wwPDB X-ray Structure Validation Summary Report (i)

Jun 25, 2024 – 03:28 PM EDT

PDB ID	:	5OPI
Title	:	Crystal structure of the TAPBPR-MHC I peptide editing complex
Authors	:	Thomas, C.; Tampe, R.
Deposited on	:	2017-08-09
Resolution	:	3.30 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

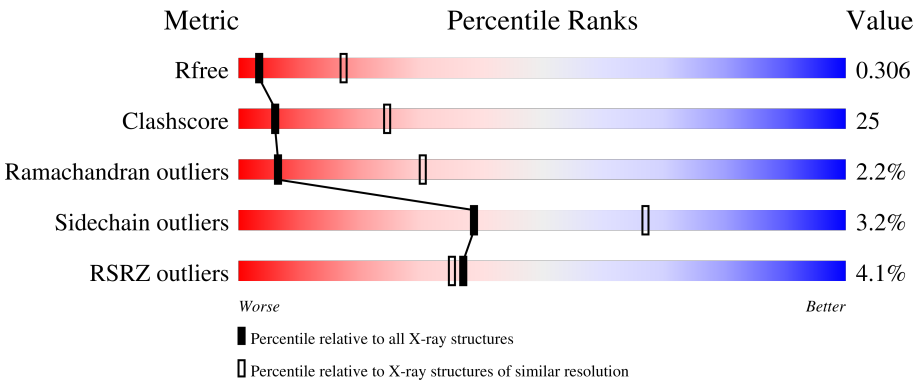
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	279	5% 72% 27% •
2	B	99	7% 65% 33% •
3	C	373	2% 49% 34% 6% 11%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5453 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	278	2257	1427	399	422	9	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	277	SER	-	expression tag	UNP P01899
A	278	SER	-	expression tag	UNP P01899
A	279	ALA	-	expression tag	UNP P01899

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	99	822	522	139	158	3	0	0	0

- Molecule 3 is a protein called TAP binding protein-like variant.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	332	2374	1490	407	466	11	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

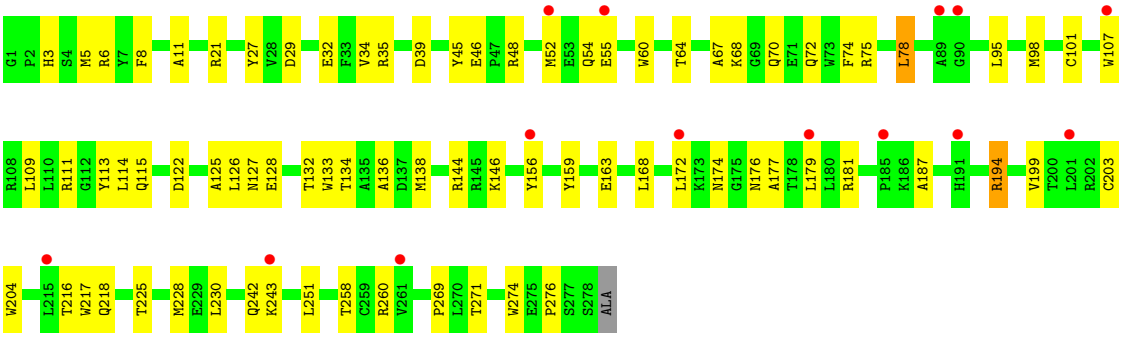
Chain	Residue	Modelled	Actual	Comment	Reference
C	11	ALA	-	expression tag	UNP Q53GH5
C	12	ASP	-	expression tag	UNP Q53GH5
C	13	PRO	-	expression tag	UNP Q53GH5
C	14	GLY	-	expression tag	UNP Q53GH5
C	97	ALA	CYS	engineered mutation	UNP Q53GH5
C	123	TRP	ARG	engineered mutation	UNP Q53GH5

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

Chain A:

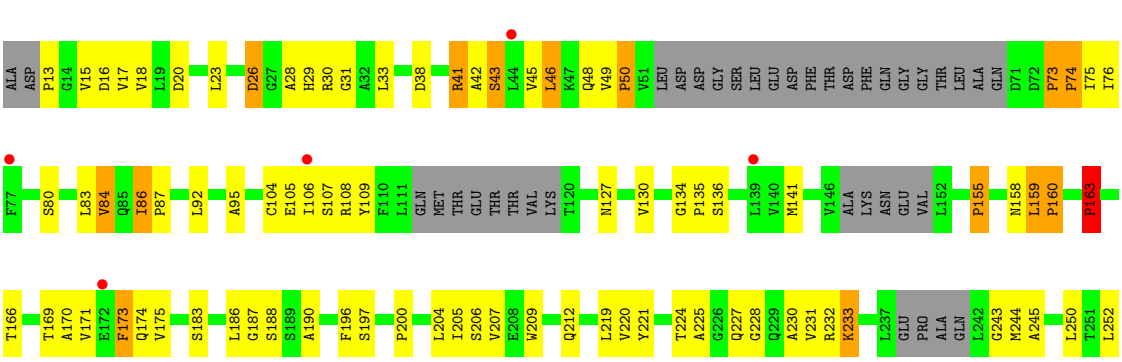


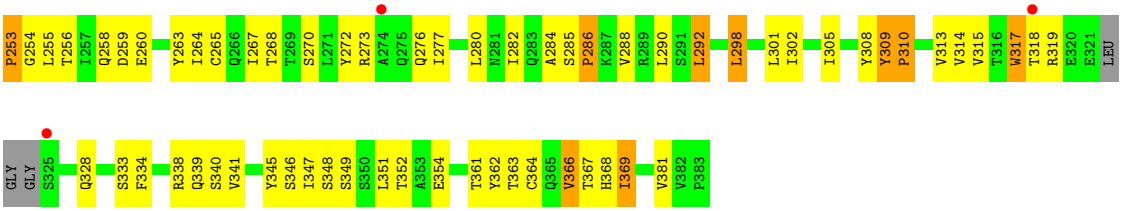
• Molecule 2: Beta-2-microglobulin

Chain B:

• Molecule 3: TAP binding protein-like variant

Chain C:




P253
G254
L255
T256
I257
Q258
D259
E260
Y263
I264
C265
Q266
I267
T268
T269
S270
L271
Y272
R273
A274
Q275
Q276
I277
L280
N281
I282
Q283
A284
S285
P286
K287
V288
R289
L290
S291
L292
L298
L301
I302
I305
Y308
Y309
P310
V313
V314
V315
T316
W317
T318
R319
E320
E321
LEU

GLY
GLY
S325
Q328
S333
F334
R338
Q339
S340
V341
Y345
S346
I347
S348
S349
S350
L351
T352
A353
E354
T361
Y362
T363
C364
Q365
V366
T367
H368
I369
V381
V382
P383

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	205.04Å 205.04Å 121.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.25 – 3.30 49.25 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.25-3.30) 100.0 (49.25-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.17 (at 3.33Å)	Xtriage
Refinement program	PHENIX (dev_2747: ???)	Depositor
R, R_{free}	0.242 , 0.303 0.255 , 0.306	Depositor DCC
R_{free} test set	1154 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	122.8	Xtriage
Anisotropy	0.461	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 137.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5453	wwPDB-VP
Average B, all atoms (Å^2)	153.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.17% 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	0.51	0/2324	0.83	1/3159 (0.0%)
2	B	0.45	0/845	0.94	0/1144
3	C	0.76	3/2410 (0.1%)	0.95	10/3286 (0.3%)
All	All	0.63	3/5579 (0.1%)	0.90	11/7589 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	310	PRO	N-CD	-15.98	1.25	1.47
3	C	74	PRO	N-CD	-9.79	1.34	1.47
3	C	317	TRP	CB-CG	-5.43	1.40	1.50

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	43	SER	N-CA-C	7.71	131.81	111.00
3	C	74	PRO	CA-N-CD	6.97	121.47	111.70
1	A	78	LEU	CB-CG-CD2	-6.50	99.95	111.00
3	C	155	PRO	N-CA-CB	6.45	111.04	103.30
3	C	309	TYR	C-N-CD	-6.26	106.82	120.60

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	2257	0	2120	68	0
2	B	822	0	774	38	0
3	C	2374	0	2267	163	0
All	All	5453	0	5161	263	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 25.

The worst 5 of 263 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:B:70:PHE:HE1	2:B:78:TYR:CZ	1.64	1.14
1:A:6:ARG:HG2	1:A:98:MET:CE	1.81	1.10
2:B:70:PHE:CE1	2:B:78:TYR:CZ	2.47	1.01
3:C:13:PRO:HB2	3:C:50:PRO:HG2	1.47	0.96
1:A:6:ARG:HG2	1:A:98:MET:HE2	1.49	0.92

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	276/279 (99%)	265 (96%)	10 (4%)	1 (0%)	34 66
2	B	97/99 (98%)	90 (93%)	6 (6%)	1 (1%)	15 46
3	C	320/373 (86%)	265 (83%)	42 (13%)	13 (4%)	3 17
All	All	693/751 (92%)	620 (90%)	58 (8%)	15 (2%)	6 30

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	C	86	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	159	LEU
3	C	160	PRO
3	C	163	PRO
3	C	41	ARG

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	231/236 (98%)	230 (100%)	1 (0%)	91	95
2	B	92/94 (98%)	91 (99%)	1 (1%)	73	85
3	C	242/306 (79%)	226 (93%)	16 (7%)	16	46
All	All	565/636 (89%)	547 (97%)	18 (3%)	39	67

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	292	LEU
3	C	369	ILE
3	C	366	VAL
3	C	206	SER
3	C	288	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	GLN
2	B	31	HIS
3	C	48	GLN
3	C	276	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 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	278/279 (99%)	0.34	14 (5%)	28	27	86, 154, 221, 259	0
2	B	99/99 (100%)	0.50	7 (7%)	16	16	98, 151, 238, 294	0
3	C	332/373 (89%)	0.12	8 (2%)	59	56	70, 141, 224, 272	0
All	All	709/751 (94%)	0.26	29 (4%)	37	35	70, 148, 228, 294	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	90	GLY	4.1
2	B	23	LEU	4.0
1	A	261	VAL	3.9
1	A	172	LEU	3.7
3	C	139	LEU	3.6

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.