

wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 09:51 PM EDT

PDB ID : 4IAV

Title: G215S, A251G, T257A, D260G, T262D mutant of carboxypeptidase T from

Thermoactinomyces vulgaris with N-Sulfamoyl-L-phenylalanine

Authors: Timofeev, V.I.; Kuznetsov, S.A.; Akparov, V.K.; Kuranova, I.P.

Deposited on : 2012-12-07

Resolution : 1.35 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
A user guide is available at

https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

MolProbity: 4.02b-467

Mogul : 2022.3.0, CSD as543be (2022)

Xtriage (Phenix) : 1.20.1

EDS : 2.37.1

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

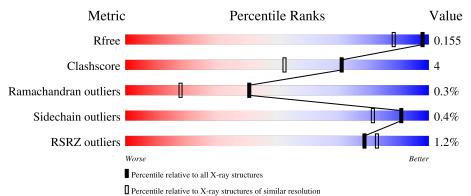
Validation Pipeline (wwPDB-VP) : 2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY\ DIFFRACTION$

The reported resolution of this entry is 1.35 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.



Whole archive Similar resolution Metric (#Entries) (#Entries, resolution range(Å)) R_{free} 1509 (1.38-1.34) 130704 Clashscore 141614 1551 (1.38-1.34) Ramachandran outliers 138981 1530 (1.38-1.34) Sidechain outliers 138945 1530 (1.38-1.34) RSRZ outliers 127900 1487 (1.38-1.34)

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 >=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 <=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	326	94%	5% •



2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 3216 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 Carboxypeptidase T.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	A	323	Total 2807	C 1757	N 473	O 565	S 12	0	26	0

There are 5 discrepancies between the modelled and reference sequences:

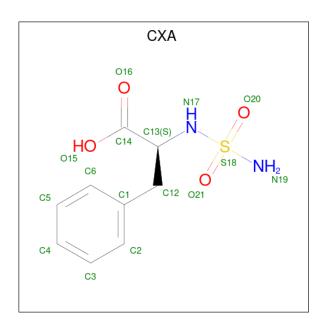
Chain	Residue	Modelled	Actual	Comment	Reference
A	215	SER	GLY	engineered mutation	UNP P29068
A	251	GLY	ALA	engineered mutation	UNP P29068
A	257	ALA	THR	engineered mutation	UNP P29068
A	260	GLY	ASP	engineered mutation	UNP P29068
A	262	ASP	THR	engineered mutation	UNP P29068

• Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atom	ns	ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0

• Molecule 3 is PHENYLALANINE-N-SULFONAMIDE (three-letter code: CXA) (formula: $C_9H_{12}N_2O_4S$).



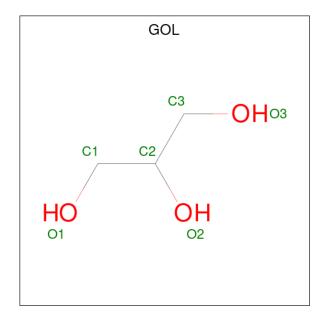


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf		
2	Λ	1	Total	С	N	О	S	0	0
)	A	1	16	9	2	4	1	0	0

• Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	5	Total Ca 5 5	0	0

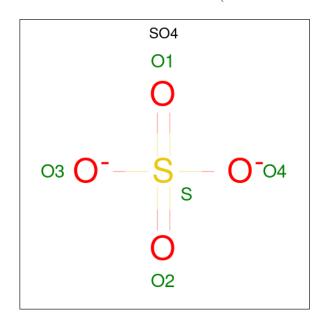
• Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 12 6 6	0	1
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0

 \bullet Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: $\mathrm{O_4S}).$



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0
6	A	1	Total O S 5 4 1	0	0

Continued on next page...



Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total O S 5 4 1	0	0

• Molecule 7 is water.

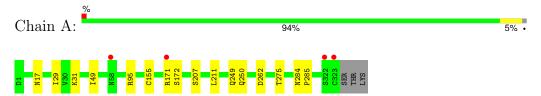
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	305	Total O 306 306	0	1



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: Carboxypeptidase T





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants	157.72Å 157.72Å 104.46Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 - 1.35	Depositor
resolution (A)	10.00 - 1.35	EDS
% Data completeness	99.7 (10.00-1.35)	Depositor
(in resolution range)	100.0 (10.00-1.35)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	3.78 (at 1.35Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
D D.	0.139 , 0.154	Depositor
R, R_{free}	0.142 , 0.155	DCC
R_{free} test set	8274 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	10.1	Xtriage
Anisotropy	0.098	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.48 , 70.0	EDS
L-test for twinning ²	$< L >=0.46, < L^2>=0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3216	wwPDB-VP
Average B, all atoms $(Å^2)$	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 2.74% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of <|L|>, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

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, ZN, CXA, SO4, CA

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		
		Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
Ī	1	A	0.33	0/2877	0.59	0/3902	

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	155	CYS	Peptide

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	2807	0	2616	19	0
2	A	1	0	0	0	0

Continued on next page...



Continued	trom	mmoninonic	maaa
COHABABACA		DIEUIUU	DUIUE
0 0 1000100000			

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	16	0	11	1	0
4	A	5	0	0	0	0
5	A	66	0	88	1	0
6	A	15	0	0	0	0
7	A	306	0	0	6	0
All	All	3216	0	2715	20	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 4.

The worst 5 of 20 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	Clash overlap (Å)	
1:A:17[B]:ASN:ND2	7:A:683:HOH:O	1.87	1.08	
1:A:262[A]:ASP:OD1	7:A:803:HOH:O	1.74	1.06	
1:A:262[B]:ASP:OD2	7:A:805:HOH:O	1.79	0.99	
1:A:31[B]:LYS:HD2	7:A:740:HOH:O	1.81	0.80	
1:A:95[B]:ARG:HH11	1:A:95[B]:ARG:HG3	1.47	0.79	

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	avoured Allowed		Percentiles	
1	A	350/326 (107%)	344 (98%)	5 (1%)	1 (0%)	41	18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	207	SER



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	310/284 (109%)	309 (100%)	1 (0%)	92	83	

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	250	GLN

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	249	GLN
1	A	250	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 21 ligands modelled in this entry, 6 are monoatomic - leaving 15 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).

Mal	Mol Type Chain Res		Link	Во	Bond lengths		Bond angles			
MIOI	Type	Chain	nes	Lilik	Counts	RMSZ	# Z > 2	Counts	RMSZ	$\mid \# Z > 2$
6	SO4	A	407	-	4,4,4	0.39	0	6,6,6	0.38	0
5	GOL	A	411	-	5,5,5	0.25	0	5,5,5	0.38	0
5	GOL	A	415	-	5,5,5	0.25	0	5, 5, 5	0.31	0
3	CXA	A	402	2	15,16,16	0.77	0	20,22,22	1.27	1 (5%)
5	GOL	A	404[B]	-	5,5,5	0.28	0	5,5,5	0.35	0
5	GOL	A	414	-	5,5,5	0.26	0	5,5,5	0.57	0
5	GOL	A	410	-	5,5,5	0.32	0	5,5,5	0.23	0
5	GOL	A	409	-	5,5,5	0.30	0	5,5,5	0.21	0
5	GOL	A	413	-	5,5,5	0.30	0	5,5,5	0.10	0
5	GOL	A	412	-	5,5,5	0.32	0	5,5,5	0.10	0
6	SO4	A	408	-	4,4,4	0.54	0	6,6,6	0.26	0
5	GOL	A	419	-	5,5,5	0.25	0	5, 5, 5	0.28	0
5	GOL	A	416	-	5,5,5	0.37	0	5,5,5	0.10	0
5	GOL	A	404[A]	-	5,5,5	0.27	0	5, 5, 5	0.21	0
6	SO4	A	418	-	4,4,4	0.49	0	6,6,6	0.18	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	A	411	-	-	1/4/4/4	-
5	GOL	A	415	-	-	0/4/4/4	-
3	CXA	A	402	2	-	0/12/13/13	0/1/1/1
5	GOL	A	404[B]	-	-	0/4/4/4	-
5	GOL	A	414	-	-	1/4/4/4	_
5	GOL	A	410	-	-	4/4/4/4	-
5	GOL	A	409	-	-	0/4/4/4	-
5	GOL	A	413	-	-	2/4/4/4	-
5	GOL	A	412	-	-	2/4/4/4	-
5	GOL	A	419	-	-	2/4/4/4	-
5	GOL	A	416	-	-	4/4/4/4	-
5	GOL	A	404[A]	_	_	2/4/4/4	-

There are no bond length outliers.



All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
3	A	402	CXA	C12-C13-N17	4.20	115.74	110.64

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	404[A]	GOL	C1-C2-C3-O3
5	A	410	GOL	O1-C1-C2-O2
5	A	410	GOL	O1-C1-C2-C3
5	A	410	GOL	C1-C2-C3-O3
5	A	410	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	402	CXA	1	0
5	A	414	GOL	1	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^{th} 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>	<RSRZ $>$ $#$ RSRZ $>$ 2		$OWAB(Å^2)$	Q < 0.9	
1	A	323/326 (99%)	-0.48	4 (1%)	79 8	3	7, 12, 24, 48	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	323	CYS	7.8
1	A	322	SER	4.5
1	A	58	ASN	2.7
1	A	171[A]	ARG	2.2

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 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^{th} 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	$ m B ext{-}factors(\AA^2)$	Q<0.9
6	SO4	A	418	5/5	0.70	0.25	29,36,40,43	5
5	GOL	A	415	6/6	0.76	0.33	31,37,44,47	6
5	GOL	A	419	6/6	0.78	0.30	36,40,44,52	0

Continued on next page...



Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
5	GOL	A	416	6/6	0.78	0.24	32,35,40,41	6
5	GOL	A	412	6/6	0.82	0.21	26,36,39,40	6
5	GOL	A	413	6/6	0.86	0.12	33,35,43,48	4
5	GOL	A	414	6/6	0.86	0.15	22,28,32,34	5
5	GOL	A	410	6/6	0.86	0.32	29,33,37,38	6
5	GOL	A	404[A]	6/6	0.90	0.23	21,23,26,26	6
6	SO4	A	408	5/5	0.90	0.16	17,23,26,31	5
5	GOL	A	404[B]	6/6	0.90	0.23	35,37,42,43	6
3	CXA	A	402	16/16	0.92	0.13	10,15,18,18	0
5	GOL	A	411	6/6	0.94	0.14	26,34,36,39	5
5	GOL	A	409	6/6	0.95	0.09	19,22,28,28	4
6	SO4	A	407	5/5	0.98	0.07	20,21,25,26	3
4	CA	A	417	1/1	0.99	0.27	31,31,31,31	0
4	CA	A	420	1/1	0.99	0.23	32,32,32,32	0
2	ZN	A	401	1/1	1.00	0.01	8,8,8,8	0
4	CA	A	403	1/1	1.00	0.06	17,17,17,17	1
4	CA	A	405	1/1	1.00	0.02	16,16,16,16	1
4	CA	A	406	1/1	1.00	0.04	19,19,19,19	0

6.5 Other polymers (i)

There are no such residues in this entry.

