



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 06:51 PM BST

PDB ID : 6VR5
Title : Complex of HLA-A2, a class I MHC, with a p53 peptide
Authors : Wu, D.; Gallagher, D.T.; Pierce, B.G.; Mariuzza, R.A.
Deposited on : 2020-02-06
Resolution : 2.38 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
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)	:	2.13.1

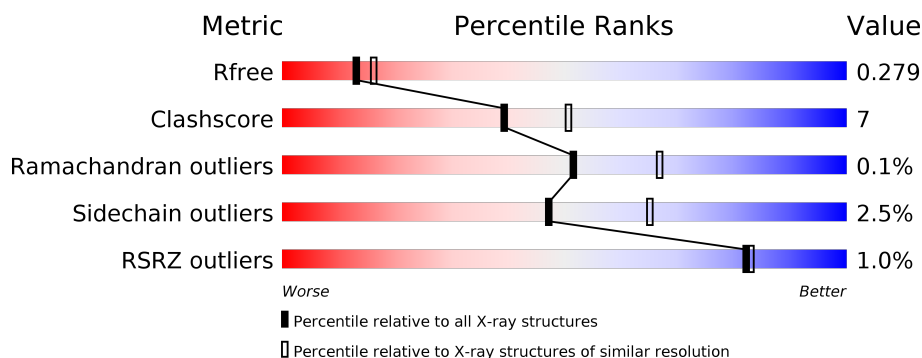
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.38 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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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 $\leq 5\%$. 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	293	<div> <div>81%</div> <div>12%</div> <div>6%</div> </div>
1	D	293	<div>2%</div> <div>77%</div> <div>16%</div> <div>6%</div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6392 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 antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2196	1378	396	413	9			
1	D	275	Total	C	N	O	S	0	0	0
			2206	1385	391	421	9			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q861F7
A	276	GLY	-	expression tag	UNP Q861F7
A	277	GLY	-	expression tag	UNP Q861F7
A	278	GLY	-	expression tag	UNP Q861F7
A	279	LEU	-	expression tag	UNP Q861F7
A	280	ASN	-	expression tag	UNP Q861F7
A	281	ASP	-	expression tag	UNP Q861F7
A	282	ILE	-	expression tag	UNP Q861F7
A	283	PHE	-	expression tag	UNP Q861F7
A	284	GLU	-	expression tag	UNP Q861F7
A	285	ALA	-	expression tag	UNP Q861F7
A	286	GLN	-	expression tag	UNP Q861F7
A	287	LYS	-	expression tag	UNP Q861F7
A	288	ILE	-	expression tag	UNP Q861F7
A	289	GLU	-	expression tag	UNP Q861F7
A	290	TRP	-	expression tag	UNP Q861F7
A	291	HIS	-	expression tag	UNP Q861F7
A	292	GLU	-	expression tag	UNP Q861F7
D	0	MET	-	initiating methionine	UNP Q861F7
D	276	GLY	-	expression tag	UNP Q861F7
D	277	GLY	-	expression tag	UNP Q861F7
D	278	GLY	-	expression tag	UNP Q861F7
D	279	LEU	-	expression tag	UNP Q861F7
D	280	ASN	-	expression tag	UNP Q861F7
D	281	ASP	-	expression tag	UNP Q861F7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	282	ILE	-	expression tag	UNP Q861F7
D	283	PHE	-	expression tag	UNP Q861F7
D	284	GLU	-	expression tag	UNP Q861F7
D	285	ALA	-	expression tag	UNP Q861F7
D	286	GLN	-	expression tag	UNP Q861F7
D	287	LYS	-	expression tag	UNP Q861F7
D	288	ILE	-	expression tag	UNP Q861F7
D	289	GLU	-	expression tag	UNP Q861F7
D	290	TRP	-	expression tag	UNP Q861F7
D	291	HIS	-	expression tag	UNP Q861F7
D	292	GLU	-	expression tag	UNP Q861F7

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	98	Total	C	N	O	S	0	0	0
			786	503	133	147	3			
2	E	100	Total	C	N	O	S	0	0	0
			809	517	137	151	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P61769
E	1	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Cellular tumor antigen p53 peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	9	Total	C	N	O	S	0	0	0
			76	45	16	13	2			
3	Q	9	Total	C	N	O	S	0	0	0
			76	45	16	13	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	8	HIS	ARG	engineered mutation	UNP P04637
Q	8	HIS	ARG	engineered mutation	UNP P04637

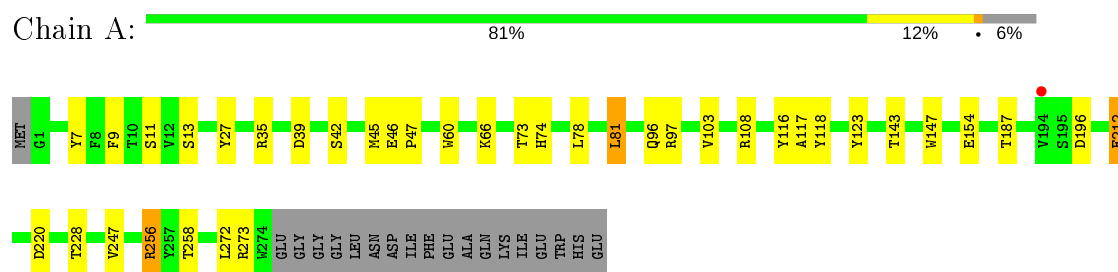
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	80	Total 80	O 80	0	0
4	B	31	Total 31	O 31	0	0
4	D	80	Total 80	O 80	0	0
4	E	46	Total 46	O 46	0	0
4	P	5	Total 5	O 5	0	0
4	Q	1	Total 1	O 1	0	0

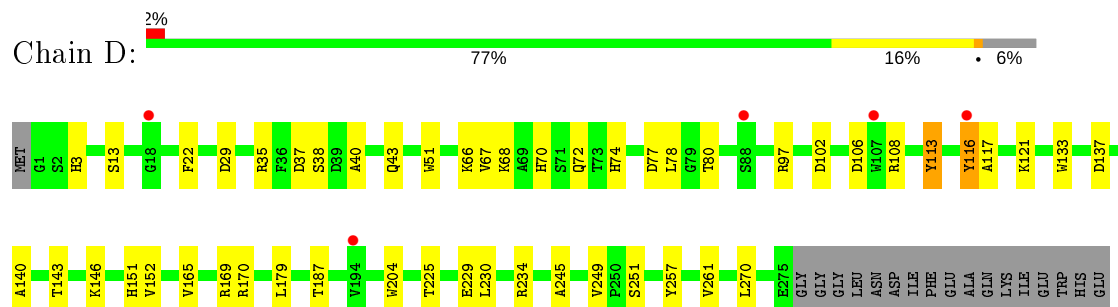
3 Residue-property plots [i](#)

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.

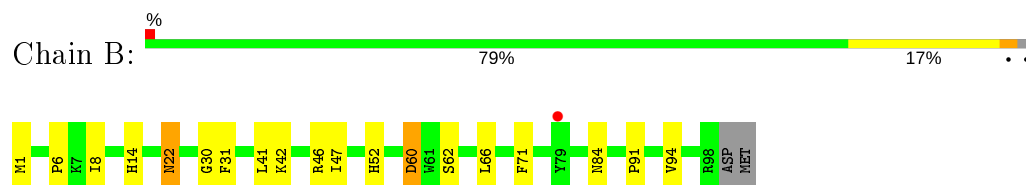
- Molecule 1: MHC class I antigen



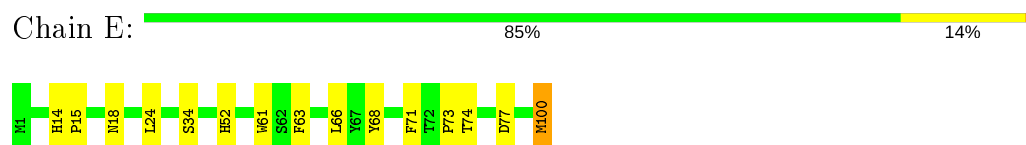
- Molecule 1: MHC class I antigen



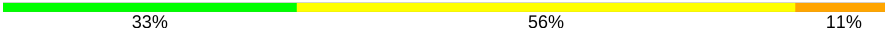
- Molecule 2: Beta-2-microglobulin



- Molecule 2: Beta-2-microglobulin



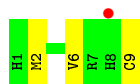
- Molecule 3: Cellular tumor antigen p53 peptide

Chain P:  33% 56% 11%



● Molecule 3: Cellular tumor antigen p53 peptide

Chain Q:  11% 67% 33%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.60Å 79.66Å 85.86Å 90.00° 102.11° 90.00°	Depositor
Resolution (Å)	20.00 – 2.38 48.10 – 2.38	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.38) 100.0 (48.10-2.38)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.94 (at 2.37Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.229 , 0.280 0.230 , 0.279	Depositor DCC
R_{free} test set	1883 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	32.6	Xtriage
Anisotropy	0.638	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 45.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	6392	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 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	0.66	3/2260 (0.1%)	1.01	2/3073 (0.1%)
1	D	0.58	0/2271	0.98	2/3089 (0.1%)
2	B	0.50	0/809	0.90	1/1103 (0.1%)
2	E	0.58	0/832	0.95	0/1131
3	P	0.92	0/77	1.30	0/101
3	Q	0.55	0/77	0.98	0/101
All	All	0.61	3/6326 (0.0%)	0.98	5/8598 (0.1%)

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 (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	212	GLU	CD-OE1	9.67	1.36	1.25
1	A	212	GLU	CD-OE2	7.49	1.33	1.25
1	A	154	GLU	CD-OE1	5.01	1.31	1.25

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	234	ARG	NE-CZ-NH2	-7.09	116.75	120.30
1	A	256	ARG	NE-CZ-NH1	-6.51	117.04	120.30
1	D	234	ARG	NE-CZ-NH1	6.15	123.37	120.30
1	A	256	ARG	NE-CZ-NH2	5.81	123.20	120.30
2	B	60	ASP	CB-CA-C	-5.34	99.71	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	212	GLU	Mainchain
1	A	27	TYR	Peptide

5.2 Too-close contacts

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	2196	0	2034	34	0
1	D	2206	0	2030	32	0
2	B	786	0	721	15	0
2	E	809	0	754	8	0
3	P	76	0	74	14	0
3	Q	76	0	74	6	0
4	A	80	0	0	2	0
4	B	31	0	0	4	0
4	D	80	0	0	2	0
4	E	46	0	0	0	0
4	P	5	0	0	0	0
4	Q	1	0	0	0	0
All	All	6392	0	5687	88	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 7.

The worst 5 of 88 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:D:165:VAL:O	1:D:169:ARG:HG3	1.86	0.73
2:B:6:PRO:HB3	2:B:31:PHE:HB3	1.71	0.73
2:B:46:ARG:NH2	4:B:202:HOH:O	2.16	0.72
1:D:68:LYS:O	1:D:72:GLN:HG2	1.90	0.72
1:A:123:TYR:HH	1:A:143:THR:HG1	1.35	0.72

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	272/293 (93%)	266 (98%)	6 (2%)	0	100	100
1	D	273/293 (93%)	269 (98%)	4 (2%)	0	100	100
2	B	96/100 (96%)	95 (99%)	1 (1%)	0	100	100
2	E	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
3	P	7/9 (78%)	6 (86%)	0	1 (14%)	0	0
3	Q	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	753/804 (94%)	738 (98%)	14 (2%)	1 (0%)	51	67

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	6	VAL

5.3.2 Protein sidechains

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	221/245 (90%)	219 (99%)	2 (1%)	78	89
1	D	223/245 (91%)	215 (96%)	8 (4%)	35	51
2	B	84/95 (88%)	82 (98%)	2 (2%)	49	66
2	E	88/95 (93%)	86 (98%)	2 (2%)	50	68
3	P	9/9 (100%)	7 (78%)	2 (22%)	1	1
3	Q	9/9 (100%)	9 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	634/698 (91%)	618 (98%)	16 (2%)	47 65

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	121	LYS
1	D	151	HIS
2	E	71	PHE
1	D	116	TYR
2	E	100	MET

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

Mol	Chain	Res	Type
1	A	218	GLN
2	E	3	GLN
3	P	8	HIS
3	Q	8	HIS

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 ⓘ

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	274/293 (93%)	0.06	1 (0%) 92 93	4, 29, 46, 66	0
1	D	275/293 (93%)	0.30	5 (1%) 68 70	15, 34, 53, 75	0
2	B	98/100 (98%)	0.25	1 (1%) 82 83	20, 37, 59, 73	0
2	E	100/100 (100%)	0.13	0 100 100	18, 30, 55, 59	0
3	P	9/9 (100%)	0.27	0 100 100	17, 25, 31, 36	0
3	Q	9/9 (100%)	0.95	1 (11%) 5 6	35, 41, 60, 64	0
All	All	765/804 (95%)	0.19	8 (1%) 82 83	4, 32, 55, 75	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	194	VAL	3.7
3	Q	8	HIS	3.5
1	D	18	GLY	3.2
1	D	88	SER	2.8
2	B	79	TYR	2.5

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.