



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 28, 2022 – 01:10 pm BST

PDB ID : 7NMY  
Title : Crystal structure of beta-2-microglobulin D76Y mutant  
Authors : Guthertz, N.  
Deposited on : 2021-02-23  
Resolution : 1.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 : 2.29  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0267  
CCP4 : 7.1.010 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.29

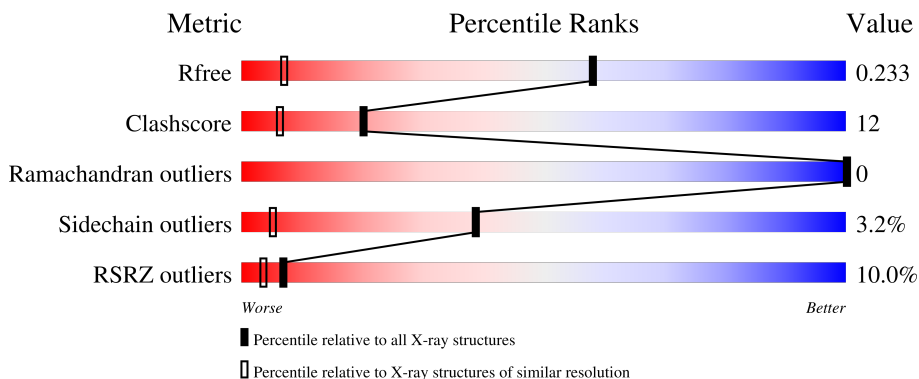
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.25 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	1023 (1.28-1.24)
Clashscore	141614	1060 (1.28-1.24)
Ramachandran outliers	138981	1029 (1.28-1.24)
Sidechain outliers	138945	1028 (1.28-1.24)
RSRZ outliers	127900	1004 (1.28-1.24)

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  $\geq 3$ , 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	AAA	100	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1800 atoms, of which 832 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 Beta-2-microglobulin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	AAA	100	1685	547	832	144	158	4	45	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP P61769
AAA	76	TYR	ASP	engineered mutation	UNP P61769

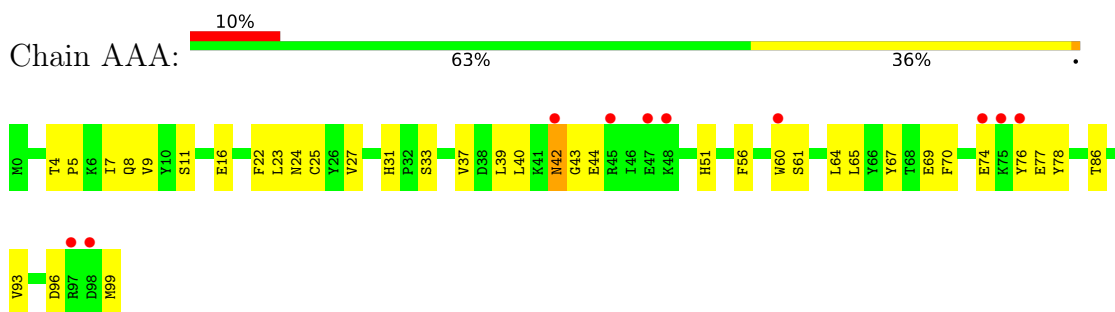
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	AAA	115	Total	O	0	0
			115	115		

### 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: Beta-2-microglobulin



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.71Å 27.69Å 64.52Å 90.00° 101.24° 90.00°	Depositor
Resolution (Å)	43.98 – 1.25 36.27 – 1.25	Depositor EDS
% Data completeness (in resolution range)	97.8 (43.98-1.25) 97.8 (36.27-1.25)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 1.25Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.178 , 0.227 0.186 , 0.233	Depositor DCC
$R_{free}$ test set	1211 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	12.0	Xtriage
Anisotropy	0.473	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	1800	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.77% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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	AAA	0.99	1/883 (0.1%)	1.19	0/1193

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	33	SER	CA-CB	-5.57	1.44	1.52

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	AAA	853	832	830	21	2
2	AAA	115	0	0	1	2
All	All	968	832	830	21	3

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 12.

All (21) 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:AAA:31:HIS:CD2	1:AAA:60:TRP:HA	2.30	0.67

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:42:ASN:ND2	1:AAA:77:GLU:H	1.99	0.61
1:AAA:42:ASN:HD22	1:AAA:42:ASN:N	2.02	0.58
1:AAA:11:SER:HA	1:AAA:22:PHE:O	2.05	0.56
1:AAA:25:CYS:HB2	1:AAA:39:LEU:HD21	1.92	0.52
1:AAA:42:ASN:HD22	1:AAA:42:ASN:H	1.59	0.51
1:AAA:42:ASN:HD21	1:AAA:76:TYR:HA	1.76	0.51
1:AAA:56:PHE:HA	1:AAA:61:SER:O	2.14	0.47
1:AAA:7:ILE:HB	1:AAA:93:VAL:HG21	1.97	0.46
1:AAA:40:LEU:O	1:AAA:78:TYR:HA	2.16	0.45
1:AAA:23:LEU:O	1:AAA:67:TYR:HA	2.18	0.44
1:AAA:40:LEU:HA	1:AAA:44:GLU:O	2.18	0.43
1:AAA:5:PRO:HB2	1:AAA:27:VAL:HG13	2.01	0.43
1:AAA:42:ASN:HD21	1:AAA:77:GLU:H	1.65	0.42
1:AAA:51:HIS:HB3	1:AAA:67:TYR:CE1	2.55	0.41
1:AAA:4:THR:HA	1:AAA:86:THR:OG1	2.20	0.41
1:AAA:65:LEU:HG	1:AAA:67:TYR:CD1	2.55	0.41
1:AAA:9:VAL:HA	1:AAA:24:ASN:O	2.20	0.41
1:AAA:8:GLN:NE2	2:AAA:109:HOH:O	2.51	0.40
1:AAA:37:VAL:HG21	1:AAA:64:LEU:HD23	2.02	0.40
1:AAA:96:ASP:HB3	1:AAA:99:MET:HG3	2.04	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:AAA:209:HOH:O	2:AAA:209:HOH:O[2_555]	1.80	0.40
1:AAA:69:GLU:OE2	1:AAA:74:GLU:OE1[4_455]	1.89	0.31
1:AAA:43:GLY:O	2:AAA:183:HOH:O[2_455]	2.05	0.15

## 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	AAA	100/100 (100%)	97 (97%)	3 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	AAA	97/95 (102%)	94 (97%)	3 (3%)	40	5

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	16	GLU
1	AAA	42	ASN
1	AAA	70	PHE

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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	100/100 (100%)	0.69	10 (10%) <b>7</b> <b>3</b>	9, 16, 37, 54	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	60	TRP	7.6
1	AAA	97	ARG	5.4
1	AAA	74	GLU	4.8
1	AAA	75	LYS	4.3
1	AAA	48	LYS	2.8
1	AAA	76	TYR	2.7
1	AAA	42	ASN	2.7
1	AAA	98	ASP	2.5
1	AAA	45	ARG	2.4
1	AAA	47	GLU	2.0

### 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.