wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 01:16 PM GMT

PDB ID	:	3TBW
Title	:	CRYSTAL STRUCTURE OF THE MURINE CLASS I MAJOR HISTOCOMPATIBILITY COMPLEX H-2DB IN COMPLEX WITH THE LCMV-DERIVED GP33 ALTERED PEPTIDE ligand (A2G, V3P, Y4S)
Authors	:	Duru, A.D.; Allerbring, E.B.; Uchtenhagen, H.; Mazumdar, P.A.; Badia-Martinez, D.; Madhurantakam, C.; Sandalova, T.; Nygren, P.; Achour, A.
Deposited on	:	2011-08-08
Resolution	:	2.15 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
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:

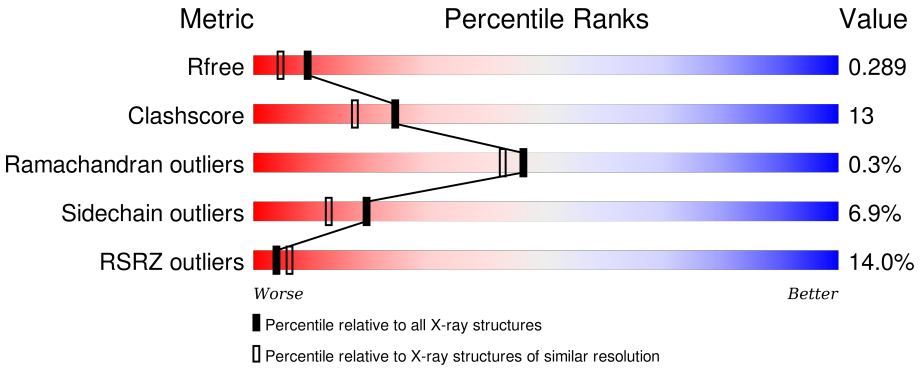
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.15 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	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (2.16-2.16)

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	338	12% 58% 20% 19%
1	C	338	15% 56% 22% 19%
1	E	338	16% 59% 18% 19%
1	G	338	13% 55% 22% 21%
2	B	99	5% 75% 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	99	5% 75% 22%
2	F	99	9% 74% 24%
2	H	99	4% 72% 27%
3	I	9	78% 11% 11%
3	J	9	44% 56%
3	K	9	67% 22% 11%
3	L	9	56% 33% 11%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 13054 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	14	0	0
			2248	1420	398	421	9			
1	C	274	Total	C	N	O	S	14	0	0
			2248	1420	398	421	9			
1	E	274	Total	C	N	O	S	30	0	0
			2248	1418	398	423	9			
1	G	267	Total	C	N	O	S	30	0	0
			2196	1387	390	410	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			820	524	138	151	7			
2	D	99	Total	C	N	O	S	0	0	0
			820	524	138	151	7			
2	F	99	Total	C	N	O	S	0	0	0
			820	524	138	151	7			
2	H	99	Total	C	N	O	S	0	0	0
			820	524	138	151	7			

- Molecule 3 is a protein called GLYCOPROTEIN GPC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	9	Total	C	N	O	S	0	0	0
			65	41	11	12	1			
3	J	9	Total	C	N	O	S	0	0	0
			65	41	11	12	1			
3	K	9	Total	C	N	O	S	0	0	0
			65	41	11	12	1			
3	L	9	Total	C	N	O	S	0	0	0
			65	41	11	12	1			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	2	GLY	ALA	ENGINEERED MUTATION	UNP P07399
I	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
I	4	SER	TYR	ENGINEERED MUTATION	UNP P07399
I	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
J	2	GLY	ALA	ENGINEERED MUTATION	UNP P07399
J	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
J	4	SER	TYR	ENGINEERED MUTATION	UNP P07399
J	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
K	2	GLY	ALA	ENGINEERED MUTATION	UNP P07399
K	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
K	4	SER	TYR	ENGINEERED MUTATION	UNP P07399
K	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
L	2	GLY	ALA	ENGINEERED MUTATION	UNP P07399
L	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
L	4	SER	TYR	ENGINEERED MUTATION	UNP P07399
L	9	MET	CYS	ENGINEERED MUTATION	UNP P07399

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	124	Total O 124 124	0	0
4	B	51	Total O 51 51	0	0
4	C	87	Total O 87 87	0	0
4	D	52	Total O 52 52	0	0
4	E	84	Total O 84 84	0	0
4	F	31	Total O 31 31	0	0
4	G	83	Total O 83 83	0	0
4	H	45	Total O 45 45	0	0
4	I	8	Total O 8 8	0	0
4	J	2	Total O 2 2	0	0
4	K	2	Total O 2 2	0	0

Continued on next page...

Continued from previous page...

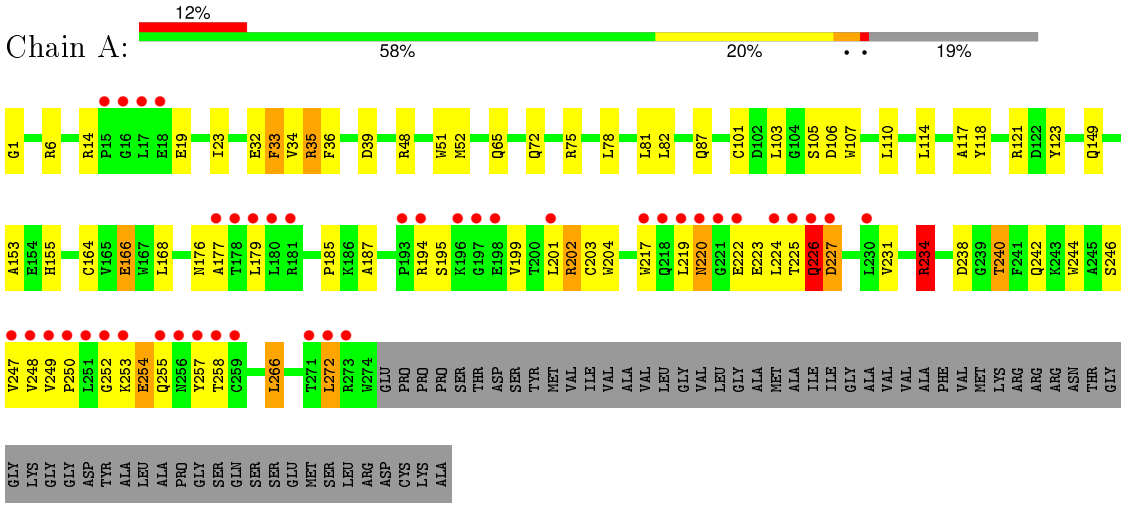
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	L	5	Total O 5 5	0	0

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 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.

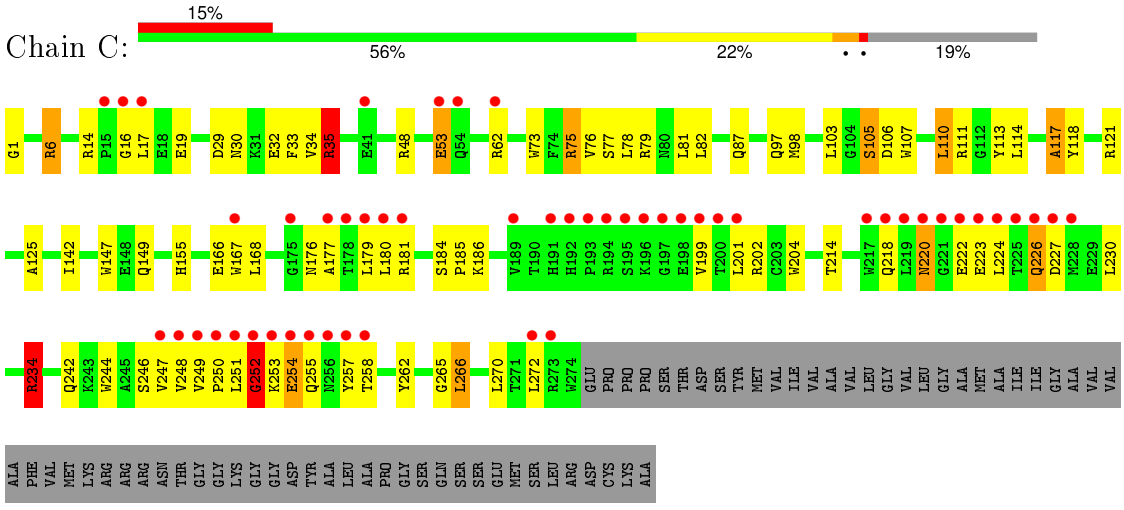
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain

Chain A:



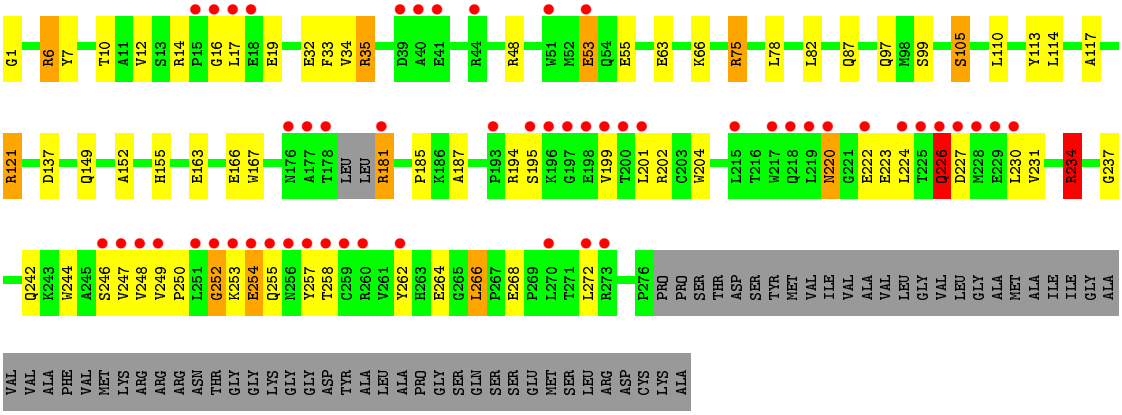
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain

Chain C:



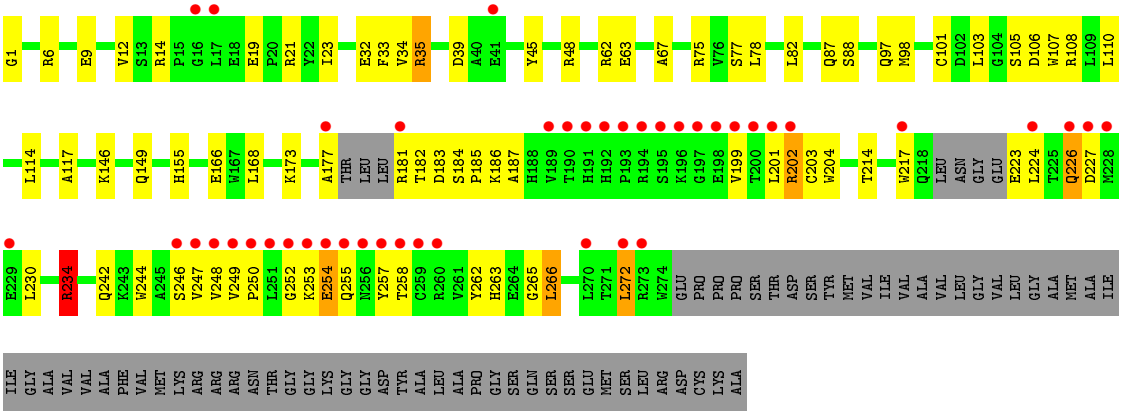
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain

Chain E:



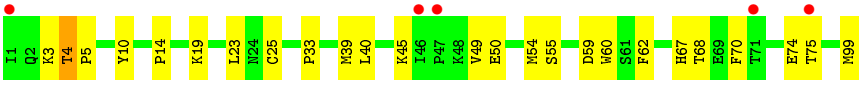
• Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain

Chain G:



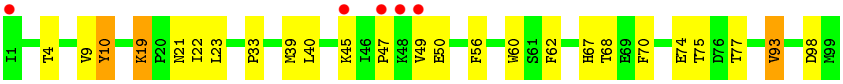
• Molecule 2: Beta-2-microglobulin

Chain B:



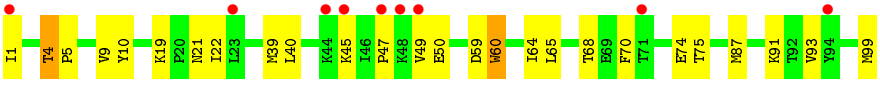
• Molecule 2: Beta-2-microglobulin

Chain D:



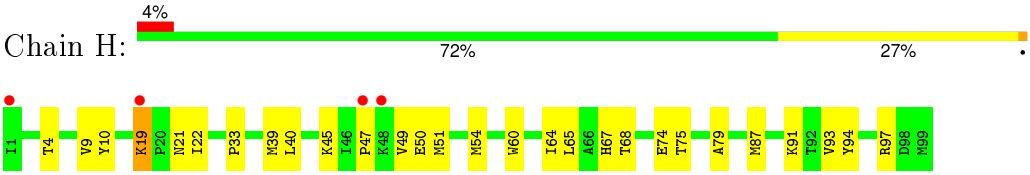
• Molecule 2: Beta-2-microglobulin

Chain F:



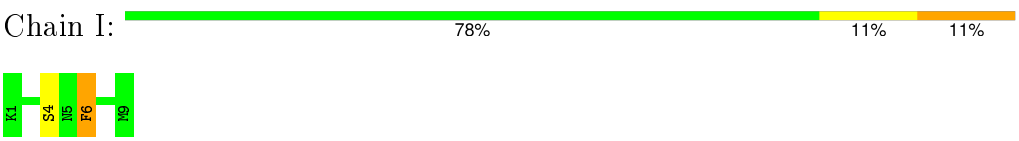
- Molecule 2: Beta-2-microglobulin

Chain H:



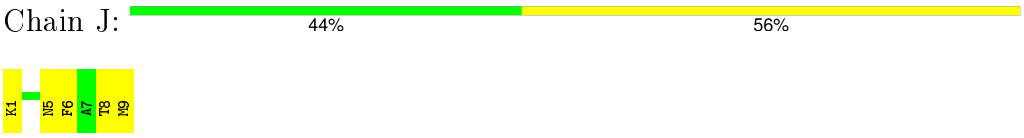
- Molecule 3: GLYCOPROTEIN GPC

Chain I:



- Molecule 3: GLYCOPROTEIN GPC

Chain J:



- Molecule 3: GLYCOPROTEIN GPC

Chain K:

- Molecule 3: GLYCOPROTEIN GPC

Chain L:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	91.40Å 124.28Å 99.89Å 90.00° 103.23° 90.00°	Depositor
Resolution (Å)	47.48 – 2.15 47.48 – 2.15	Depositor EDS
% Data completeness (in resolution range)	92.6 (47.48-2.15) 98.4 (47.48-2.15)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.24 (at 2.16Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.232 , 0.283 0.244 , 0.289	Depositor DCC
R_{free} test set	5819 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	35.4	Xtriage
Anisotropy	0.422	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 56.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Outliers	3 of 116037 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13054	wwPDB-VP
Average B, all atoms (Å^2)	56.0	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 58.42 % 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 2.0489e-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, 0.375 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	1.00	4/2314 (0.2%)	1.05	10/3142 (0.3%)
1	C	0.92	3/2314 (0.1%)	0.90	10/3142 (0.3%)
1	E	0.93	2/2314 (0.1%)	0.93	12/3141 (0.4%)
1	G	0.91	2/2260 (0.1%)	0.89	6/3065 (0.2%)
2	B	1.01	1/846 (0.1%)	0.89	1/1148 (0.1%)
2	D	1.02	3/846 (0.4%)	0.89	0/1148
2	F	0.93	1/846 (0.1%)	0.87	1/1148 (0.1%)
2	H	0.91	0/846	0.89	1/1148 (0.1%)
3	I	1.96	2/66 (3.0%)	1.16	0/87
3	J	1.34	0/66	1.17	0/87
3	K	1.37	0/66	1.20	0/87
3	L	1.35	0/66	1.27	1/87 (1.1%)
All	All	0.96	18/12850 (0.1%)	0.94	42/17430 (0.2%)

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
1	C	0	1
1	E	0	2
1	G	0	1
All	All	0	5

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	4	SER	CA-CB	9.23	1.66	1.52
2	B	25	CYS	CB-SG	-7.80	1.69	1.82
1	E	152	ALA	CA-CB	7.47	1.68	1.52
1	A	164	CYS	CB-SG	7.12	1.94	1.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	7	TYR	CD2-CE2	6.79	1.49	1.39

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	234	ARG	NE-CZ-NH1	19.05	129.82	120.30
1	A	234	ARG	NE-CZ-NH2	-15.87	112.36	120.30
1	A	35	ARG	NE-CZ-NH2	-15.47	112.56	120.30
1	A	35	ARG	NE-CZ-NH1	12.80	126.70	120.30
1	E	234	ARG	NE-CZ-NH2	-12.28	114.16	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	252	GLY	Peptide
1	C	252	GLY	Peptide
1	E	252	GLY	Peptide
1	E	53	GLU	Peptide
1	G	252	GLY	Peptide

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	2248	0	2123	57	0
1	C	2248	0	2123	65	0
1	E	2248	0	2113	57	0
1	G	2196	0	2066	74	0
2	B	820	0	796	20	0
2	D	820	0	796	23	0
2	F	820	0	796	23	0
2	H	820	0	796	33	0
3	I	65	0	66	1	0
3	J	65	0	66	9	0
3	K	65	0	66	8	0
3	L	65	0	66	7	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	124	0	0	11	0
4	B	51	0	0	2	0
4	C	87	0	0	5	0
4	D	52	0	0	5	0
4	E	84	0	0	4	0
4	F	31	0	0	3	0
4	G	83	0	0	9	0
4	H	45	0	0	3	0
4	I	8	0	0	0	0
4	J	2	0	0	0	0
4	K	2	0	0	0	0
4	L	5	0	0	0	0
All	All	13054	0	11873	315	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 13.

The worst 5 of 315 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:H:39:MET:HE2	2:H:49:VAL:HG13	1.41	1.02
2:D:39:MET:SD	4:D:190:HOH:O	2.18	1.01
2:F:39:MET:HE2	2:F:49:VAL:HG13	1.46	0.97
1:G:184:SER:HA	4:G:389:HOH:O	1.73	0.89
2:D:39:MET:HE2	2:D:49:VAL:HG13	1.56	0.87

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	272/338 (80%)	254 (93%)	17 (6%)	1 (0%)	39 34
1	C	272/338 (80%)	254 (93%)	18 (7%)	0	100 100
1	E	270/338 (80%)	253 (94%)	16 (6%)	1 (0%)	39 34
1	G	261/338 (77%)	246 (94%)	15 (6%)	0	100 100
2	B	97/99 (98%)	93 (96%)	4 (4%)	0	100 100
2	D	97/99 (98%)	94 (97%)	2 (2%)	1 (1%)	19 11
2	F	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	19 11
2	H	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	19 11
3	I	7/9 (78%)	5 (71%)	2 (29%)	0	100 100
3	J	7/9 (78%)	6 (86%)	1 (14%)	0	100 100
3	K	7/9 (78%)	6 (86%)	1 (14%)	0	100 100
3	L	7/9 (78%)	6 (86%)	1 (14%)	0	100 100
All	All	1491/1784 (84%)	1403 (94%)	83 (6%)	5 (0%)	46 42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	GLN
1	E	226	GLN
2	H	47	PRO
2	D	47	PRO
2	F	47	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	232/280 (83%)	215 (93%)	17 (7%)	17 11
1	C	232/280 (83%)	214 (92%)	18 (8%)	16 9
1	E	232/280 (83%)	216 (93%)	16 (7%)	19 13
1	G	226/280 (81%)	212 (94%)	14 (6%)	23 17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	B	94/94 (100%)	87 (93%)	7 (7%)	17 11
2	D	94/94 (100%)	88 (94%)	6 (6%)	22 15
2	F	94/94 (100%)	87 (93%)	7 (7%)	17 11
2	H	94/94 (100%)	88 (94%)	6 (6%)	22 15
3	I	7/7 (100%)	7 (100%)	0	100 100
3	J	7/7 (100%)	7 (100%)	0	100 100
3	K	7/7 (100%)	6 (86%)	1 (14%)	4 1
3	L	7/7 (100%)	7 (100%)	0	100 100
All	All	1326/1524 (87%)	1234 (93%)	92 (7%)	19 13

5 of 92 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	272	LEU
1	E	114	LEU
2	H	4	THR
2	D	4	THR
2	D	74	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	30	ASN
1	E	97	GLN
1	G	97	GLN
1	C	220	ASN
1	G	30	ASN

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

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	273/338 (80%)	1.02	41 (15%) 3 5	20, 49, 123, 195	0
1	C	273/338 (80%)	1.34	52 (19%) 2 2	22, 51, 130, 169	0
1	E	271/338 (80%)	1.15	53 (19%) 1 2	25, 52, 125, 160	0
1	G	264/338 (78%)	1.32	43 (16%) 2 4	22, 51, 123, 159	0
2	B	99/99 (100%)	0.65	5 (5%) 32 42	26, 46, 74, 100	0
2	D	99/99 (100%)	0.58	5 (5%) 32 42	27, 43, 76, 102	0
2	F	99/99 (100%)	0.99	9 (9%) 11 17	30, 51, 78, 103	0
2	H	99/99 (100%)	0.69	4 (4%) 42 52	28, 48, 79, 102	0
3	I	9/9 (100%)	0.28	0 100 100	21, 27, 36, 37	0
3	J	9/9 (100%)	0.17	0 100 100	25, 33, 44, 61	0
3	K	9/9 (100%)	0.26	0 100 100	31, 37, 42, 74	0
3	L	9/9 (100%)	0.29	0 100 100	27, 32, 40, 48	0
All	All	1513/1784 (84%)	1.06	212 (14%) 4 6	20, 49, 118, 195	0

The worst 5 of 212 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	ALA	12.6
1	A	17	LEU	12.6
1	E	178	THR	12.0
1	G	177	ALA	11.1
1	G	199	VAL	10.7

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 (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.