wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:58 AM GMT

PDB ID	:	2JCC
Title	:	AH3 RECOGNITION OF MUTANT HLA-A2 W167A
Authors	:	Miller, P.; Benhar, Y.P.; Biddison, W.; Collins, E.J.
Deposited on	:	2006-12-21
Resolution	:	2.50 Å(reported)

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 (i)

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

The reported resolution of this entry is 2.50 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	Percentile Ranks	Value
Rfree		0.298
Clashscore		8
Ramachandran outliers		1.5%
Sidechain outliers		3.2%
RSRZ outliers		15.5%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
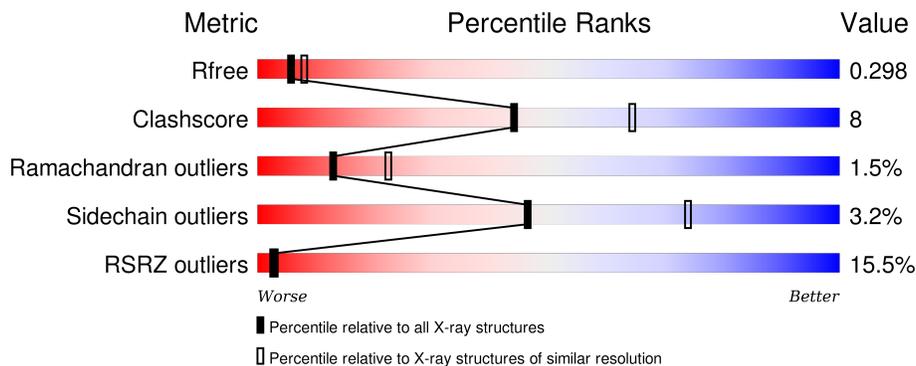
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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	275	20% / 83% / 16% ·
1	H	275	12% / 85% / 14% ·
2	B	100	11% / 88% / 10% ·
2	I	100	11% / 89% / 10% ·
3	C	9	67% / 89% / 11%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	9	67% ... 89% ... 11%
4	E	194	15% ... 68% ... 27% ... • •
4	L	194	18% ... 72% ... 21% ... 5% ... •
5	F	238	18% ... 76% ... 16% ... 5% •
5	M	238	9% ... 84% ... 14% ... •

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 13180 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 I HISTOCOMPATIBILITY ANTIGEN, A-2 ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	275	Total 2238	C 1395	N 408	O 426	S 9	4	0	0
1	H	275	Total 2238	C 1395	N 408	O 426	S 9	3	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	167	ALA	TRP	ENGINEERED MUTATION	UNP P01892
H	167	ALA	TRP	ENGINEERED MUTATION	UNP P01892

- Molecule 2 is a protein called BETA-2-MICROGLOBULIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total 837	C 533	N 141	O 159	S 4	0	0	0
2	I	100	Total 837	C 533	N 141	O 159	S 4	0	0	0

- Molecule 3 is a protein called P1049.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total 76	C 56	N 10	O 10	0	0	0
3	J	9	Total 76	C 56	N 10	O 10	0	0	0

- Molecule 4 is a protein called TCR ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	194	Total 1521	C 965	N 245	O 302	S 9	86	0	0
4	L	194	Total 1521	C 965	N 245	O 302	S 9	87	0	0

- Molecule 5 is a protein called TCR BETA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	237	Total 1891	C 1194	N 331	O 361	S 5	0	0	0
5	M	237	Total 1891	C 1194	N 331	O 361	S 5	3	0	0

- Molecule 6 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	9	Total 9	O 9	0	0
6	B	1	Total 1	O 1	0	0
6	E	10	Total 10	O 10	0	0
6	F	5	Total 5	O 5	0	0
6	H	9	Total 9	O 9	0	0
6	I	6	Total 6	O 6	0	0
6	L	9	Total 9	O 9	0	0
6	M	5	Total 5	O 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 I HISTOCOMPATIBILITY ANTIGEN, A-2 ALPHA CHAIN
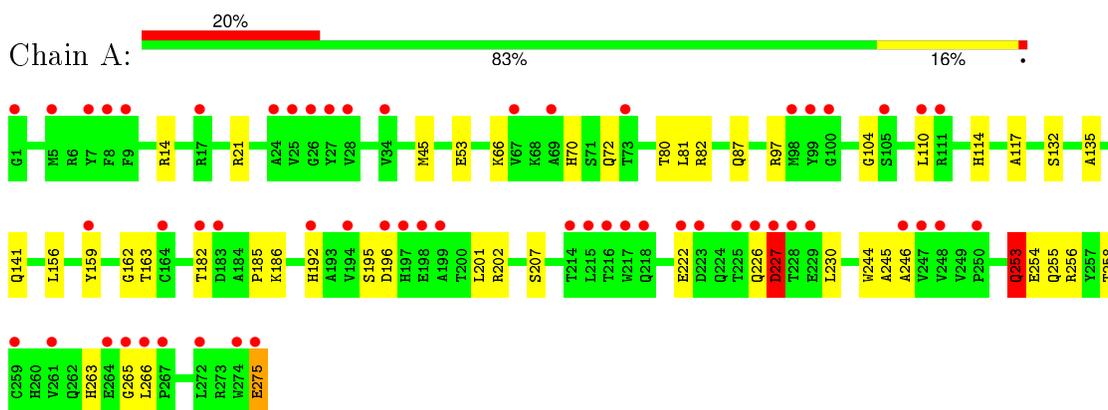
- Molecule 1: HLA CLASS I HISTOCOMPATIBILITY ANTIGEN, A-2 ALPHA CHAIN
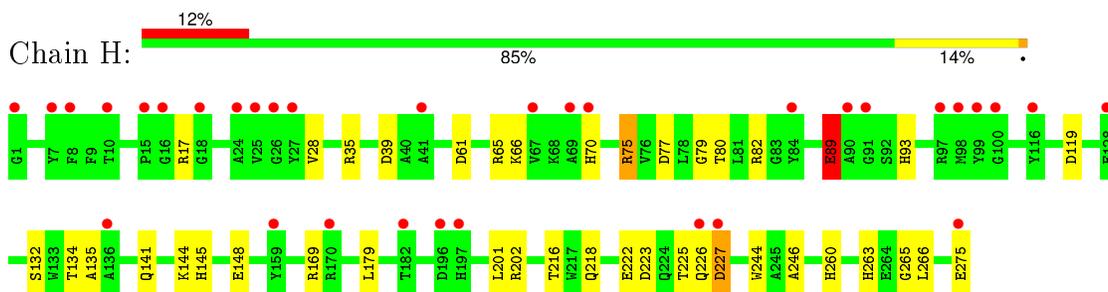

- Molecule 2: BETA-2-MICROGLOBULIN
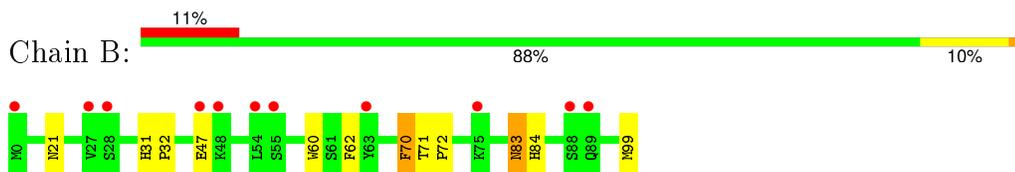

- Molecule 2: BETA-2-MICROGLOBULIN
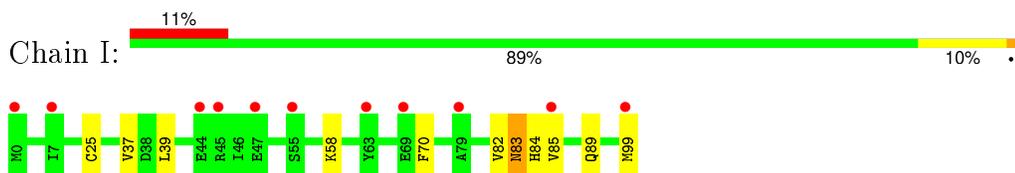
- Molecule 3: P1049
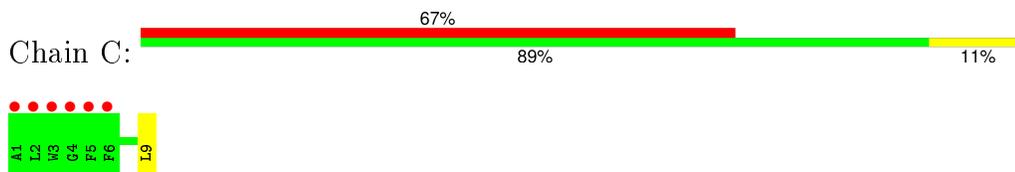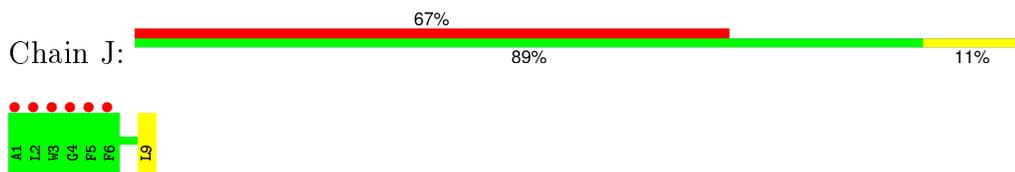
Chain C:

- Molecule 3: P1049

Chain J:

- Molecule 4: TCR ALPHA

Chain E:

- Molecule 4: TCR ALPHA

Chain L:

- Molecule 5: TCR BETA

Chain F:

- Molecule 5: TCR BETA

Chain M:
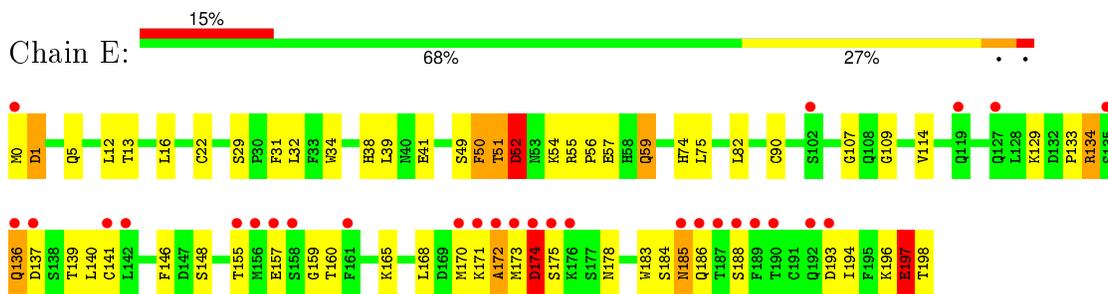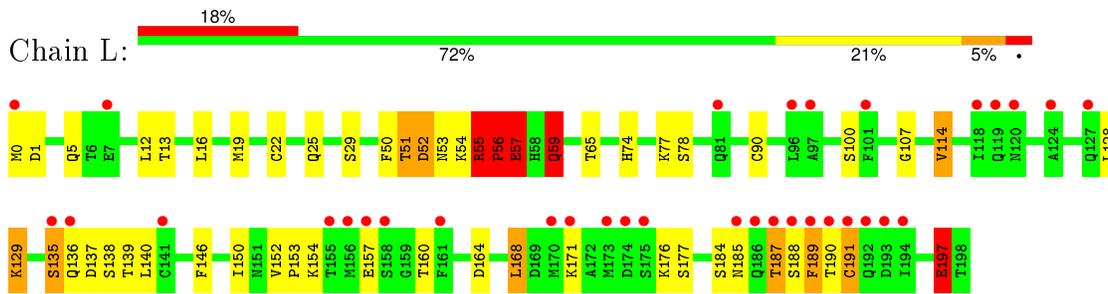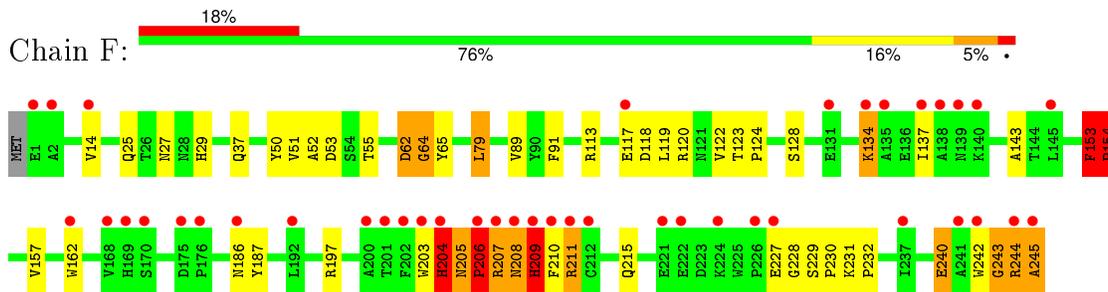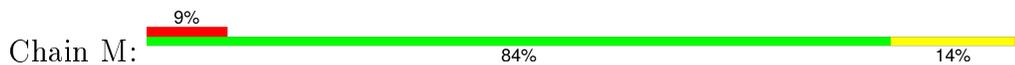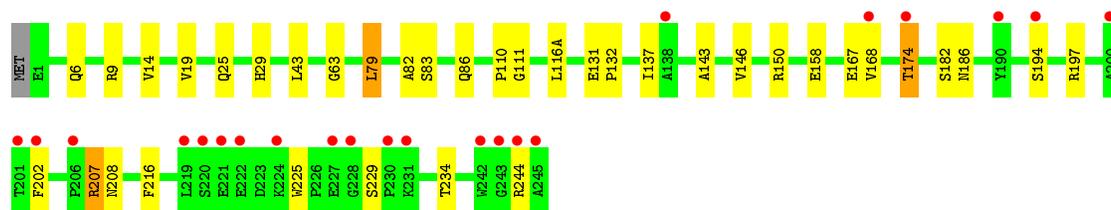
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	94.28Å 84.35Å 122.47Å 90.00° 92.53° 90.00°	Depositor
Resolution (Å)	122.17 – 2.50 40.78 – 2.50	Depositor EDS
% Data completeness (in resolution range)	92.5 (122.17-2.50) 92.5 (40.78-2.50)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.242 , 0.292 0.268 , 0.298	Depositor DCC
R_{free} test set	3091 reflections (5.02%)	DCC
Wilson B-factor (Å^2)	47.7	Xtriage
Anisotropy	0.558	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 32.0	EDS
Estimated twinning fraction	0.000 for h,-k,-l	Xtriage
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Outliers	2 of 61770 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	13180	wwPDB-VP
Average B, all atoms (Å^2)	50.0	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 42.51 % 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 2.0198e-04. 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 ⓘ

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	1.07	14/2301 (0.6%)	0.72	7/3121 (0.2%)
1	H	0.76	7/2301 (0.3%)	0.72	1/3121 (0.0%)
2	B	0.73	2/860 (0.2%)	0.65	0/1162
2	I	0.82	3/860 (0.3%)	0.85	1/1162 (0.1%)
3	C	0.81	0/80	0.62	0/108
3	J	0.76	0/80	0.62	0/108
4	E	1.40	24/1555 (1.5%)	1.12	12/2106 (0.6%)
4	L	1.22	14/1557 (0.9%)	1.54	31/2112 (1.5%)
5	F	1.60	29/1947 (1.5%)	1.13	18/2649 (0.7%)
5	M	0.68	3/1947 (0.2%)	0.59	1/2649 (0.0%)
All	All	1.10	96/13488 (0.7%)	0.95	71/18298 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
4	E	0	2
4	L	0	4
5	F	0	4
All	All	0	11

The worst 5 of 96 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	275	GLU	C-O	25.85	1.72	1.23
5	F	62	ASP	C-N	22.68	1.73	1.33
4	E	56	PRO	N-CD	21.81	1.78	1.47
5	F	117	GLU	CD-OE1	20.21	1.47	1.25
5	F	64	GLY	N-CA	19.33	1.75	1.46

The worst 5 of 71 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	244	ARG	NE-CZ-NH1	23.41	132.00	120.30
4	L	52	ASP	CA-C-N	-21.03	70.94	117.20
5	F	207	ARG	NE-CZ-NH2	-19.11	110.75	120.30
4	E	134	ARG	NE-CZ-NH1	19.05	129.82	120.30
2	I	99	MET	CA-C-O	18.29	158.52	120.10

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	253	GLN	Sidechain
4	E	197	GLU	Sidechain,Mainchain
5	F	153	PHE	Peptide
5	F	204	HIS	Mainchain
5	F	206	PRO	Mainchain

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	2238	0	2091	36	0
1	H	2238	0	2091	25	1
2	B	837	0	803	9	0
2	I	837	0	803	6	0
3	C	76	0	76	2	0
3	J	76	0	76	2	0
4	E	1521	0	1472	32	1
4	L	1521	0	1476	36	0
5	F	1891	0	1793	55	0
5	M	1891	0	1794	21	0
6	A	9	0	0	0	0
6	B	1	0	0	0	0
6	E	10	0	0	0	0
6	F	5	0	0	0	0
6	H	9	0	0	0	0
6	I	6	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	L	9	0	0	1	0
6	M	5	0	0	0	0
All	All	13180	0	12475	202	1

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 8.

The worst 5 of 202 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)
5:F:134:LYS:CE	5:F:134:LYS:NZ	1.70	1.51
5:F:64:GLY:CA	5:F:64:GLY:N	1.75	1.47
5:F:154:PRO:N	5:F:154:PRO:CD	1.68	1.46
1:A:253:GLN:CD	1:A:253:GLN:NE2	1.69	1.43
5:F:62:ASP:C	5:F:64:GLY:N	1.73	1.41

All (1) 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:E:198:THR:O	1:H:169:ARG:NH2[2_645]	2.14	0.06

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	273/275 (99%)	268 (98%)	3 (1%)	2 (1%)	26	46
1	H	273/275 (99%)	262 (96%)	9 (3%)	2 (1%)	26	46
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
2	I	98/100 (98%)	96 (98%)	2 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	J	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
4	E	189/194 (97%)	165 (87%)	17 (9%)	7 (4%)	4	5
4	L	192/194 (99%)	164 (85%)	17 (9%)	11 (6%)	2	2
5	F	235/238 (99%)	219 (93%)	14 (6%)	2 (1%)	21	37
5	M	235/238 (99%)	227 (97%)	8 (3%)	0	100	100
All	All	1607/1632 (98%)	1509 (94%)	74 (5%)	24 (2%)	13	22

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	E	51	THR
4	E	52	ASP
4	E	174	ASP
5	F	154	PRO
4	L	51	THR

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	230/230 (100%)	226 (98%)	4 (2%)	68	89
1	H	230/230 (100%)	223 (97%)	7 (3%)	48	76
2	B	95/95 (100%)	93 (98%)	2 (2%)	61	85
2	I	95/95 (100%)	92 (97%)	3 (3%)	46	74
3	C	7/7 (100%)	7 (100%)	0	100	100
3	J	7/7 (100%)	7 (100%)	0	100	100
4	E	177/177 (100%)	169 (96%)	8 (4%)	34	59
4	L	177/177 (100%)	166 (94%)	11 (6%)	23	41
5	F	205/206 (100%)	200 (98%)	5 (2%)	57	82
5	M	205/206 (100%)	199 (97%)	6 (3%)	50	77

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1428/1430 (100%)	1382 (97%)	46 (3%)	46 74

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	39	ASP
1	H	266	LEU
5	M	174	THR
1	H	75	ARG
1	H	132	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	139	ASN
1	H	93	HIS
5	M	86	GLN
1	H	70	HIS
1	H	141	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)

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(Å²)	Q<0.9
1	A	275/275 (100%)	1.18	56 (20%)	1	1	40, 50, 58, 67	1 (0%)
1	H	275/275 (100%)	0.98	33 (12%)	6	6	38, 50, 59, 74	1 (0%)
2	B	100/100 (100%)	0.81	11 (11%)	7	7	42, 49, 59, 73	0
2	I	100/100 (100%)	0.90	11 (11%)	7	7	43, 51, 60, 76	0
3	C	9/9 (100%)	2.13	6 (66%)	0	0	44, 45, 51, 54	0
3	J	9/9 (100%)	2.37	6 (66%)	0	0	48, 50, 54, 57	0
4	E	184/194 (94%)	1.20	29 (15%)	3	2	37, 49, 66, 75	0
4	L	185/194 (95%)	1.17	34 (18%)	2	2	36, 49, 67, 73	2 (1%)
5	F	237/238 (99%)	1.09	42 (17%)	2	2	35, 50, 63, 69	0
5	M	237/238 (99%)	0.80	22 (9%)	11	11	37, 49, 58, 66	1 (0%)
All	All	1611/1632 (98%)	1.05	250 (15%)	3	3	35, 50, 62, 76	5 (0%)

The worst 5 of 250 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	E	189	PHE	9.6
5	M	245	ALA	9.0
4	E	173	MET	8.6
5	F	245	ALA	8.1
4	L	156	MET	7.7

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.