



wwPDB X-ray Structure Validation Summary Report

Aug 26, 2024 – 12:16 PM EDT

PDB ID : 9C6X
Title : Crystal Structure of a single chain trimer composed of HLA-B*39:01 Y84C variant, beta-2microglobulin, and NRVMLPKAA peptide from NLRP2
Authors : Sharma, R.; Amdare, N.P.; Celikgil, A.; Garforth, S.J.; DiLorenzo, T.P.; Almo, S.C.; Ghosh, A.
Deposited on : 2024-06-09
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  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
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.20.1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.3

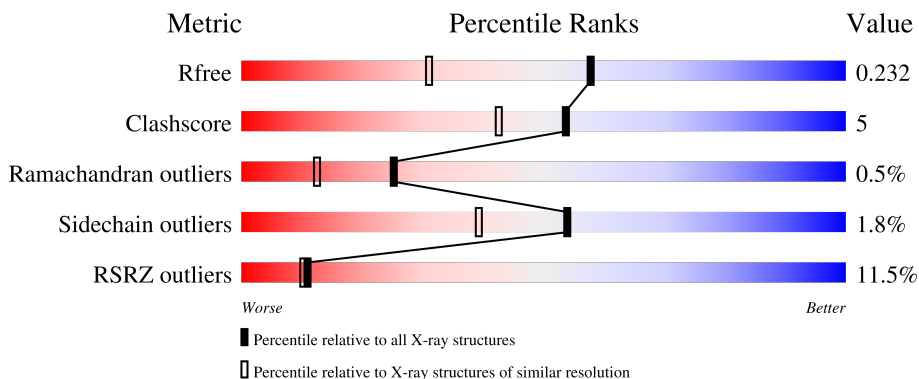
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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 (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (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 $\leq 5\%$. 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	418	

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 3472 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 NACHT, LRR and PYD domains-containing protein 2, Beta-2-microglobulin, MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	390	3206	2003	581	609	13	0	6	0

There are 37 discrepancies between the modelled and reference sequences:

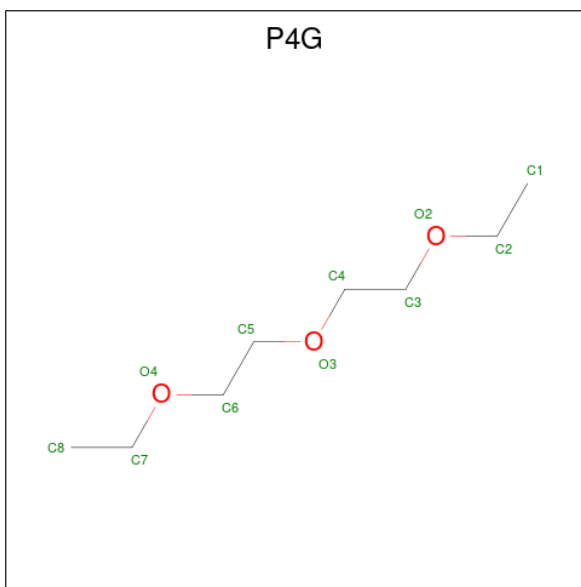
Chain	Residue	Modelled	Actual	Comment	Reference
A	332	GLY	-	linker	UNP Q9NX02
A	333	CYS	-	linker	UNP Q9NX02
A	334	GLY	-	linker	UNP Q9NX02
A	335	ALA	-	linker	UNP Q9NX02
A	990	SER	-	linker	UNP Q9NX02
A	991	GLY	-	linker	UNP Q9NX02
A	992	GLY	-	linker	UNP Q9NX02
A	993	GLY	-	linker	UNP Q9NX02
A	994	GLY	-	linker	UNP Q9NX02
A	995	SER	-	linker	UNP Q9NX02
A	996	GLY	-	linker	UNP Q9NX02
A	997	GLY	-	linker	UNP Q9NX02
A	998	GLY	-	linker	UNP Q9NX02
A	999	GLY	-	linker	UNP Q9NX02
A	1000	SER	-	linker	UNP Q9NX02
A	1001	MET	-	linker	UNP Q9NX02
A	1981	GLY	-	linker	UNP P61769
A	1982	GLY	-	linker	UNP P61769
A	1983	GLY	-	linker	UNP P61769
A	1984	GLY	-	linker	UNP P61769
A	1985	SER	-	linker	UNP P61769
A	1986	GLY	-	linker	UNP P61769
A	1987	GLY	-	linker	UNP P61769
A	1988	GLY	-	linker	UNP P61769
A	1989	GLY	-	linker	UNP P61769
A	1990	SER	-	linker	UNP P61769

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1991	GLY	-	linker	UNP P61769
A	1992	GLY	-	linker	UNP P61769
A	1993	GLY	-	linker	UNP P61769
A	1994	GLY	-	linker	UNP P61769
A	1995	SER	-	linker	UNP P61769
A	1996	GLY	-	linker	UNP P61769
A	1997	GLY	-	linker	UNP P61769
A	1998	GLY	-	linker	UNP P61769
A	1999	GLY	-	linker	UNP P61769
A	2000	SER	-	linker	UNP P61769
A	2084	CYS	TYR	variant	UNP R5AK10

- Molecule 2 is 1-ETHOXY-2-(2-ETHOXYETHOXY)ETHANE (three-letter code: P4G) (formula: C₈H₁₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			11	8	3		

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0

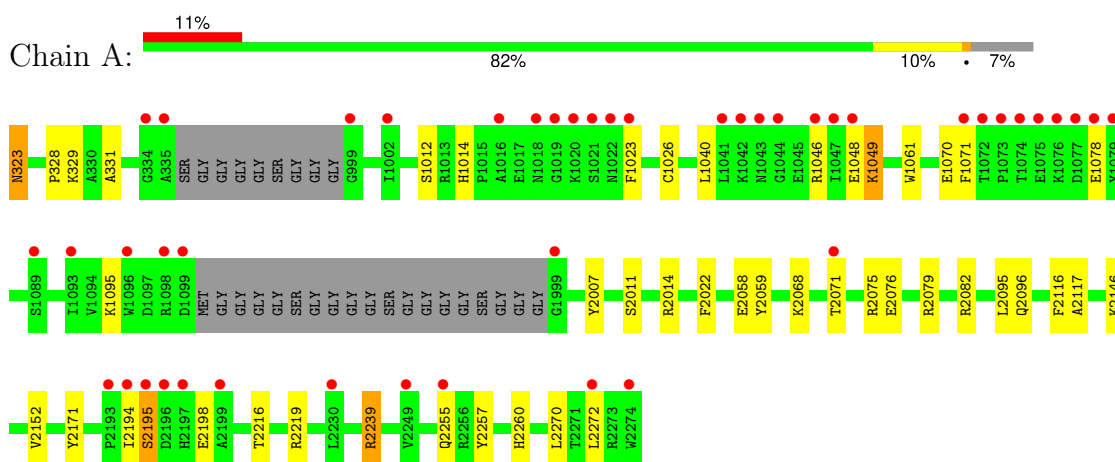
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	250	Total O 251 251	0	1

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: NACHT, LRR and PYD domains-containing protein 2, Beta-2-microglobulin, MHC class I antigen



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	75.94Å 84.12Å 150.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.83 – 1.70 24.83 – 1.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.83-1.70) 99.9 (24.83-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.88 (at 1.70Å)	Xtrriage
Refinement program	PHENIX (1.21rc1_5156)	Depositor
R, R_{free}	0.202 , 0.232 0.202 , 0.232	Depositor DCC
R_{free} test set	2733 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å ²)	22.3	Xtrriage
Anisotropy	0.293	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3472	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.77% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: P4G, EDO

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	0.38	0/3308	0.62	1/4484 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2272	LEU	CA-CB-CG	5.34	127.57	115.30

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	3206	0	3049	29	0
2	A	11	0	18	2	0
3	A	4	0	6	0	0
4	A	251	0	0	4	0
All	All	3472	0	3073	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 5.

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:A:329:LYS:HD3	1:A:2076:GLU:HG2	1.68	0.75
1:A:2270:LEU:HA	2:A:2301:P4G:H21	1.81	0.63
1:A:1048:GLU:N	1:A:1048:GLU:OE1	2.33	0.62
1:A:2014:ARG:CZ	4:A:2414:HOH:O	2.49	0.58
1:A:323:ASN:N	1:A:2007:TYR:HH	2.02	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	390/418 (93%)	381 (98%)	7 (2%)	2 (0%)	25 12

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2239	ARG
1	A	2195	SER

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	345/345 (100%)	339 (98%)	6 (2%)	56 41

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2075	ARG
1	A	2116	PHE
1	A	2239	ARG
1	A	1049	LYS
1	A	323	ASN

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

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

2 ligands are modelled in this entry.

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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	P4G	A	2301	-	10,10,10	0.36	0	9,9,9	0.15	0
3	EDO	A	2302	-	3,3,3	0.23	0	2,2,2	0.40	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
2	P4G	A	2301	-	-	7/8/8/8	-
3	EDO	A	2302	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2302	EDO	O1-C1-C2-O2
2	A	2301	P4G	O3-C5-C6-O4
2	A	2301	P4G	C1-C2-O2-C3
2	A	2301	P4G	C8-C7-O4-C6
2	A	2301	P4G	C3-C4-O3-C5

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2301	P4G	2	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, 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	390/418 (93%)	0.37	45 (11%) 11 10	11, 28, 70, 92	6 (1%)

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	999	GLY	4.8
1	A	2197	HIS	4.4
1	A	2195	SER	3.9
1	A	1047	ILE	3.8
1	A	1074	THR	3.7

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, 95th 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	B-factors(Å ²)	Q<0.9
2	P4G	A	2301	11/11	0.72	0.19	45,47,51,55	2
3	EDO	A	2302	4/4	0.88	0.10	31,38,38,43	0

6.5 Other polymers [i](#)

There are no such residues in this entry.