



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 01:30 AM UTC

PDB ID : 6PJ5 / pdb\_00006pj5  
Title : Time-resolved structural snapshot of proteolysis by GlpG inside the membrane  
Authors : Urban, S.; Cho, S.  
Deposited on : 2019-06-27  
Resolution : 2.40 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 : 4-5-2 with Phenix2.0  
Xtrriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.49

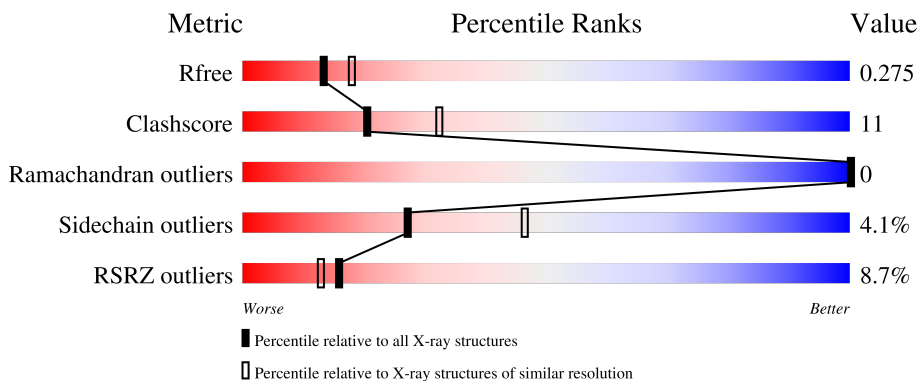
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



The reported resolution of this entry is 2.40 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))
$R_{free}$	180053	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	A	211	 8% 63% 20% .. 15%
2	B	4	 50% 50%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1493 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 Rhomboid protease GlpG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	179	1437	968	232	227	10	0	1	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	66	MET	-	initiating methionine	UNP A0A0J2E248
A	67	GLY	-	expression tag	UNP A0A0J2E248
A	68	SER	-	expression tag	UNP A0A0J2E248
A	69	SER	-	expression tag	UNP A0A0J2E248
A	70	HIS	-	expression tag	UNP A0A0J2E248
A	71	HIS	-	expression tag	UNP A0A0J2E248
A	72	HIS	-	expression tag	UNP A0A0J2E248
A	73	HIS	-	expression tag	UNP A0A0J2E248
A	74	HIS	-	expression tag	UNP A0A0J2E248
A	75	HIS	-	expression tag	UNP A0A0J2E248
A	76	SER	-	expression tag	UNP A0A0J2E248
A	77	SER	-	expression tag	UNP A0A0J2E248
A	78	GLY	-	expression tag	UNP A0A0J2E248
A	79	LEU	-	expression tag	UNP A0A0J2E248
A	80	VAL	-	expression tag	UNP A0A0J2E248
A	81	PRO	-	expression tag	UNP A0A0J2E248
A	82	ARG	-	expression tag	UNP A0A0J2E248
A	83	GLY	-	expression tag	UNP A0A0J2E248
A	84	SER	-	expression tag	UNP A0A0J2E248
A	85	HIS	-	expression tag	UNP A0A0J2E248
A	86	MET	-	expression tag	UNP A0A0J2E248
A	205	PHE	TYR	engineered mutation	UNP A0A0J2E248

- Molecule 2 is a protein called Peptide aldehyde inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	4	31	19	7	4	1	0	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	25	Total	O	0	0
			25	25		



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.77Å 98.44Å 62.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.79 – 2.40 29.79 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.4 (29.79-2.40) 96.4 (29.79-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.86 (at 2.31Å)	Xtrriage
Refinement program	REFMAC 5.8.0230	Depositor
R, $R_{free}$	0.223 , 0.268 0.234 , 0.275	Depositor DCC
$R_{free}$ test set	487 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	29.6	Xtrriage
Anisotropy	0.905	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 59.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	1493	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.99% 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	A	0.68	1/1485 (0.1%)	1.30	12/2020 (0.6%)
2	B	0.58	0/30	1.32	0/38
All	All	0.68	1/1515 (0.1%)	1.30	12/2058 (0.6%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	96	VAL	C-O	-5.07	1.18	1.24

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	253	ALA	CA-C-N	10.01	134.05	120.54
1	A	253	ALA	C-N-CA	10.01	134.05	120.54
1	A	255	ILE	CA-C-N	8.03	131.36	120.44
1	A	255	ILE	C-N-CA	8.03	131.36	120.44
1	A	256	ALA	O-C-N	6.72	129.35	122.09
1	A	96	VAL	O-C-N	-6.46	115.18	121.90
1	A	91	GLU	CA-C-N	6.43	131.47	121.76
1	A	91	GLU	C-N-CA	6.43	131.47	121.76
1	A	92	ARG	CA-C-N	6.07	133.14	121.54
1	A	92	ARG	C-N-CA	6.07	133.14	121.54
1	A	96	VAL	CA-C-O	5.76	127.45	121.05
1	A	92	ARG	N-CA-C	5.60	117.36	107.61

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	1437	0	1439	33	0
2	B	31	0	35	5	0
3	A	25	0	0	2	0
All	All	1493	0	1474	33	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 11.

All (33) 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:201:SER:OG	2:B:500:ALA:C	1.89	1.16
1:A:248:SER:HA	1:A:254:HIS:CE1	2.18	0.78
1:A:201:SER:CB	2:B:500:ALA:C	2.66	0.68
1:A:248:SER:HA	1:A:254:HIS:HE1	1.58	0.65
1:A:101:MET:HG2	1:A:155:LEU:HD22	1.80	0.64
1:A:211:VAL:HG13	1:A:223:ILE:HG23	1.81	0.62
1:A:91:GLU:O	1:A:167:LYS:HD3	2.02	0.59
1:A:201:SER:OG	2:B:500:ALA:CA	2.51	0.58
1:A:94:GLY:HA3	1:A:166:GLU:OE1	2.04	0.57
1:A:91:GLU:O	1:A:167:LYS:CD	2.54	0.56
1:A:92:ARG:HG3	1:A:163:GLY:HA3	1.88	0.55
1:A:117:GLN:HA	1:A:120:MET:HE3	1.90	0.52
1:A:233:ALA:O	1:A:237:ILE:HG13	2.12	0.50
1:A:251:ASN:O	1:A:255:ILE:HG22	2.12	0.49
1:A:108:PHE:O	1:A:112:GLN:HG2	2.13	0.48
1:A:168:ARG:HE	1:A:222:GLY:HA3	1.78	0.48
1:A:94:GLY:HA3	1:A:166:GLU:CD	2.38	0.48
1:A:231:ILE:O	1:A:235:ILE:HG12	2.14	0.48
1:A:125:TRP:CD1	1:A:190:GLN:HG3	2.49	0.48
1:A:108:PHE:HB2	3:A:305:HOH:O	2.12	0.47
1:A:204:VAL:HG12	1:A:208:MET:HE2	1.97	0.47
1:A:205:PHE:CE1	1:A:237:ILE:HG12	2.49	0.47
1:A:165:VAL:HA	1:A:223:ILE:HD11	1.98	0.45
1:A:91:GLU:HG3	1:A:92:ARG:N	2.33	0.44
1:A:166:GLU:HG3	1:A:171:SER:HA	2.00	0.42
1:A:185:SER:HB2	1:A:253:ALA:HA	2.00	0.42
1:A:213:LEU:C	1:A:213:LEU:HD23	2.44	0.42
1:A:230:ILE:O	1:A:234:LEU:HD13	2.19	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:249:MET:HB2	2:B:499:MET:H	1.84	0.42
1:A:125:TRP:CD1	1:A:126:PRO:HD2	2.54	0.41
1:A:213:LEU:HD21	1:A:217:ARG:HD3	2.03	0.41
1:A:150[A]:HIS:CE1	2:B:499:MET:HG2	2.56	0.41
1:A:133:PHE:HB2	3:A:311:HOH:O	2.20	0.40

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	176/211 (83%)	170 (97%)	6 (3%)	0	100	100
2	B	2/4 (50%)	2 (100%)	0	0	100	100
All	All	178/215 (83%)	172 (97%)	6 (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	A	144/169 (85%)	138 (96%)	6 (4%)	26	45
2	B	3/3 (100%)	3 (100%)	0	100	100
All	All	147/172 (86%)	141 (96%)	6 (4%)	27	46

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	91	GLU
1	A	92	ARG
1	A	131	LEU
1	A	237	ILE
1	A	244	LEU
1	A	260	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	190	GLN
1	A	254	HIS

### 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](#)

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	A	179/211 (84%)	0.40	16 (8%) 15 12	15, 33, 68, 124	1 (0%)
2	B	4/4 (100%)	0.40	0 100 100	40, 42, 46, 52	0
All	All	183/215 (85%)	0.40	16 (8%) 16 13	15, 34, 66, 124	1 (0%)

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	150[A]	HIS	6.3
1	A	248	SER	5.4
1	A	91	GLU	4.8
1	A	92	ARG	4.7
1	A	269	SER	3.5
1	A	93	ALA	3.4
1	A	149	MET	3.2
1	A	227	ARG	2.9
1	A	215	GLY	2.9
1	A	224	TYR	2.6
1	A	117	GLN	2.5
1	A	249	MET	2.5
1	A	270	LEU	2.3
1	A	272	ALA	2.3
1	A	230	ILE	2.1
1	A	212	TRP	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 oligosaccharides 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.