



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 14, 2022 – 02:49 PM EST

PDB ID : 7SRK
Title : Single chain trimer HLA-A*24:02 (Y108C, A163C) with 8mer peptide YP-PVPETF
Authors : Finton, K.A.K.; Rupert, P.B.
Deposited on : 2021-11-08
Resolution : 2.50 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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	:	2.31.2
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)	:	2.31.2

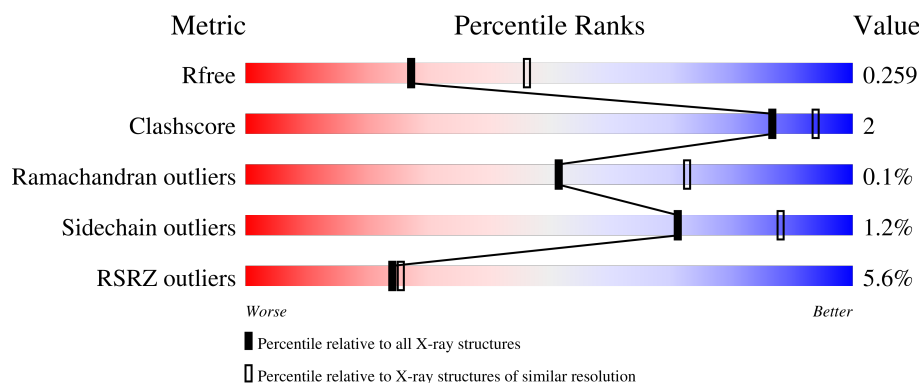
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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	423	<div> <div>4%</div> <div>85%</div> <div>12%</div> </div>
1	C	423	<div> <div>7%</div> <div>85%</div> <div>5%</div> <div>10%</div> </div>
2	B	116	<div> <div>3%</div> <div>91%</div> <div>8%</div> </div>
2	D	116	<div> <div>3%</div> <div>87%</div> <div>11%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7685 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 RPA-related protein RADX peptide,Beta-2-microglobulin,MH C class I antigen chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	0	0
			2952	1860	514	565	13			
1	C	381	Total	C	N	O	S	0	0	0
			2972	1868	522	568	14			

There are 86 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	GLY	-	linker	UNP Q6NSI4
A	11	GLY	-	linker	UNP Q6NSI4
A	12	GLY	-	linker	UNP Q6NSI4
A	13	GLY	-	linker	UNP Q6NSI4
A	14	SER	-	linker	UNP Q6NSI4
A	15	GLY	-	linker	UNP Q6NSI4
A	16	GLY	-	linker	UNP Q6NSI4
A	17	GLY	-	linker	UNP Q6NSI4
A	18	GLY	-	linker	UNP Q6NSI4
A	19	SER	-	linker	UNP Q6NSI4
A	20	GLY	-	linker	UNP Q6NSI4
A	21	GLY	-	linker	UNP Q6NSI4
A	22	GLY	-	linker	UNP Q6NSI4
A	23	GLY	-	linker	UNP Q6NSI4
A	24	SER	-	linker	UNP Q6NSI4
A	124	GLY	-	linker	UNP P16213
A	125	GLY	-	linker	UNP P16213
A	126	GLY	-	linker	UNP P16213
A	127	GLY	-	linker	UNP P16213
A	128	SER	-	linker	UNP P16213
A	129	GLY	-	linker	UNP P16213
A	130	GLY	-	linker	UNP P16213
A	131	GLY	-	linker	UNP P16213
A	132	GLY	-	linker	UNP P16213

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	133	SER	-	linker	UNP P16213
A	134	GLY	-	linker	UNP P16213
A	135	GLY	-	linker	UNP P16213
A	136	GLY	-	linker	UNP P16213
A	137	GLY	-	linker	UNP P16213
A	138	SER	-	linker	UNP P16213
A	139	GLY	-	linker	UNP P16213
A	140	GLY	-	linker	UNP P16213
A	141	GLY	-	linker	UNP P16213
A	142	GLY	-	linker	UNP P16213
A	143	SER	-	linker	UNP P16213
A	227	CYS	TYR	engineered mutation	UNP A0A411J078
A	282	CYS	ALA	engineered mutation	UNP A0A411J078
A	419	HIS	-	expression tag	UNP A0A411J078
A	420	HIS	-	expression tag	UNP A0A411J078
A	421	HIS	-	expression tag	UNP A0A411J078
A	422	HIS	-	expression tag	UNP A0A411J078
A	423	HIS	-	expression tag	UNP A0A411J078
A	424	HIS	-	expression tag	UNP A0A411J078
C	10	GLY	-	linker	UNP Q6NSI4
C	11	GLY	-	linker	UNP Q6NSI4
C	12	GLY	-	linker	UNP Q6NSI4
C	13	GLY	-	linker	UNP Q6NSI4
C	14	SER	-	linker	UNP Q6NSI4
C	15	GLY	-	linker	UNP Q6NSI4
C	16	GLY	-	linker	UNP Q6NSI4
C	17	GLY	-	linker	UNP Q6NSI4
C	18	GLY	-	linker	UNP Q6NSI4
C	19	SER	-	linker	UNP Q6NSI4
C	20	GLY	-	linker	UNP Q6NSI4
C	21	GLY	-	linker	UNP Q6NSI4
C	22	GLY	-	linker	UNP Q6NSI4
C	23	GLY	-	linker	UNP Q6NSI4
C	24	SER	-	linker	UNP Q6NSI4
C	124	GLY	-	linker	UNP P16213
C	125	GLY	-	linker	UNP P16213
C	126	GLY	-	linker	UNP P16213
C	127	GLY	-	linker	UNP P16213
C	128	SER	-	linker	UNP P16213
C	129	GLY	-	linker	UNP P16213
C	130	GLY	-	linker	UNP P16213
C	131	GLY	-	linker	UNP P16213

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	132	GLY	-	linker	UNP P16213
C	133	SER	-	linker	UNP P16213
C	134	GLY	-	linker	UNP P16213
C	135	GLY	-	linker	UNP P16213
C	136	GLY	-	linker	UNP P16213
C	137	GLY	-	linker	UNP P16213
C	138	SER	-	linker	UNP P16213
C	139	GLY	-	linker	UNP P16213
C	140	GLY	-	linker	UNP P16213
C	141	GLY	-	linker	UNP P16213
C	142	GLY	-	linker	UNP P16213
C	143	SER	-	linker	UNP P16213
C	227	CYS	TYR	engineered mutation	UNP A0A411J078
C	282	CYS	ALA	engineered mutation	UNP A0A411J078
C	419	HIS	-	expression tag	UNP A0A411J078
C	420	HIS	-	expression tag	UNP A0A411J078
C	421	HIS	-	expression tag	UNP A0A411J078
C	422	HIS	-	expression tag	UNP A0A411J078
C	423	HIS	-	expression tag	UNP A0A411J078
C	424	HIS	-	expression tag	UNP A0A411J078

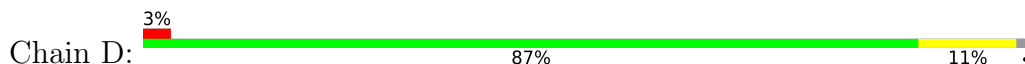
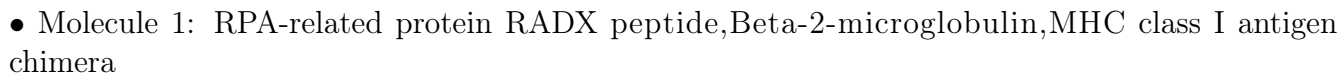
- Molecule 2 is a protein called VHH.

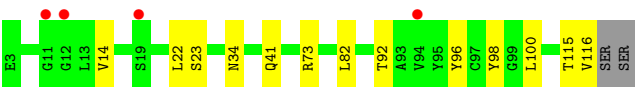
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	115	Total	C	N	O	S	0	0	0
			854	528	148	174	4			
2	D	114	Total	C	N	O	S	0	0	0
			824	507	143	170	4			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	46	Total	O	0	0
			46	46		
3	B	11	Total	O	0	0
			11	11		
3	C	24	Total	O	0	0
			24	24		
3	D	2	Total	O	0	0
			2	2		

- Molecule 1: RPA-related protein RADX peptide, Beta-2-microglobulin, MHC class I antigen chimera





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	69.92Å 76.07Å 226.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.83 – 2.50 29.83 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.83-2.50) 99.9 (29.83-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.33 (at 2.51Å)	Xtriage
Refinement program	PHENIX 1.19.1_4122	Depositor
R, R_{free}	0.229 , 0.260 0.228 , 0.259	Depositor DCC
R_{free} test set	2155 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	39.1	Xtriage
Anisotropy	0.365	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 32.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	7685	wwPDB-VP
Average B, all atoms (Å ²)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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 [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	0.24	0/3033	0.49	0/4126
1	C	0.24	0/3053	0.49	0/4155
2	B	0.24	0/869	0.49	0/1178
2	D	0.24	0/839	0.49	0/1141
All	All	0.24	0/7794	0.49	0/10600

There are no bond length outliers.

There are no bond angle outliers.

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	2952	0	2713	7	0
1	C	2972	0	2690	11	0
2	B	854	0	802	4	0
2	D	824	0	747	8	0
3	A	46	0	0	1	0
3	B	11	0	0	0	0
3	C	24	0	0	0	0
3	D	2	0	0	0	0
All	All	7685	0	6952	30	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 2.

The worst 5 of 30 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:C:163:PRO:O	1:C:218:ARG:NH2	2.28	0.67
1:A:3:PRO:O	3:A:501:HOH:O	2.17	0.58
2:B:34:ASN:O	2:B:73:ARG:NH1	2.35	0.57
2:D:92:THR:HG22	2:D:115:THR:HA	1.85	0.57
1:C:151:PHE:HE1	1:C:241:MET:HE2	1.69	0.57

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	365/423 (86%)	353 (97%)	11 (3%)	1 (0%)	41	61
1	C	375/423 (89%)	364 (97%)	11 (3%)	0	100	100
2	B	113/116 (97%)	109 (96%)	4 (4%)	0	100	100
2	D	112/116 (97%)	107 (96%)	5 (4%)	0	100	100
All	All	965/1078 (90%)	933 (97%)	31 (3%)	1 (0%)	51	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	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	304/347 (88%)	301 (99%)	3 (1%)	76	90
1	C	297/347 (86%)	293 (99%)	4 (1%)	69	87
2	B	90/97 (93%)	89 (99%)	1 (1%)	73	89
2	D	84/97 (87%)	83 (99%)	1 (1%)	71	88
All	All	775/888 (87%)	766 (99%)	9 (1%)	71	88

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	375	GLU
2	D	23	SER
2	B	32	SER
1	C	94	PHE
1	C	201	GLU

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	41	ASN
1	A	284	GLN
1	A	340	HIS
2	D	85	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

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	373/423 (88%)	0.05	17 (4%) 32 34	22, 39, 90, 115	0
1	C	381/423 (90%)	0.39	30 (7%) 12 12	31, 54, 98, 139	0
2	B	115/116 (99%)	0.01	4 (3%) 44 47	27, 36, 66, 124	0
2	D	114/116 (98%)	0.36	4 (3%) 44 47	37, 60, 92, 102	0
All	All	983/1078 (91%)	0.21	55 (5%) 24 25	22, 46, 94, 139	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	12	GLY	6.2
2	B	29	SER	3.7
1	C	373	LEU	3.6
1	A	400	TYR	3.6
1	A	360	TRP	3.5

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.