



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 12, 2026 – 04:32 PM UTC

PDB ID : 8DF2 / pdb\_00008df2  
Title : The structure of the 'ALT' construct of the Amuc\_1438 glycopeptidase  
Authors : Medley, B.J.; Boraston, A.B.  
Deposited on : 2022-06-21  
Resolution : 2.35 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](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  
Xtriage (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

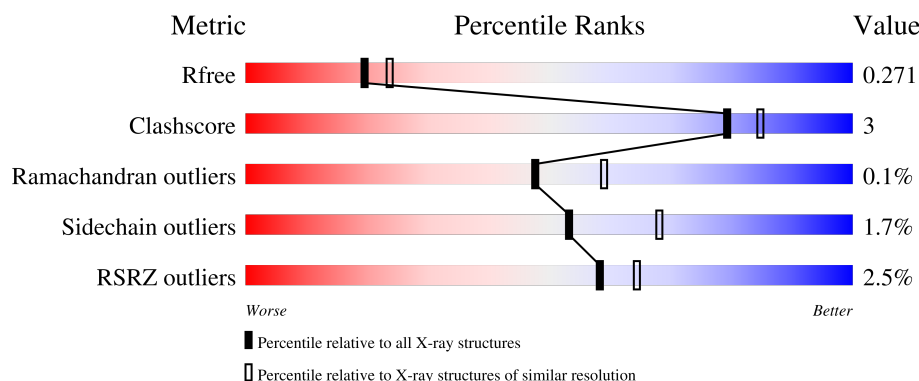
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1596 (2.36-2.36)
Clashscore	190562	1663 (2.36-2.36)
Ramachandran outliers	187476	1646 (2.36-2.36)
Sidechain outliers	187428	1646 (2.36-2.36)
RSRZ outliers	180081	1598 (2.36-2.36)

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	486	
1	B	486	
1	C	486	
1	D	486	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 11578 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 NPCBM/NEW2 domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	374	Total	C	N	O	S	0	0	0
			2798	1773	483	529	13			
1	B	374	Total	C	N	O	S	0	0	0
			2816	1784	487	532	13			
1	C	374	Total	C	N	O	S	0	1	0
			2807	1778	487	529	13			
1	D	374	Total	C	N	O	S	0	0	0
			2827	1790	491	532	14			

- Molecule 2 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Zn	0	0
			1	1		
2	B	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	D	1	Total	Ca	0	0
			1	1		

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total 3	Na 3	0	0
4	B	4	Total 4	Na 4	0	0
4	C	4	Total 4	Na 4	0	0
4	D	1	Total 1	Na 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	79	Total 79	O 79	0	0
5	B	89	Total 89	O 89	0	0
5	C	68	Total 68	O 68	0	0
5	D	74	Total 74	O 74	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.58Å 146.09Å 147.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.93 – 2.35 24.93 – 2.35	Depositor EDS
% Data completeness (in resolution range)	95.8 (24.93-2.35) 95.7 (24.93-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 2.36Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.222 , 0.270 0.223 , 0.271	Depositor DCC
$R_{free}$ test set	3874 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.7	Xtriage
Anisotropy	0.290	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 20.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	11578	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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](#)

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

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	0.22	0/2870	0.36	0/3899
1	B	0.23	0/2888	0.38	0/3920
1	C	0.23	0/2879	0.38	0/3902
1	D	0.21	0/2900	0.36	0/3939
All	All	0.23	0/11537	0.37	0/15660

There are no bond length outliers.

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	2798	0	2637	20	0
1	B	2816	0	2677	12	0
1	C	2807	0	2628	15	0
1	D	2827	0	2695	14	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	3	0	0	0	0
4	B	4	0	0	0	0
4	C	4	0	0	0	0
4	D	1	0	0	0	0
5	A	79	0	0	2	0
5	B	89	0	0	1	0
5	C	68	0	0	1	0
5	D	74	0	0	0	0
All	All	11578	0	10637	60	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 3.

The worst 5 of 60 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:366:CYS:SG	5:A:667:HOH:O	2.39	0.80
1:B:319:LEU:HD11	1:B:395:VAL:HG11	1.65	0.77
1:C:258:GLU:C	1:C:260:ARG:H	2.07	0.61
1:D:77:ARG:NH2	1:D:119:ASP:OD2	2.34	0.59
1:B:366:CYS:SG	5:B:684:HOH:O	2.48	0.55

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	370/486 (76%)	358 (97%)	12 (3%)	0	<b>100</b> <b>100</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	370/486 (76%)	361 (98%)	9 (2%)	0	100	100
1	C	371/486 (76%)	360 (97%)	11 (3%)	0	100	100
1	D	370/486 (76%)	359 (97%)	10 (3%)	1 (0%)	36	43
All	All	1481/1944 (76%)	1438 (97%)	42 (3%)	1 (0%)	48	59

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	259	LEU

### 5.3.2 Protein sidechains ⓘ

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	278/390 (71%)	272 (98%)	6 (2%)	45	59
1	B	280/390 (72%)	274 (98%)	6 (2%)	47	61
1	C	273/390 (70%)	269 (98%)	4 (2%)	57	72
1	D	285/390 (73%)	282 (99%)	3 (1%)	65	79
All	All	1116/1560 (72%)	1097 (98%)	19 (2%)	53	68

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	238	TRP
1	D	208	VAL
1	D	215	THR
1	D	163	GLN
1	B	163	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	88	GLN
1	C	213	ASN
1	D	163	GLN
1	A	383	ASN
1	B	86	ASN

### 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](#)

Of 20 ligands modelled in this entry, 20 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

### 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	374/486 (76%)	0.21	8 (2%) 63 68	24, 31, 42, 57	0
1	B	374/486 (76%)	0.24	7 (1%) 66 71	24, 31, 46, 64	0
1	C	374/486 (76%)	0.38	15 (4%) 42 48	16, 35, 50, 65	1 (0%)
1	D	374/486 (76%)	0.21	7 (1%) 66 71	25, 33, 42, 49	0
All	All	1496/1944 (76%)	0.26	37 (2%) 58 64	16, 32, 46, 65	1 (0%)

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	208	VAL	4.2
1	C	259	LEU	3.9
1	D	399	ALA	3.3
1	B	198	THR	3.2
1	C	374[A]	PHE	3.2

### 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](#)

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<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
4	NA	B	505	1/1	0.84	0.30	53,53,53,53	0
4	NA	B	506	1/1	0.84	0.24	48,48,48,48	0
4	NA	A	504	1/1	0.87	0.32	45,45,45,45	0
4	NA	B	503	1/1	0.88	0.19	40,40,40,40	0
4	NA	A	505	1/1	0.89	0.40	50,50,50,50	0
4	NA	A	503	1/1	0.91	0.26	50,50,50,50	0
4	NA	C	506	1/1	0.92	0.11	41,41,41,41	0
4	NA	C	503	1/1	0.93	0.29	52,52,52,52	0
4	NA	B	504	1/1	0.93	0.19	47,47,47,47	0
4	NA	D	503	1/1	0.93	0.21	44,44,44,44	0
4	NA	C	504	1/1	0.96	0.25	48,48,48,48	0
3	CA	C	502	1/1	0.98	0.06	34,34,34,34	0
2	ZN	A	501	1/1	0.98	0.07	35,35,35,35	0
4	NA	C	505	1/1	0.98	0.12	20,20,20,20	0
3	CA	A	502	1/1	0.98	0.03	31,31,31,31	0
3	CA	B	502	1/1	0.98	0.06	32,32,32,32	0
2	ZN	C	501	1/1	0.99	0.05	39,39,39,39	0
2	ZN	B	501	1/1	0.99	0.06	34,34,34,34	0
3	CA	D	502	1/1	0.99	0.03	36,36,36,36	0
2	ZN	D	501	1/1	1.00	0.03	36,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.