



wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 02:18 am BST

PDB ID : 3NWM
Title : Crystal structure of a single chain construct composed of MHC class I H-2Kd, beta-2microglobulin and a peptide which is an autoantigen for type 1 diabetes
Authors : Ramagopal, U.A.; Samanta, D.; Nathenson, S.G.; Almo, S.C.
Deposited on : 2010-07-09
Resolution : 2.70 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

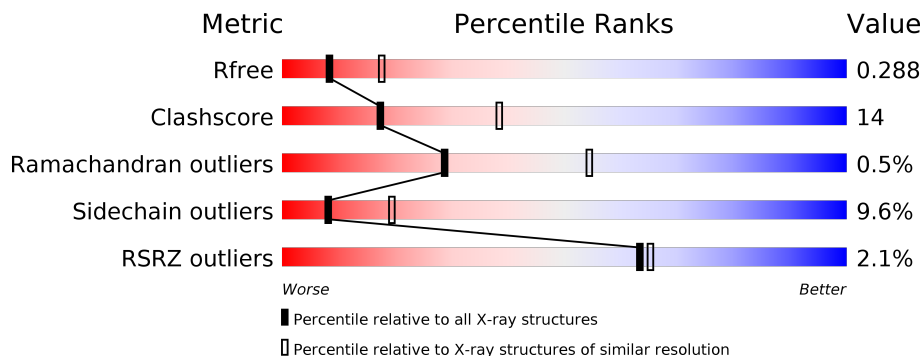
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 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	443	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3174 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 Peptide/beta-2microglobulin/MHC class I H-2Kd chimeric protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	384	3160	2014	552	580	14	0	1	0

There are 62 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0P	MET	-	EXPRESSION TAG	PDB 3NWM
A	10L	GLY	-	LINKER	PDB 3NWM
A	11L	GLY	-	LINKER	PDB 3NWM
A	12L	GLY	-	LINKER	PDB 3NWM
A	13L	ALA	-	LINKER	PDB 3NWM
A	14L	SER	-	LINKER	PDB 3NWM
A	15L	GLY	-	LINKER	PDB 3NWM
A	16L	GLY	-	LINKER	PDB 3NWM
A	17L	GLY	-	LINKER	PDB 3NWM
A	18L	GLY	-	LINKER	PDB 3NWM
A	19L	SER	-	LINKER	PDB 3NWM
A	20L	GLY	-	LINKER	PDB 3NWM
A	21L	GLY	-	LINKER	PDB 3NWM
A	22L	GLY	-	LINKER	PDB 3NWM
A	23L	GLY	-	LINKER	PDB 3NWM
A	-1L	SER	-	LINKER	PDB 3NWM
A	0B	MET	-	EXPRESSION TAG	UNP Q91XJ8
A	100L	GLY	-	LINKER	UNP Q91XJ8
A	101L	GLY	-	LINKER	UNP Q91XJ8
A	102L	GLY	-	LINKER	UNP Q91XJ8
A	103L	GLY	-	LINKER	UNP Q91XJ8
A	104L	SER	-	LINKER	UNP Q91XJ8
A	105L	GLY	-	LINKER	UNP Q91XJ8
A	106L	GLY	-	LINKER	UNP Q91XJ8
A	107L	GLY	-	LINKER	UNP Q91XJ8
A	108L	GLY	-	LINKER	UNP Q91XJ8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	109L	SER	-	LINKER	UNP Q91XJ8
A	110L	GLY	-	LINKER	UNP Q91XJ8
A	111L	GLY	-	LINKER	UNP Q91XJ8
A	112L	GLY	-	LINKER	UNP Q91XJ8
A	113L	GLY	-	LINKER	UNP Q91XJ8
A	114L	SER	-	LINKER	UNP Q91XJ8
A	115L	GLY	-	LINKER	UNP Q91XJ8
A	116L	GLY	-	LINKER	UNP Q91XJ8
A	117L	GLY	-	LINKER	UNP Q91XJ8
A	118L	GLY	-	LINKER	UNP Q91XJ8
A	119L	SER	-	LINKER	UNP Q91XJ8
A	1H	GLY	-	LINKER	UNP Q91XJ8
A	275H	ALA	-	EXPRESSION TAG	UNP Q5KTQ2
A	276H	ALA	-	EXPRESSION TAG	UNP Q5KTQ2
A	277H	ALA	-	EXPRESSION TAG	UNP Q5KTQ2
A	278H	GLY	-	EXPRESSION TAG	UNP Q5KTQ2
A	279H	GLY	-	EXPRESSION TAG	UNP Q5KTQ2
A	280H	THR	-	EXPRESSION TAG	UNP Q5KTQ2
A	281H	GLY	-	EXPRESSION TAG	UNP Q5KTQ2
A	282H	GLY	-	EXPRESSION TAG	UNP Q5KTQ2
A	283H	LEU	-	EXPRESSION TAG	UNP Q5KTQ2
A	284H	ASN	-	EXPRESSION TAG	UNP Q5KTQ2
A	285H	ASP	-	EXPRESSION TAG	UNP Q5KTQ2
A	286H	ILE	-	EXPRESSION TAG	UNP Q5KTQ2
A	287H	PHE	-	EXPRESSION TAG	UNP Q5KTQ2
A	288H	GLU	-	EXPRESSION TAG	UNP Q5KTQ2
A	289H	ALA	-	EXPRESSION TAG	UNP Q5KTQ2
A	290H	GLN	-	EXPRESSION TAG	UNP Q5KTQ2
A	291H	LYS	-	EXPRESSION TAG	UNP Q5KTQ2
A	292H	ILE	-	EXPRESSION TAG	UNP Q5KTQ2
A	293H	GLU	-	EXPRESSION TAG	UNP Q5KTQ2
A	294H	TRP	-	EXPRESSION TAG	UNP Q5KTQ2
A	295H	HIS	-	EXPRESSION TAG	UNP Q5KTQ2
A	296H	GLU	-	EXPRESSION TAG	UNP Q5KTQ2
A	297H	GLY	-	EXPRESSION TAG	UNP Q5KTQ2
A	298H	LEU	-	EXPRESSION TAG	UNP Q5KTQ2

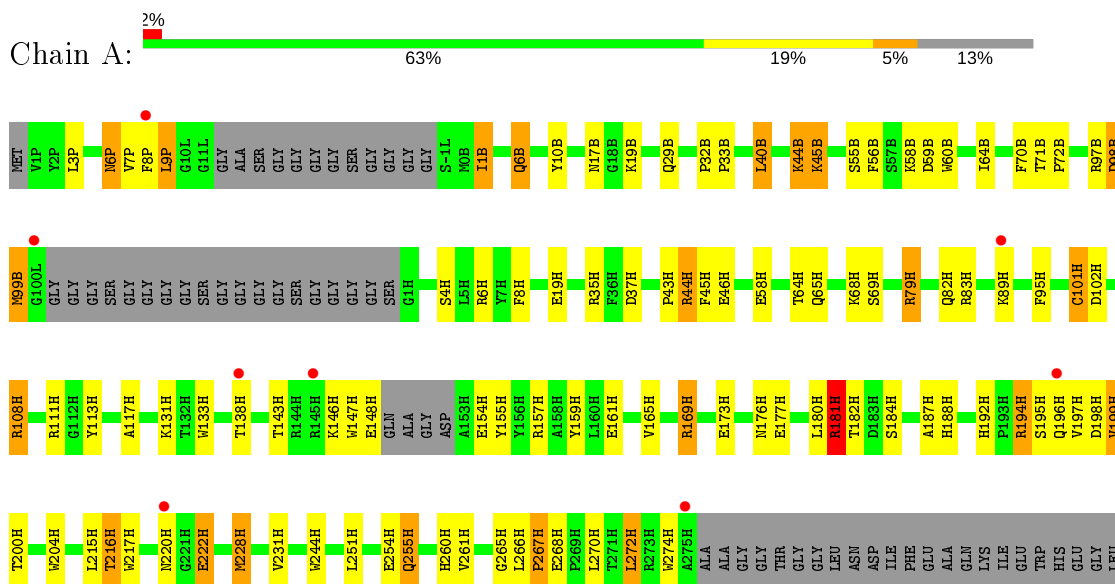
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	14	Total O 14 14	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Peptide/beta-2microglobulin/MHC class I H-2Kd chimeric protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	41.35Å 88.53Å 61.20Å 90.00° 102.92° 90.00°	Depositor
Resolution (Å)	59.65 – 2.70 49.47 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.5 (59.65-2.70) 99.5 (49.47-2.70)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.40 (at 2.69Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.202 , 0.284 0.203 , 0.288	Depositor DCC
R_{free} test set	564 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å ²)	44.6	Xtrriage
Anisotropy	0.158	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 30.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3174	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.22% 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	0.45	1/3258 (0.0%)	0.58	0/4427

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	101(H)	CYS	CB-SG	-6.11	1.71	1.82

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	3160	0	3008	85	0
2	A	14	0	0	1	0
All	All	3174	0	3008	85	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 14.

The worst 5 of 85 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:217(H):TRP:H	1:A:228(H):MET:HE1	1.09	1.10

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:194(H):ARG:HG2	1:A:195(H):SER:N	1.70	1.04
1:A:217(H):TRP:N	1:A:228(H):MET:HE1	1.80	0.97
1:A:182(H):THR:HG23	1:A:182(H):THR:O	1.67	0.93
1:A:165(H):VAL:O	1:A:169(H):ARG:HG3	1.73	0.89

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	377/443 (85%)	339 (90%)	36 (10%)	2 (0%)	29 54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	181(H)	ARG
1	A	267(H)	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	333/359 (93%)	301 (90%)	32 (10%)	8 19

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45(H)	PHE
1	A	69(H)	SER
1	A	255(H)	GLN
1	A	65(H)	GLN
1	A	79(H)	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	65(H)	GLN
1	A	255(H)	GLN
1	A	188(H)	HIS
1	A	67(B)	HIS
1	A	87(H)	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 carbohydrates 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	384/443 (86%)	-0.11	8 (2%) 63 65	24, 41, 69, 77	6 (1%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	8(P)	PHE	3.4
1	A	275(H)	ALA	3.2
1	A	138(H)	THR	2.7
1	A	89(H)	LYS	2.4
1	A	100(L)	GLY	2.4

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 carbohydrates 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.