WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 19, 2023 – 06:29 PM JST

|		
PDB ID	:	8HN4
Title	:	Complex structure of HLA2402 with recognizing SARS-CoV-2 epitope QYIK-WPWYI
Authors	:	Deng, S.S.; Jin, T.C.; Xu, Z.H.
Deposited on	:	2022-12-07
Resolution	:	2.85 Å(reported)

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 2.85 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.
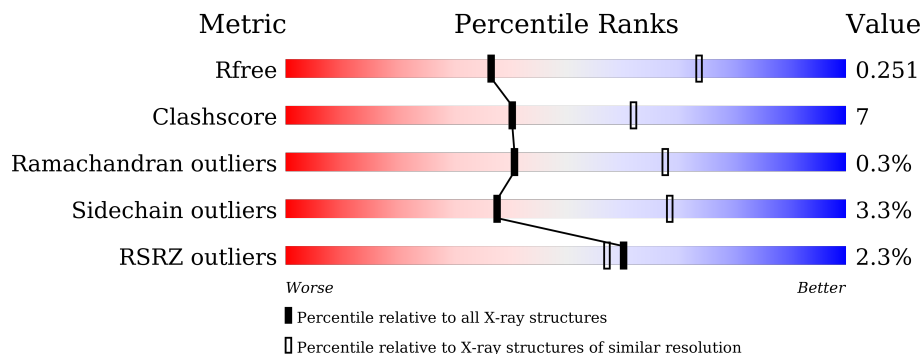
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	308	3% / 74% 15% 11%
2	B	100	% / 88% 12%
3	E	9	44% 33% 22%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 3156 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
			Total	C	N	O	S			
1	A	274	2226	1385	403	428	10	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	initiating methionine	UNP F6IR24
A	21	ASN	-	expression tag	UNP F6IR24
A	22	SER	-	expression tag	UNP F6IR24
A	23	VAL	-	expression tag	UNP F6IR24
A	24	ASP	-	expression tag	UNP F6IR24
A	301	GLY	-	expression tag	UNP F6IR24
A	302	SER	-	expression tag	UNP F6IR24
A	303	GLY	-	expression tag	UNP F6IR24
A	304	LEU	-	expression tag	UNP F6IR24
A	305	ASN	-	expression tag	UNP F6IR24
A	306	ASP	-	expression tag	UNP F6IR24
A	307	ILE	-	expression tag	UNP F6IR24
A	308	PHE	-	expression tag	UNP F6IR24
A	309	GLU	-	expression tag	UNP F6IR24
A	310	ALA	-	expression tag	UNP F6IR24
A	311	GLN	-	expression tag	UNP F6IR24
A	312	LYS	-	expression tag	UNP F6IR24
A	313	ILE	-	expression tag	UNP F6IR24
A	314	GLU	-	expression tag	UNP F6IR24
A	315	TRP	-	expression tag	UNP F6IR24
A	316	HIS	-	expression tag	UNP F6IR24
A	317	ALA	-	expression tag	UNP F6IR24
A	318	ALA	-	expression tag	UNP F6IR24
A	319	ALA	-	expression tag	UNP F6IR24
A	320	LEU	-	expression tag	UNP F6IR24
A	321	GLU	-	expression tag	UNP F6IR24
A	322	HIS	-	expression tag	UNP F6IR24

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	323	HIS	-	expression tag	UNP F6IR24
A	324	HIS	-	expression tag	UNP F6IR24
A	325	HIS	-	expression tag	UNP F6IR24
A	326	HIS	-	expression tag	UNP F6IR24
A	327	HIS	-	expression tag	UNP F6IR24

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	100	837	533	141	159	4	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called SARS-CoV-2 T-cell epitope QYIKWPWYI.

| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	9	93	68	13	12	0	0	0

3 Residue-property plots ⓘ

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
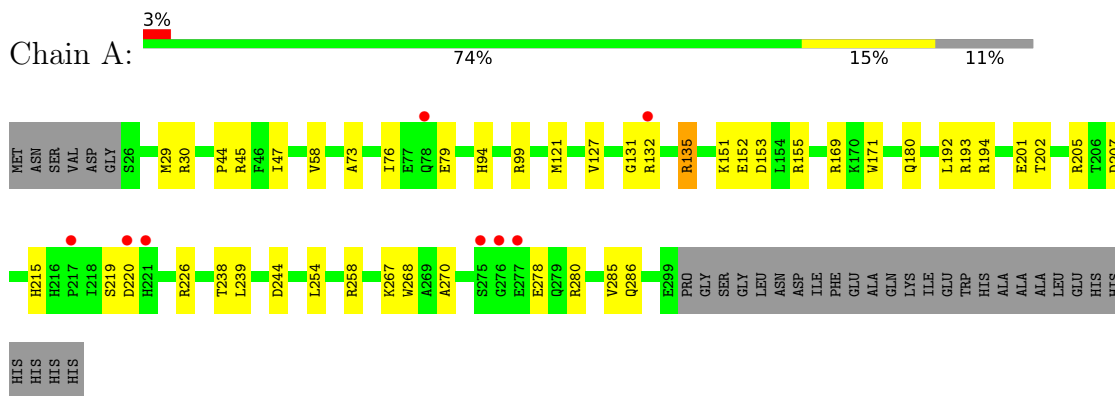
• Molecule 2: Beta-2-microglobulin
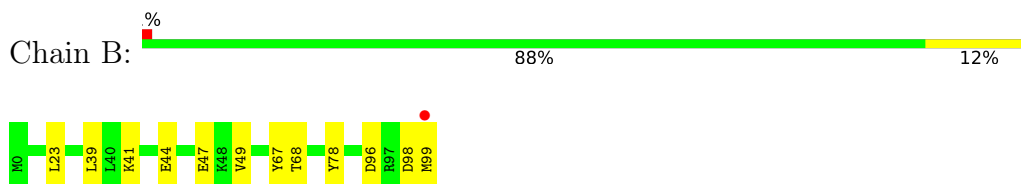
• Molecule 3: SARS-CoV-2 T-cell epitope QYIKWPWYI

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.07Å 67.25Å 116.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.64 – 2.85 39.64 – 2.85	Depositor EDS
% Data completeness (in resolution range)	98.8 (39.64-2.85) 99.5 (39.64-2.85)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 2.86Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.212 , 0.251 0.212 , 0.251	Depositor DCC
R_{free} test set	458 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	39.2	Xtriage
Anisotropy	0.557	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 31.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	3156	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.42% 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 ⓘ

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	0.46	0/2286	0.55	0/3099
2	B	0.60	0/860	0.54	0/1162
3	E	1.39	0/99	1.15	0/136
All	All	0.55	0/3245	0.57	0/4397

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 ⓘ

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	2226	0	2082	33	0
2	B	837	0	803	14	0
3	E	93	0	90	6	0
All	All	3156	0	2975	44	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 44 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:171:TRP:CH2	3:E:5:TRP:HH2	1.77	1.01

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:171:TRP:CH2	3:E:5:TRP:CH2	2.69	0.79
1:A:153:ASP:HB2	1:A:155:ARG:HH11	1.50	0.74
1:A:171:TRP:CZ2	3:E:5:TRP:HH2	2.12	0.67
2:B:23:LEU:HD23	2:B:39:LEU:HD22	1.77	0.67

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	272/308 (88%)	263 (97%)	9 (3%)	0	100	100
2	B	98/100 (98%)	96 (98%)	2 (2%)	0	100	100
3	E	7/9 (78%)	5 (71%)	1 (14%)	1 (14%)	0	0
All	All	377/417 (90%)	364 (97%)	12 (3%)	1 (0%)	41	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	6	PRO

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	231/258 (90%)	225 (97%)	6 (3%)	46	75
2	B	95/95 (100%)	94 (99%)	1 (1%)	73	90
3	E	9/9 (100%)	5 (56%)	4 (44%)	0	0
All	All	335/362 (92%)	324 (97%)	11 (3%)	38	68

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	E	4	LYS
3	E	5	TRP
3	E	9	ILE
3	E	6	PRO
1	A	151	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	274/308 (88%)	0.03	8 (2%)	51	47	14, 31, 51, 77	0
2	B	100/100 (100%)	-0.35	1 (1%)	82	81	15, 25, 44, 49	0
3	E	9/9 (100%)	-0.11	0	100	100	22, 28, 38, 41	0
All	All	383/417 (91%)	-0.07	9 (2%)	60	57	14, 30, 49, 77	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	220	ASP	4.7
1	A	275	SER	3.6
1	A	276	GLY	3.4
1	A	221	HIS	3.4
1	A	277	GLU	2.6

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.