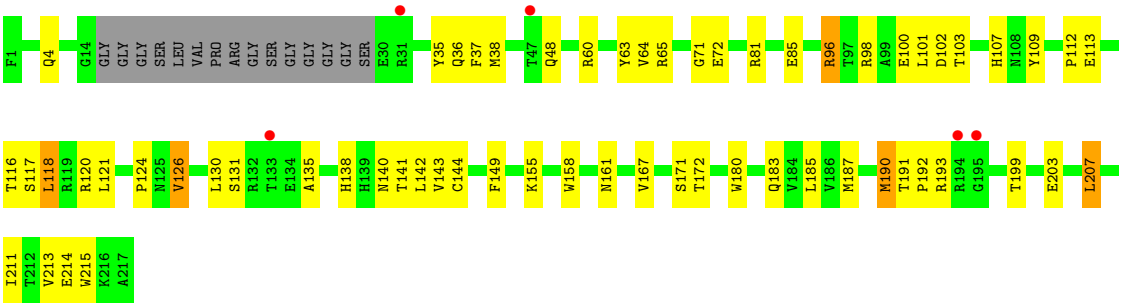
Chain D:



- Molecule 4: 3K peptide, Linker, and H-2 class II histocompatibility antigen (A beta chain)

Chain H:

2%
65%
26%
7%


F1 Q4 G14 GLY GLY GLY SER LEU VAL PRO ARG GLY SER GLY GLY GLY GLY SER E30 R31 Y35 Q36 F37 M38 T47 Q48 R60 Y63 V64 R65 G71 E72 R81 E85 R96 T97 R98 A99 E100 L101 D102 T103 H107 N108 Y109 P112 E113
T116 S117 L118 R119 R120 L121 P124 N125 V126 L130 S131 R132 T133 E134 A135 H138 H139 N140 T141 L142 V143 C144 F149 K155 W158 N161 V167 S171 T172 W180 Q183 V184 L185 V186 M187 M190 T191 P192 R193 R194 G195 T199 E203 L207
I211 T212 V213 E214 W215 K216 A217

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.30Å 117.80Å 136.20Å 90.00° 99.56° 90.00°	Depositor
Resolution (Å)	40.00 – 2.55 33.90 – 2.54	Depositor EDS
% Data completeness (in resolution range)	86.0 (40.00-2.55) 85.2 (33.90-2.54)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 2.54Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.246 , 0.281 0.246 , 0.280	Depositor DCC
R_{free} test set	2886 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	49.2	Xtriage
Anisotropy	0.821	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	13288	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.49% 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 (i)

5.1 Standard geometry (i)

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	0.45	0/1618	0.63	0/2195
1	E	0.45	0/1609	0.63	0/2183
2	B	0.46	0/1953	0.61	0/2660
2	F	0.44	0/1953	0.61	0/2660
3	C	0.51	0/1504	0.65	0/2054
3	G	0.53	0/1504	0.66	0/2054
4	D	0.52	0/1713	0.66	0/2321
4	H	0.56	0/1713	0.66	0/2321
All	All	0.49	0/13567	0.64	0/18448

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 (i)

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	1581	0	1483	33	0
1	E	1573	0	1476	38	0
2	B	1902	0	1802	59	0
2	F	1902	0	1802	63	0
3	C	1459	0	1386	48	0
3	G	1459	0	1386	38	0
4	D	1673	0	1603	60	0
4	H	1673	0	1603	54	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	6	0	0	0	0
5	B	6	0	0	0	0
5	C	10	0	0	1	0
5	D	10	0	0	2	0
5	E	7	0	0	0	0
5	F	4	0	0	0	0
5	G	8	0	0	0	0
5	H	15	0	0	3	0
All	All	13288	0	12541	367	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 14.

The worst 5 of 367 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:D:112:PRO:O	4:D:116:THR:HG22	1.64	0.98
4:H:112:PRO:O	4:H:116:THR:HG22	1.65	0.94
4:D:199:THR:HG22	4:D:214:GLU:HG3	1.53	0.90
4:H:116:THR:HG23	4:H:117:SER:H	1.37	0.90
4:H:199:THR:HG22	4:H:214:GLU:HG3	1.53	0.89

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	200/202 (99%)	174 (87%)	19 (10%)	7 (4%)	3 2
1	E	199/202 (98%)	176 (88%)	16 (8%)	7 (4%)	3 2
2	B	238/240 (99%)	212 (89%)	23 (10%)	3 (1%)	12 16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	238/240 (99%)	214 (90%)	21 (9%)	3 (1%)	12	16
3	C	180/182 (99%)	166 (92%)	12 (7%)	2 (1%)	14	19
3	G	180/182 (99%)	167 (93%)	11 (6%)	2 (1%)	14	19
4	D	198/217 (91%)	184 (93%)	13 (7%)	1 (0%)	29	40
4	H	198/217 (91%)	183 (92%)	14 (7%)	1 (0%)	29	40
All	All	1631/1682 (97%)	1476 (90%)	129 (8%)	26 (2%)	9	12

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	117	ASP
1	A	165	SER
4	D	161	ASN
1	E	165	SER
1	A	7	GLU

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	179/179 (100%)	172 (96%)	7 (4%)	32	44
1	E	178/179 (99%)	169 (95%)	9 (5%)	24	32
2	B	208/208 (100%)	198 (95%)	10 (5%)	25	34
2	F	208/208 (100%)	198 (95%)	10 (5%)	25	34
3	C	163/163 (100%)	152 (93%)	11 (7%)	16	21
3	G	163/163 (100%)	152 (93%)	11 (7%)	16	21
4	D	182/189 (96%)	175 (96%)	7 (4%)	33	45
4	H	182/189 (96%)	173 (95%)	9 (5%)	25	34
All	All	1463/1478 (99%)	1389 (95%)	74 (5%)	24	32

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	79	ASN
4	H	190	MET
3	G	142	ARG
4	H	4	GLN
3	C	152	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	83	GLN
3	G	102	GLN
2	F	115	ASN
3	G	15	GLN
3	G	168	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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 (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	202/202 (100%)	0.82	23 (11%) 5 6	38, 70, 100, 123	0
1	E	201/202 (99%)	0.67	13 (6%) 18 22	42, 69, 100, 124	0
2	B	240/240 (100%)	0.42	9 (3%) 40 47	40, 63, 95, 124	0
2	F	240/240 (100%)	0.64	20 (8%) 11 13	39, 64, 94, 110	0
3	C	182/182 (100%)	0.41	8 (4%) 34 41	30, 54, 91, 112	0
3	G	182/182 (100%)	0.40	10 (5%) 25 30	29, 54, 92, 111	0
4	D	202/217 (93%)	0.48	13 (6%) 19 22	35, 53, 86, 114	0
4	H	202/217 (93%)	0.41	5 (2%) 57 63	33, 52, 86, 113	0
All	All	1651/1682 (98%)	0.53	101 (6%) 21 25	29, 61, 95, 124	0

The worst 5 of 101 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	194	ARG	6.2
4	D	14	GLY	6.0
1	A	127	LYS	5.8
2	F	240	ASP	5.3
1	E	81	GLU	5.2

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.