wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 07:39 am BST

PDB ID	:	5M00
Title	:	Crystal structure of murine P14 TCR complex with H-2Db and Y4A, modified gp33 peptide from LCMV
Authors	:	Achour, A.; Sandalova, T.; Sun, R.; Han, X.
Deposited on	:	2016-10-03
Resolution	:	1.95 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:

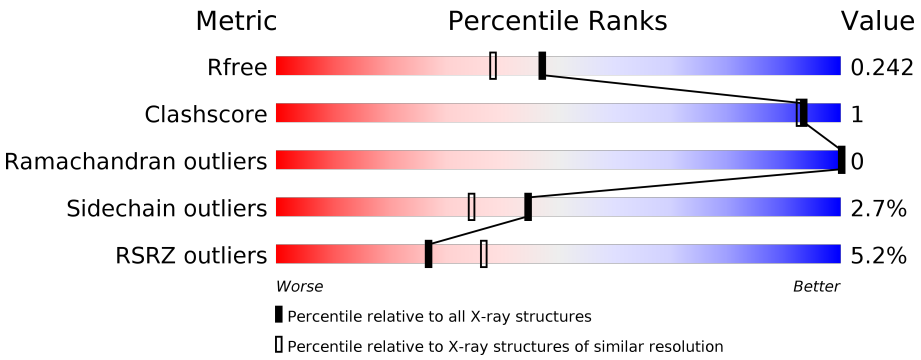
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 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	276	4%, 92%, 6% •
2	B	119	5%, 78%, 6% •, 15%
3	G	205	9%, 84%, 5% •, 9%
4	H	238	3%, 96%, •
5	P	9	100%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6906 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
			Total	C	N	O	S			
1	A	276	2265	1430	400	426	9	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	101	833	531	140	154	8	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-19	MET	-	initiating methionine	UNP P01887
B	-18	ALA	-	expression tag	UNP P01887
B	-17	ARG	-	expression tag	UNP P01887
B	-16	SER	-	expression tag	UNP P01887
B	-15	VAL	-	expression tag	UNP P01887
B	-14	THR	-	expression tag	UNP P01887
B	-13	LEU	-	expression tag	UNP P01887
B	-12	VAL	-	expression tag	UNP P01887
B	-11	PHE	-	expression tag	UNP P01887
B	-10	LEU	-	expression tag	UNP P01887
B	-9	VAL	-	expression tag	UNP P01887
B	-8	LEU	-	expression tag	UNP P01887
B	-7	VAL	-	expression tag	UNP P01887
B	-6	SER	-	expression tag	UNP P01887
B	-5	LEU	-	expression tag	UNP P01887
B	-4	THR	-	expression tag	UNP P01887
B	-3	GLY	-	expression tag	UNP P01887
B	-2	LEU	-	expression tag	UNP P01887
B	-1	MET	-	cloning artifact	UNP P01887
B	0	GLY	-	cloning artifact	UNP P01887

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	85	ASP	ALA	conflict	UNP P01887

- Molecule 3 is a protein called Protein Trav14-1,Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	G	187	1468	933	236	292	7	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	100	LEU	-	linker	UNP A0A0G2JF94
G	101	TYR	-	linker	UNP A0A0G2JF94
G	102	GLY	-	linker	UNP A0A0G2JF94
G	103	ASN	-	linker	UNP A0A0G2JF94
G	104	GLU	-	linker	UNP A0A0G2JF94
G	105	LYS	-	linker	UNP A0A0G2JF94
G	106	ILE	-	linker	UNP A0A0G2JF94
G	107	THR	-	linker	UNP A0A0G2JF94
G	108	PHE	-	linker	UNP A0A0G2JF94
G	109	GLY	-	linker	UNP A0A0G2JF94
G	110	ALA	-	linker	UNP A0A0G2JF94
G	111	GLY	-	linker	UNP A0A0G2JF94
G	112	THR	-	linker	UNP A0A0G2JF94
G	113	LYS	-	linker	UNP A0A0G2JF94
G	114	LEU	-	linker	UNP A0A0G2JF94
G	115	THR	-	linker	UNP A0A0G2JF94
G	116	ILE	-	linker	UNP A0A0G2JF94
G	117	LYS	-	linker	UNP A0A0G2JF94
G	118	PRO	-	linker	UNP A0A0G2JF94
G	119	ASN	-	linker	UNP A0A0G2JF94
G	166	CYS	THR	conflict	UNP Q6PIR9

- Molecule 4 is a protein called T-cell receptor beta chain V region C5,Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	H	237	1893	1187	337	363	6	0	5	0

There are 19 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	93	ASP	-	linker	UNP P04213
H	94	ALA	-	linker	UNP P04213
H	95	GLY	-	linker	UNP P04213
H	96	GLY	-	linker	UNP P04213
H	97	ARG	-	linker	UNP P04213
H	98	ASN	-	linker	UNP P04213
H	99	THR	-	linker	UNP P04213
H	100	LEU	-	linker	UNP P04213
H	101	TYR	-	linker	UNP P04213
H	102	PHE	-	linker	UNP P04213
H	103	GLY	-	linker	UNP P04213
H	104	ALA	-	linker	UNP P04213
H	105	GLY	-	linker	UNP P04213
H	106	THR	-	linker	UNP P04213
H	107	ARG	-	linker	UNP P04213
H	108	LEU	-	linker	UNP P04213
H	109	SER	-	linker	UNP P04213
H	168	CYS	SER	conflict	UNP Q7TND8
H	182	SER	CYS	conflict	UNP Q7TND8

- Molecule 5 is a protein called LCMV-DERIVED GP33 ALTERED PEPTIDE LIGAND Y4A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	P	9	66	42	11	12	1	0	0	0

- Molecule 6 is water.

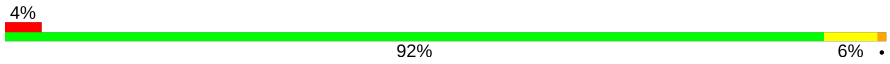
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	150	150	150	0	0
6	B	47	47	47	0	0
6	G	62	62	62	0	0
6	H	117	117	117	0	0
6	P	5	5	5	0	0

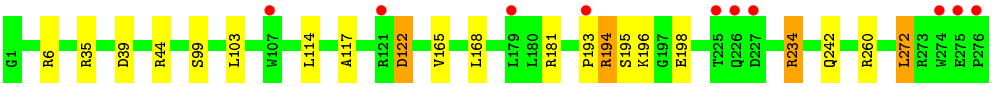
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: H-2 class I histocompatibility antigen, D-B alpha chain

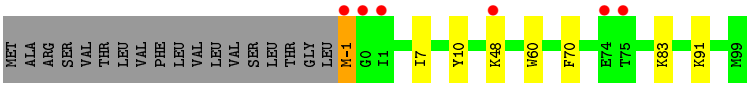
Chain A:





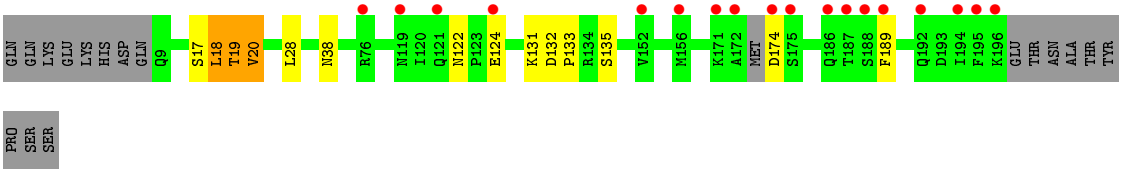
- Molecule 2: Beta-2-microglobulin

Chain B:



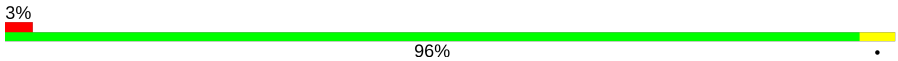
- Molecule 3: Protein Trav14-1,Uncharacterized protein

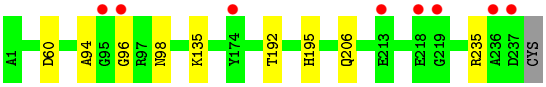
Chain G:



- Molecule 4: T-cell receptor beta chain V region C5,Uncharacterized protein

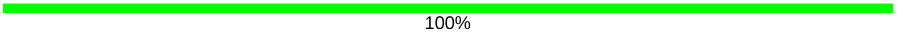
Chain H:





- Molecule 5: LCMV-DERIVED GP33 ALTERED PEPTIDE LIGAND Y4A

Chain P:



There are no outlier residues recorded for this chain.

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	258.45Å 46.94Å 90.25Å 90.00° 94.24° 90.00°	Depositor
Resolution (Å)	128.87 – 1.95 54.33 – 1.95	Depositor EDS
% Data completeness (in resolution range)	97.6 (128.87-1.95) 97.6 (54.33-1.95)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.85 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.200 , 0.237 0.208 , 0.242	Depositor DCC
R_{free} test set	3894 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	32.9	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 35.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6906	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.26% 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	0.76	1/2332 (0.0%)	0.95	10/3166 (0.3%)
2	B	0.68	0/859	0.79	0/1163
3	G	0.71	1/1502 (0.1%)	0.81	3/2036 (0.1%)
4	H	0.71	0/1956	0.82	1/2657 (0.0%)
5	P	0.59	0/66	0.73	0/86
All	All	0.72	2/6715 (0.0%)	0.86	14/9108 (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	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	99	SER	CB-OG	-5.42	1.35	1.42
3	G	133	PRO	N-CD	5.15	1.55	1.47

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	234	ARG	NE-CZ-NH1	11.30	125.95	120.30
1	A	234	ARG	NE-CZ-NH2	-11.27	114.67	120.30
1	A	6	ARG	NE-CZ-NH1	7.38	123.99	120.30
1	A	260	ARG	NE-CZ-NH1	6.32	123.46	120.30
1	A	44	ARG	NE-CZ-NH1	5.87	123.23	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	193	PRO	Peptide
1	A	195	SER	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	2265	0	2136	9	0
2	B	833	0	808	7	0
3	G	1468	0	1410	3	0
4	H	1893	0	1789	4	0
5	P	66	0	70	0	0
6	A	150	0	0	0	0
6	B	47	0	0	0	0
6	G	62	0	0	0	0
6	H	117	0	0	1	0
6	P	5	0	0	0	0
All	All	6906	0	6213	18	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 1.

The worst 5 of 18 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:A:234:ARG:HD3	2:B:10:TYR:CE2	2.36	0.61
1:A:194:ARG:O	1:A:198:GLU:O	2.21	0.58
3:G:18:LEU:HD12	3:G:20:VAL:HG12	1.89	0.55
1:A:103:LEU:HD22	1:A:168:LEU:HD23	1.94	0.50
1:A:234:ARG:HD3	2:B:10:TYR:CZ	2.48	0.49

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	274/276 (99%)	268 (98%)	6 (2%)	0	100	100
2	B	99/119 (83%)	97 (98%)	2 (2%)	0	100	100
3	G	183/205 (89%)	178 (97%)	5 (3%)	0	100	100
4	H	240/238 (101%)	232 (97%)	8 (3%)	0	100	100
5	P	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	803/847 (95%)	781 (97%)	22 (3%)	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	234/234 (100%)	229 (98%)	5 (2%)	53	46
2	B	95/111 (86%)	92 (97%)	3 (3%)	39	27
3	G	166/184 (90%)	157 (95%)	9 (5%)	22	10
4	H	205/204 (100%)	203 (99%)	2 (1%)	76	74
5	P	6/6 (100%)	6 (100%)	0	100	100
All	All	706/739 (96%)	687 (97%)	19 (3%)	44	34

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	17	SER
3	G	19	THR
3	G	174	ASP
2	B	70	PHE
3	G	189	PHE

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

Mol	Chain	Res	Type
4	H	134	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 (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%)	0.29	10 (3%) 42 52	19, 35, 60, 87	0
2	B	101/119 (84%)	0.29	6 (5%) 22 30	21, 41, 66, 76	2 (1%)
3	G	187/205 (91%)	0.59	18 (9%) 8 13	25, 51, 92, 127	0
4	H	237/238 (99%)	0.28	8 (3%) 45 55	28, 37, 57, 80	0
5	P	9/9 (100%)	0.10	0 100 100	21, 24, 29, 31	0
All	All	810/847 (95%)	0.36	42 (5%) 27 37	19, 39, 74, 127	2 (0%)

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	172	ALA	6.6
3	G	187	THR	6.3
4	H	219	GLY	5.4
2	B	-1	MET	5.1
3	G	192	GLN	5.1

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.