



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 18, 2022 – 03:28 am GMT

PDB ID : 7Q8L  
Title : Peptide VPCGTAHE in complex with human cathepsin V C25A mutant  
Authors : Loboda, J.; Sosnowski, P.; Tusar, L.; Vidmar, R.; Vizovisek, M.; Horvat, J.;  
Kosec, G.; Impens, F.; Demol, H.; Turk, B.; Gevaert, K.; Turk, D.  
Deposited on : 2021-11-11  
Resolution : 1.80 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.31.2  
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) : 2.31.2

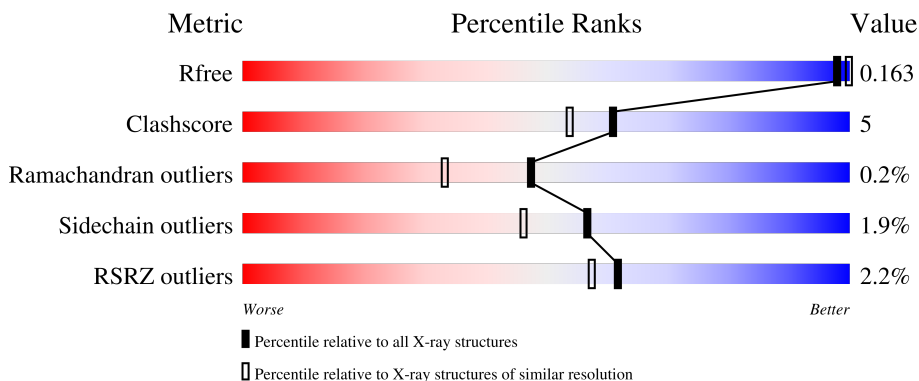
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	AA	222	
1	BA	222	
2	PA	10	
2	PB	10	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	GOL	BA	503	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8292 atoms, of which 4288 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 Cathepsin L2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	AA	222	3323	1069	1626	291	327	10	1636	0	0
1	BA	221	3311	1065	1622	290	324	10	1626	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AA	26	ALA	CYS	conflict	UNP O60911
AA	109	GLN	ASN	conflict	UNP O60911
AA	180	GLN	ASN	conflict	UNP O60911
BA	247	ALA	CYS	conflict	UNP O60911
BA	330	GLN	ASN	conflict	UNP O60911
BA	401	GLN	ASN	conflict	UNP O60911

- Molecule 2 is a protein called VPCGTAHE Peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
2	PA	5	60	18	28	7	7	28	0	1
2	PB	5	60	18	28	7	7	28	0	1

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AA	1	Total C O 8 6 2	0	0
3	AA	1	Total C O 8 6 2	0	0
3	BA	1	Total C O 8 6 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AA	1	Total Cl 1 1	0	0
4	BA	1	Total Cl 1 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AA	1	Total C O 6 3 3	0	0
5	AA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0
5	BA	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AA	249	Total H O 747 498 249	498	0
6	BA	240	Total H O 720 480 240	480	0
6	PA	1	Total H O 3 2 1	2	0
6	PB	2	Total H O 6 4 2	4	0

### 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: Cathepsin L2

Chain AA:  91% 9%




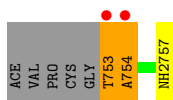
- Molecule 1: Cathepsin L2

Chain BA:  3% 93% 6%




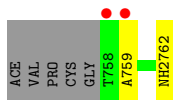
- Molecule 2: VPCGTAHE Peptide

Chain PA:  20% 20% 10% 20% 50%



- Molecule 2: VPCGTAHE Peptide

Chain PB:  20% 30% 20% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	94.35Å 94.35Å 127.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.18 – 1.80 47.18 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.18-1.80) 99.9 (47.18-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.50 (at 1.81Å)	Xtriage
Refinement program	MAIN	Depositor
R, $R_{free}$	0.167 , 0.188 0.164 , 0.163	Depositor DCC
$R_{free}$ test set	2672 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.7	Xtriage
Anisotropy	0.070	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>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8292	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: GOL, NH2, MPD, CL

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	AA	0.65	0/1739	0.75	0/2350
1	BA	0.61	0/1731	0.74	0/2339
2	PA	1.02	0/31	0.84	0/41
2	PB	1.21	0/31	0.89	0/41
All	All	0.64	0/3532	0.75	0/4771

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	AA	1697	1626	1626	15	0
1	BA	1689	1622	1619	10	0
2	PA	32	28	24	7	0
2	PB	32	28	24	5	0
3	AA	16	0	28	0	0
3	BA	8	0	14	2	0
4	AA	1	0	0	0	0
4	BA	1	0	0	0	0
5	AA	12	0	16	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	BA	24	0	28	0	0
6	AA	249	498	0	5	0
6	BA	240	480	0	1	0
6	PA	1	2	0	0	0
6	PB	2	4	0	0	0
All	All	4004	4288	3379	29	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 29 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:AA:165:HIS:HD1	2:PA:757:NH2:N	1.76	0.83
1:BA:291:PHE:CZ	2:PB:759:ALA:HB2	2.21	0.75
1:AA:142:HIS:HD2	6:AA:602:HOH:O	1.69	0.74
1:BA:291:PHE:CZ	2:PB:759:ALA:CB	2.72	0.73
1:BA:386:HIS:HD1	2:PB:762:NH2:N	1.90	0.68

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	AA	220/222 (99%)	214 (97%)	6 (3%)	0	100	100
1	BA	219/222 (99%)	210 (96%)	9 (4%)	0	100	100
2	PA	3/10 (30%)	1 (33%)	1 (33%)	1 (33%)	0	0
2	PB	3/10 (30%)	1 (33%)	2 (67%)	0	100	100
All	All	445/464 (96%)	426 (96%)	18 (4%)	1 (0%)	47	33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	PA	754	ALA

### 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	AA	179/179 (100%)	176 (98%)	3 (2%)	60	51
1	BA	178/179 (99%)	175 (98%)	3 (2%)	60	51
2	PA	3/6 (50%)	2 (67%)	1 (33%)	0	0
2	PB	3/6 (50%)	3 (100%)	0	100	100
All	All	363/370 (98%)	356 (98%)	7 (2%)	57	46

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BA	274	ASN
1	BA	311	TYR
2	PA	753	THR
1	BA	379	CYS
1	AA	158	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	AA	162	ASN
1	BA	383	ASN
1	BA	401	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 2 are monoatomic - leaving 9 for Mogul analysis.

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	$\# Z  > 2$
5	GOL	AA	304	-	5,5,5	0.59	0	5,5,5	1.28	1 (20%)
5	GOL	BA	505	-	5,5,5	1.40	0	5,5,5	1.26	0
3	MPD	AA	303	-	7,7,7	0.41	0	9,10,10	0.24	0
5	GOL	BA	504	-	5,5,5	0.16	0	5,5,5	0.54	0
5	GOL	BA	506	-	5,5,5	1.20	1 (20%)	5,5,5	0.25	0
5	GOL	AA	305	-	5,5,5	0.47	0	5,5,5	0.51	0
3	MPD	BA	502	-	7,7,7	0.33	0	9,10,10	0.30	0
5	GOL	BA	503	-	5,5,5	0.25	0	5,5,5	0.14	0
3	MPD	AA	301	-	7,7,7	0.56	0	9,10,10	0.38	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
5	GOL	AA	304	-	-	0/4/4/4	-
5	GOL	BA	505	-	-	3/4/4/4	-
3	MPD	AA	303	-	-	1/5/5/5	-
5	GOL	BA	504	-	-	0/4/4/4	-
5	GOL	BA	506	-	-	0/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	AA	305	-	-	0/4/4/4	-
3	MPD	BA	502	-	-	3/5/5/5	-
5	GOL	BA	503	-	-	0/4/4/4	-
3	MPD	AA	301	-	-	0/5/5/5	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	BA	506	GOL	O1-C1	-2.01	1.33	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	AA	304	GOL	C3-C2-C1	-2.84	100.67	111.70

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	BA	502	MPD	C2-C3-C4-O4
5	BA	505	GOL	O1-C1-C2-C3
5	BA	505	GOL	O1-C1-C2-O2
5	BA	505	GOL	O2-C2-C3-O3
3	BA	502	MPD	CM-C2-C3-C4

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	BA	502	MPD	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, 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	AA	222/222 (100%)	-0.64	0 <b>100</b> <b>100</b>	30, 37, 58, 86	5 (2%)
1	BA	221/222 (99%)	-0.35	6 (2%) 54 49	32, 40, 64, 109	3 (1%)
2	PA	4/10 (40%)	1.55	2 (50%) <b>0</b> <b>0</b>	65, 76, 103, 131	0
2	PB	4/10 (40%)	4.06	2 (50%) <b>0</b> <b>0</b>	84, 90, 144, 148	0
All	All	451/464 (97%)	-0.44	10 (2%) 62 57	30, 39, 66, 148	8 (1%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	PB	758	THR	9.3
2	PB	759	ALA	4.7
1	BA	402	ASN	4.5
2	PA	754	ALA	3.0
1	BA	398	ALA	2.9

### 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, 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
5	GOL	BA	503	6/6	0.70	0.57	127,128,130,131	0
5	GOL	BA	504	6/6	0.81	0.23	91,101,107,107	0
5	GOL	AA	305	6/6	0.82	0.21	88,90,93,101	0
5	GOL	BA	505	6/6	0.83	0.20	93,98,102,105	0
3	MPD	AA	303	8/8	0.84	0.22	81,88,103,108	0
5	GOL	BA	506	6/6	0.87	0.18	60,69,75,78	0
3	MPD	AA	301	8/8	0.91	0.20	54,59,79,87	0
5	GOL	AA	304	6/6	0.93	0.31	87,93,105,114	0
3	MPD	BA	502	8/8	0.95	0.22	66,69,73,75	0
4	CL	AA	302	1/1	0.99	0.06	42,42,42,42	0
4	CL	BA	501	1/1	0.99	0.04	42,42,42,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.