

wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 04:51 AM EDT

PDB ID : 2GTZ

Title: Human Class I MHC HLA-A2 in complex with the nonameric Melan-A/MA

RT-1(27-35) peptide having A28L substitution

Authors: Borbulevych, O.Y.; Baker, B.M.

Deposited on : 2006-04-28

Resolution : 1.70 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:

 $Mol Probity \quad : \quad 4.02b\text{--}467$

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.35

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

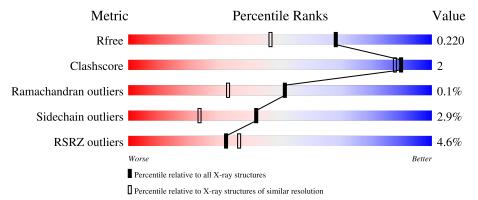
Validation Pipeline (wwPDB-VP) : 2.35

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 1.70 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	275	95%	
1	D		5%	
1		275	92%	7% •
2	В	100	96%	•
2	Е	100	94%	5% •
3	С	9	11%	



Mol	Chain	Length	Quality of chain				
			33%				
3	F	9	67%	33%			



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 7210 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 HLA-A*0201 heavy chain.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	275	Total 2262	C 1414	- 1	O 428	S 10	0	4	0
1	D	275	Total 2286	C 1428	N 416	O 433	S 9	0	9	0

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

Mol	Chain	Residues	\mathbf{Atoms}			ZeroOcc	AltConf	Trace		
2	В	100	Total	С	N	О	S	0	2	0
			839	534	141	160	4	O		0
2	E	100	Total	С	N	О	S	0	9	0
	2 E	100	843	537	141	160	5	U		

There are 2 discrepancies between the modelled and reference sequences:

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

• Molecule 3 is a protein called octapeptide from Melan-A/MART-1.

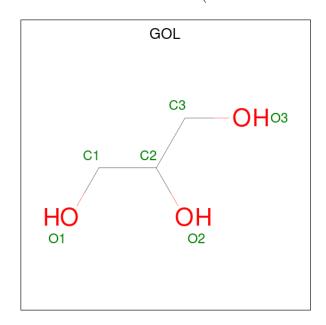
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	С	0	Total	С	N	О	0	2	0
3		9	66	43	11	12			
2	E	0	Total	С	N	О	0	1	0
)	I'	9	64	44	9	11	0	1	

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
С	2	LEU	ALA	engineered mutation	UNP Q16655
F	2	LEU	ALA	engineered mutation	UNP Q16655



 \bullet Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $\mathrm{C_3H_8O_3}).$



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
1	A	1	Total C O	0	0
-	11	1	6 3 3	Ü	
1	A	1	Total C O	0	0
4	Λ	1	6 3 3		
1	A	1	Total C O	0	0
4	Λ	1	6 3 3		
1	A	1	Total C O	0	0
-1	11	1	6 3 3	O	0
1	D	1	Total C O	0	0
-1	D	1	6 3 3	U	0
1	D	1	Total C O	0	0
4		1	6 3 3	U	U

• Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	В	1	Total Na 1 1	0	0
5	E	1	Total Na 1 1	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	268	Total O 268 268	0	0



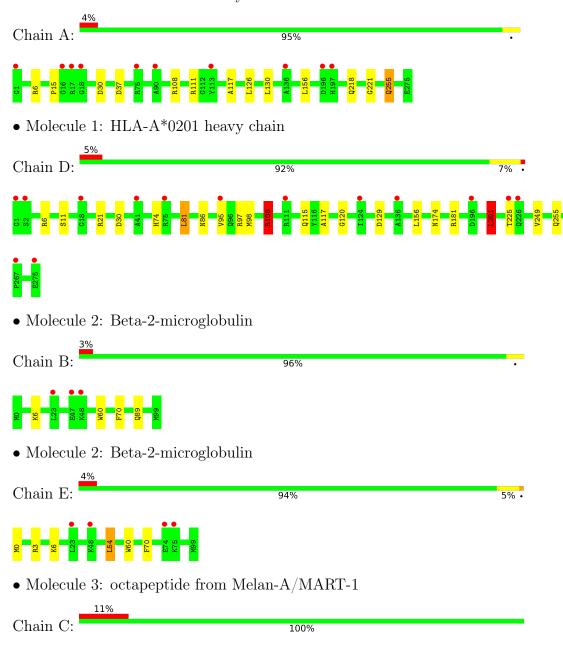
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	В	130	Total O 130 130	0	0
6	С	10	Total O 10 10	0	0
6	D	269	Total O 269 269	0	0
6	E	126	Total O 126 126	0	0
6	F	9	Total O 9 9	0	0



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: HLA-A*0201 heavy chain







 \bullet Molecule 3: octapeptide from Melan-A/MART-1

Chain F: 67% 33%





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	58.40Å 84.31Å 84.14Å	Donositon
a, b, c, α , β , γ	90.00° 90.13° 90.00°	Depositor
Resolution (Å)	20.00 - 1.70	Depositor
rtesolution (A)	19.85 - 1.70	EDS
% Data completeness	97.0 (20.00-1.70)	Depositor
(in resolution range)	97.0 (19.85-1.70)	EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.03 (at 1.69Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, R_{free}	0.172 , 0.217	Depositor
	0.174 , 0.220	DCC
R_{free} test set	4382 reflections $(5.01%)$	wwPDB-VP
Wilson B-factor (Å ²)	15.7	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.38, 43.0	EDS
L-test for twinning ²	$< L > = 0.49, < L^2> = 0.33$	Xtriage
	0.015 for -h,-l,-k	
Estimated twinning fraction	0.006 for -h,l,k	Xtriage
	0.125 for h,-k,-l	
F_o, F_c correlation	0.95	EDS
Total number of atoms	7210	wwPDB-VP
Average B, all atoms (\mathring{A}^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.78% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, GOL

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		
WIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	A	0.77	0/2343	0.84	1/3179 (0.0%)	
1	D	0.74	0/2384	0.89	$6/3235 \ (0.2\%)$	
2	В	0.79	0/866	0.82	0/1170	
2	Е	0.77	0/874	0.85	2/1180 (0.2%)	
3	С	0.81	0/71	1.20	0/94	
3	F	0.72	0/67	0.83	0/89	
All	All	0.76	0/6605	0.86	9/8947 (0.1%)	

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\mathrm{Ideal}(^{o})$
1	D	108	ARG	NE-CZ-NH1	6.16	123.38	120.30
2	Е	54	LEU	CB-CG-CD2	5.84	120.93	111.00
1	D	201	LEU	CB-CG-CD1	5.79	120.85	111.00
1	D	129	ASP	CB-CG-OD1	5.57	123.31	118.30
1	A	37	ASP	CB-CG-OD1	5.55	123.30	118.30

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	2262	0	2113	9	0
1	D	2286	0	2130	12	0
2	В	839	0	800	1	0
2	Е	843	0	807	4	0
3	С	66	0	77	0	0
3	F	64	0	80	1	0
4	A	24	0	32	2	0
4	D	12	0	16	1	0
5	В	1	0	0	0	0
5	Е	1	0	0	0	0
6	A	268	0	0	1	0
6	В	130	0	0	0	0
6	С	10	0	0	0	0
6	D	269	0	0	0	0
6	Ε	126	0	0	1	0
6	F	9	0	0	1	0
All	All	7210	0	6055	23	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 23 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:D:201:LEU:HD22	1:D:249:VAL:HG21	1.84	0.59
1:A:218:GLN:HE21	1:A:221:GLY:HA2	1.71	0.55
1:A:6:ARG:NH2	6:A:1213:HOH:O	2.40	0.55
1:D:74[B]:HIS:HE2	1:D:97:ARG:HE	1.58	0.51
1:D:11[A]:SER:OG	1:D:95[A]:VAL:CG2	2.60	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	Perce	ntiles
1	A	277/275 (101%)	270 (98%)	6 (2%)	1 (0%)	34	18
1	D	282/275 (102%)	278 (99%)	4 (1%)	0	100	100
2	В	99/100 (99%)	98 (99%)	1 (1%)	0	100	100
2	E	99/100 (99%)	98 (99%)	1 (1%)	0	100	100
3	С	9/9 (100%)	7 (78%)	2 (22%)	0	100	100
3	F	8/9 (89%)	8 (100%)	0	0	100	100
All	All	774/768 (101%)	759 (98%)	14 (2%)	1 (0%)	51	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	15	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	Perce	ntiles
1	A	$235/231 \; (102\%)$	232 (99%)	3 (1%)	69	56
1	D	240/231 (104%)	231 (96%)	9 (4%)	33	14
2	В	96/95 (101%)	93 (97%)	3 (3%)	40	21
2	Е	97/95 (102%)	93 (96%)	4 (4%)	30	12
3	С	7/6 (117%)	7 (100%)	0	100	100
3	F	7/6 (117%)	4 (57%)	3 (43%)	0	0
All	All	682/664 (103%)	660 (97%)	22 (3%)	42	20

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	255	GLN
2	Е	54	LEU
2	Е	0[B]	MET
2	Е	70	PHE



Mol	Chain	Res	Type
1	D	81	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	218	GLN
1	A	255	GLN
1	D	255	GLN
1	D	115	GLN
1	A	155	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)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Mol Type Chain Res Lin		Link	Bond lengths			Bond angles			
WIOI	Type	Chain	nes	LIIIK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	GOL	D	1004	-	5,5,5	0.26	0	5,5,5	0.87	0
4	GOL	A	1002	-	5,5,5	0.42	0	5,5,5	1.19	1 (20%)



Mol	Type	Chain	Chain Res	Link	Bond lengths			Bond angles		
MIOI				5 LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
4	GOL	D	1003	-	5,5,5	0.30	0	5,5,5	1.24	1 (20%)
4	GOL	A	1006	-	5,5,5	0.37	0	5,5,5	0.49	0
4	GOL	A	1005	-	5,5,5	0.46	0	5,5,5	0.91	0
4	GOL	A	1001	-	5,5,5	0.40	0	5,5,5	0.44	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	D	1004	-	-	0/4/4/4	-
4	GOL	A	1002	-	-	2/4/4/4	-
4	GOL	D	1003	-	-	3/4/4/4	-
4	GOL	A	1006	-	-	4/4/4/4	-
4	GOL	A	1005	-	-	4/4/4/4	-
4	GOL	A	1001	-	-	0/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

	Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\mathbf{Observed}(^o)$	$\operatorname{Ideal}({}^{o})$
ſ	4	D	1003	GOL	C3-C2-C1	-2.49	102.02	111.70
	4	A	1002	GOL	C3-C2-C1	-2.30	102.77	111.70

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1002	GOL	O1-C1-C2-C3
4	A	1005	GOL	C1-C2-C3-O3
4	A	1006	GOL	O1-C1-C2-C3
4	A	1006	GOL	C1-C2-C3-O3
4	A	1006	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1002	GOL	2	0



\mathbf{Mol}	Chain	Res	Type	Clashes	Symm-Clashes
4	D	1003	GOL	1	0

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, 95^{th} 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>	$\# \mathrm{RSRZ}{>}2$	$OWAB(A^2)$	Q<0.9
1	A	275/275 (100%)	0.40	10 (3%) 42 47	14, 19, 27, 37	0
1	D	275/275 (100%)	0.42	14 (5%) 28 31	14, 19, 27, 32	0
2	В	100/100 (100%)	0.36	3 (3%) 50 54	16, 19, 28, 33	0
2	Е	100/100 (100%)	0.35	4 (4%) 38 42	15, 19, 27, 31	0
3	С	9/9 (100%)	1.07	1 (11%) 5 6	18, 21, 34, 35	0
3	F	9/9 (100%)	1.64	3 (33%) 0 0	19, 23, 34, 36	0
All	All	768/768 (100%)	0.42	35 (4%) 32 36	14, 19, 28, 37	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	5	GLY	5.0
1	A	1	GLY	4.9
2	В	48	LYS	4.8
1	A	196	ASP	4.6
1	D	1	GLY	3.8

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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
4	GOL	A	1006	6/6	0.67	0.17	37,38,38,39	6
4	GOL	A	1005	6/6	0.72	0.24	27,33,35,37	6
4	GOL	D	1003	6/6	0.79	0.17	33,36,37,37	0
4	GOL	A	1002	6/6	0.81	0.20	35,37,38,38	0
4	GOL	D	1004	6/6	0.91	0.18	27,28,31,31	0
5	NA	Ε	5002	1/1	0.93	0.27	40,40,40,40	0
4	GOL	A	1001	6/6	0.96	0.17	20,22,24,24	0
5	NA	В	5001	1/1	0.97	0.21	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.

