W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:51 AM GMT

PDB ID	:	3JTT
Title	:	Cystal structure of Rhesus macaque MHC class I:Mamu-A*02
Authors	:	Dai, L.; Feng, Y.; Qi, J.; Gao, G.F.
Deposited on	:	2009-09-14
Resolution	:	2.80 Å(reported)

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

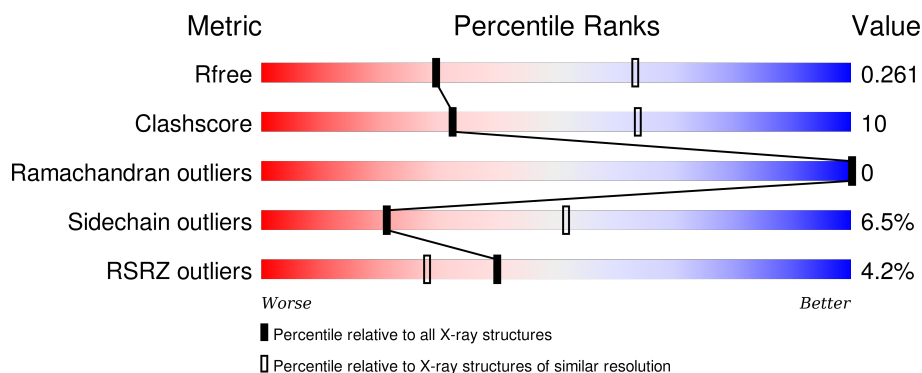|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.80 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		0.261
Clashscore		10
Ramachandran outliers		0
Sidechain outliers		6.5%
RSRZ outliers		4.2%

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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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. 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	276	3% 79% 18% •
1	D	276	5% 78% 19% •
1	G	276	5% 80% 17% •
2	B	100	6% 84% 15% •
2	E	100	3% 82% 16% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	100	82% 14% •
3	C	9	56% 22% 22%
3	F	9	33% 44% 22%
3	I	9	22% 56% 22% 22%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9756 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 MHC class I Mamu-A*02.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total 2261	C 1402	N 415	O 434	S 10	0	0	0
1	D	276	Total 2261	C 1402	N 415	O 434	S 10	0	0	0
1	G	276	Total 2261	C 1402	N 415	O 434	S 10	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total 830	C 528	N 141	O 156	S 5	0	0	0
2	E	100	Total 830	C 528	N 141	O 156	S 5	0	0	0
2	H	100	Total 830	C 528	N 141	O 156	S 5	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	EXPRESSION TAG	UNP Q6V7J5
E	1	MET	-	EXPRESSION TAG	UNP Q6V7J5
H	1	MET	-	EXPRESSION TAG	UNP Q6V7J5

- Molecule 3 is a protein called peptide of Protein Nef.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total 71	C 46	N 12	O 13	0	0	0
3	F	9	Total 71	C 46	N 12	O 13	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	9	Total 71	C 46	N 12	O 13	0	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	68	Total 68	O 68	0	0
4	B	27	Total 27	O 27	0	0
4	C	2	Total 2	O 2	0	0
4	D	47	Total 47	O 47	0	0
4	E	27	Total 27	O 27	0	0
4	F	4	Total 4	O 4	0	0
4	G	54	Total 54	O 54	0	0
4	H	37	Total 37	O 37	0	0
4	I	4	Total 4	O 4	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.
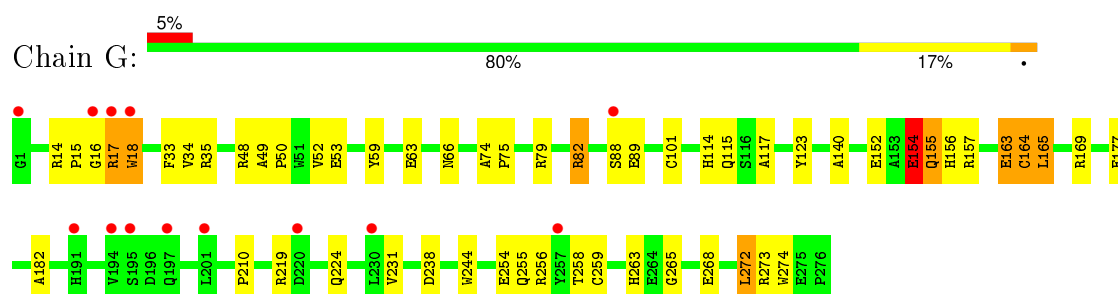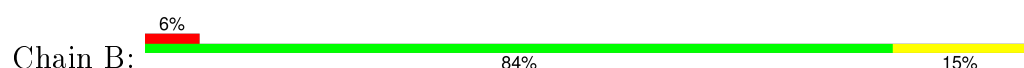
- Molecule 1: MHC class I Mamu-A*02
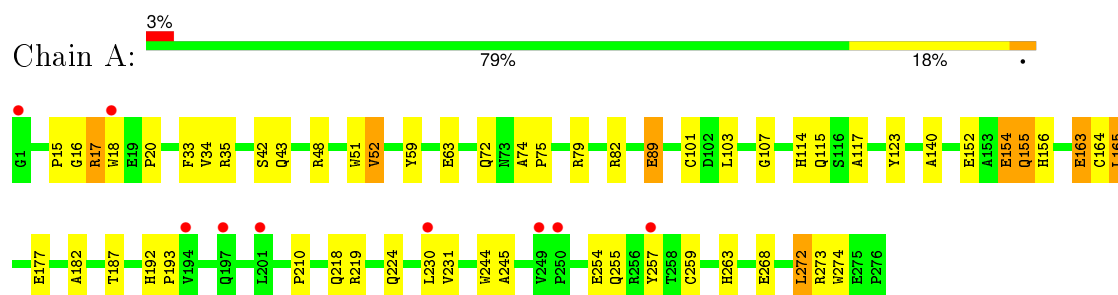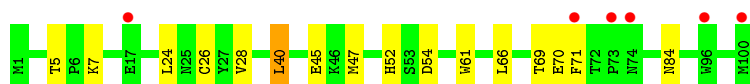

- Molecule 1: MHC class I Mamu-A*02
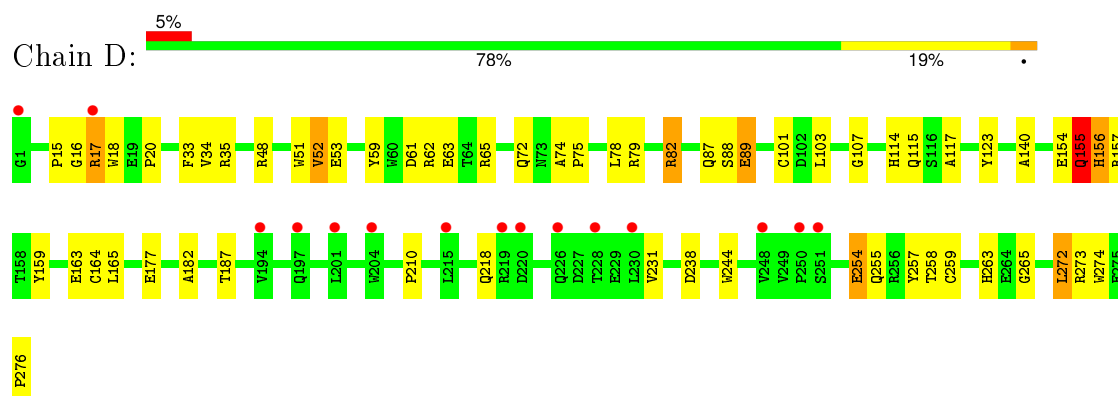
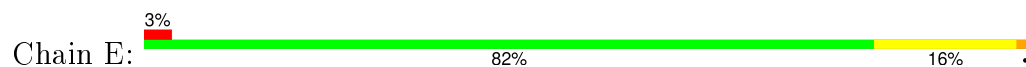- Molecule 1: MHC class I Mamu-A*02

- Molecule 2: Beta-2-microglobulin

• Molecule 2: Beta-2-microglobulin

Chain E:

• Molecule 2: Beta-2-microglobulin

Chain H:

• Molecule 3: peptide of Protein Nef

Chain C:
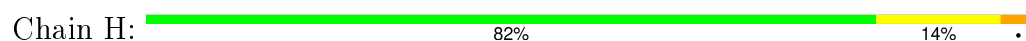
• Molecule 3: peptide of Protein Nef

Chain F:

• Molecule 3: peptide of Protein Nef

Chain I:
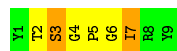
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	128.99Å　　129.01Å　　129.03Å 90.00°　　　90.00°　　　90.00°	Depositor
Resolution (Å)	31.29　–　2.80 31.29　–　2.80	Depositor EDS
% Data completeness (in resolution range)	95.2 (31.29-2.80) 99.5 (31.29-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$<I/\sigma(I)>$ [1]	4.44 (at 2.81Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.214　,　0.255 0.216　,　0.261	Depositor DCC
R_{free} test set	2702 reflections (5.07%)	DCC
Wilson B-factor (Å^2)	60.5	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 40.9	EDS
Estimated twinning fraction	0.026 for -h,l,k 0.026 for -l,-k,-h 0.026 for k,h,-l 0.488 for k,l,h 0.488 for l,h,k	Xtriage
L-test for twinning [2]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Outliers	0 of 53358 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9756	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.27	0/2324	0.42	0/3156
1	D	0.29	0/2324	0.48	3/3156 (0.1%)
1	G	0.27	0/2324	0.51	2/3156 (0.1%)
2	B	0.22	0/855	0.52	2/1156 (0.2%)
2	E	0.22	0/855	0.42	1/1156 (0.1%)
2	H	0.31	0/855	0.46	1/1156 (0.1%)
3	C	0.23	0/73	0.44	0/98
3	F	0.88	0/73	2.72	4/98 (4.1%)
3	I	0.54	0/73	1.53	3/98 (3.1%)
All	All	0.28	0/9756	0.54	16/13230 (0.1%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	3	SER	CB-CA-C	-19.24	73.54	110.10
1	G	155	GLN	N-CA-CB	-13.95	85.50	110.60
3	F	5	PRO	N-CA-C	-11.68	81.72	112.10
1	D	155	GLN	CB-CA-C	-9.85	90.70	110.40
2	B	47	MET	CB-CA-C	-9.61	91.17	110.40

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	2261	0	2103	58	0
1	D	2261	0	2103	56	0
1	G	2261	0	2103	53	0
2	B	830	0	792	8	0
2	E	830	0	792	10	0
2	H	830	0	792	11	0
3	C	71	0	69	8	0
3	F	71	0	69	5	0
3	I	71	0	69	9	0
4	A	68	0	0	0	0
4	B	27	0	0	0	0
4	C	2	0	0	0	0
4	D	47	0	0	0	0
4	E	27	0	0	1	0
4	F	4	0	0	0	0
4	G	54	0	0	0	0
4	H	37	0	0	1	0
4	I	4	0	0	0	0
All	All	9756	0	8892	190	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.

The worst 5 of 190 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:G:101:CYS:CB	1:G:164:CYS:SG	2.21	1.27
1:A:101:CYS:CB	1:A:164:CYS:SG	2.33	1.16
1:D:89:GLU:H	1:D:89:GLU:CD	1.51	1.09
3:F:7:ILE:HD13	3:F:7:ILE:N	1.61	1.09
1:D:101:CYS:CB	1:D:164:CYS:SG	2.41	1.07

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	274/276 (99%)	259 (94%)	15 (6%)	0	100	100
1	D	274/276 (99%)	258 (94%)	16 (6%)	0	100	100
1	G	274/276 (99%)	261 (95%)	13 (5%)	0	100	100
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
2	E	98/100 (98%)	93 (95%)	5 (5%)	0	100	100
2	H	98/100 (98%)	94 (96%)	4 (4%)	0	100	100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
3	I	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	1137/1155 (98%)	1078 (95%)	59 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	233/233 (100%)	221 (95%)	12 (5%)	29	62
1	D	233/233 (100%)	219 (94%)	14 (6%)	24	56
1	G	233/233 (100%)	218 (94%)	15 (6%)	22	52
2	B	94/94 (100%)	88 (94%)	6 (6%)	22	52
2	E	94/94 (100%)	88 (94%)	6 (6%)	22	52
2	H	94/94 (100%)	87 (93%)	7 (7%)	17	43
3	C	7/7 (100%)	5 (71%)	2 (29%)	0	1
3	F	7/7 (100%)	6 (86%)	1 (14%)	4	12
3	I	7/7 (100%)	5 (71%)	2 (29%)	0	1
All	All	1002/1002 (100%)	937 (94%)	65 (6%)	21	52

wwPDB X-ray Structure Validation Summary Report

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	155	GLN
2	E	45	GLU
2	H	71	PHE
1	D	165	LEU
1	D	272	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	80	ASN
1	D	156	HIS
1	G	263	HIS
1	D	115	GLN
1	D	155	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(Å^2)	Q<0.9
1	A	276/276 (100%)	0.64	9 (3%)	50	38	35, 53, 97, 113	0
1	D	276/276 (100%)	0.57	15 (5%)	29	19	35, 53, 97, 115	0
1	G	276/276 (100%)	0.60	13 (4%)	35	24	35, 53, 97, 114	0
2	B	100/100 (100%)	0.60	6 (6%)	25	15	36, 51, 85, 97	0
2	E	100/100 (100%)	0.63	3 (3%)	54	41	37, 51, 85, 102	0
2	H	100/100 (100%)	0.52	0	100	100	36, 51, 84, 96	0
3	C	9/9 (100%)	0.65	0	100	100	41, 52, 60, 65	0
3	F	9/9 (100%)	0.64	0	100	100	41, 56, 68, 77	0
3	I	9/9 (100%)	1.05	2 (22%)	1	1	41, 57, 65, 67	0
All	All	1155/1155 (100%)	0.60	48 (4%)	40	28	35, 53, 95, 115	0

The worst 5 of 48 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	GLY	4.6
1	D	197	GLN	4.2
1	D	1	GLY	4.2
1	G	197	GLN	3.9
1	A	230	LEU	3.7

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.