



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2021 – 10:57 AM EDT

PDB ID : 2ET2  
Title : Crystal structure of an Asn to Ala mutant of Winged Bean Chymotrypsin Inhibitor protein  
Authors : Dattagupta, J.K.; Sen, U.; Dasgupta, J.; Khamrui, S.  
Deposited on : 2005-10-27  
Resolution : 2.10 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

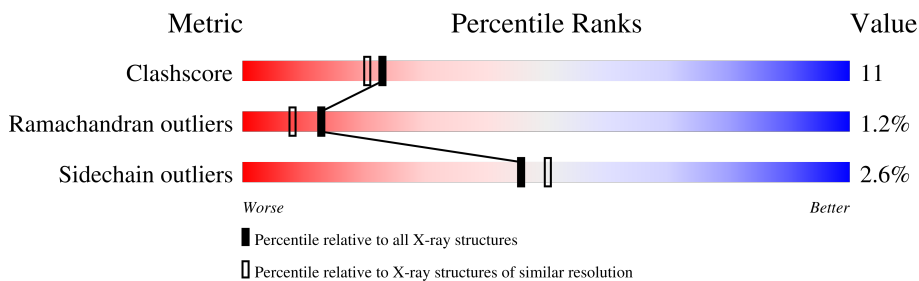
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	186	 72% 19% •• 6%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 1589 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 Chymotrypsin inhibitor 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	175	1366	869	235	258	4	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	cloning artifact	UNP P10822
A	2	GLU	-	cloning artifact	UNP P10822
A	3	PHE	-	cloning artifact	UNP P10822
A	17	ALA	ASN	engineered mutation	UNP P10822

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ni	0	0
			1	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	217	Total	O	0	0
			217	217		

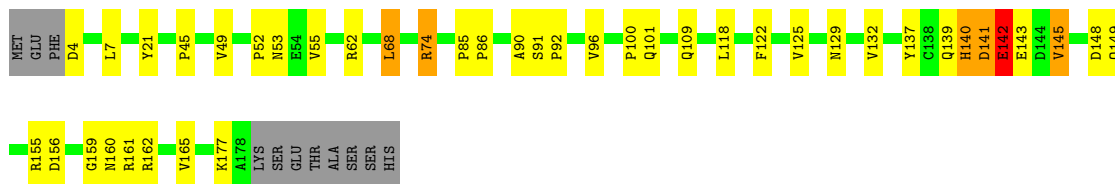
### 3 Residue-property plots [i](#)

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. 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. 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.

Note EDS was not executed.

- Molecule 1: Chymotrypsin inhibitor 3

Chain A:  72% 19% •• 6%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.39Å 60.39Å 212.18Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.93 – 2.10	Depositor
% Data completeness (in resolution range)	97.4 (19.93-2.10)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.227 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1589	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

## 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: SO4, NI

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	0.34	0/1400	0.77	4/1907 (0.2%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	142	GLU	N-CA-C	6.22	127.80	111.00
1	A	143	GLU	N-CA-C	-5.44	96.31	111.00
1	A	141	ASP	C-N-CA	5.03	134.28	121.70
1	A	141	ASP	N-CA-C	-5.01	97.48	111.00

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	1366	0	1346	31	0
2	A	1	0	0	0	0
3	A	5	0	0	1	0
4	A	217	0	0	7	0
All	All	1589	0	1346	31	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 (31) 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:53:ASN:OD1	1:A:55:VAL:HG12	1.85	0.75
1:A:159:GLY:HA2	4:A:389:HOH:O	1.90	0.72
1:A:155:ARG:HB3	4:A:389:HOH:O	1.97	0.64
1:A:52:PRO:HG2	4:A:330:HOH:O	1.97	0.64
1:A:148:ASP:OD2	1:A:148:ASP:N	2.27	0.63
1:A:140:HIS:C	1:A:142:GLU:H	2.06	0.58
1:A:141:ASP:H	1:A:145:VAL:HA	1.70	0.56
1:A:74:ARG:NE	4:A:359:HOH:O	2.39	0.55
1:A:109:GLN:NE2	4:A:302:HOH:O	2.38	0.53
1:A:62:ARG:HB2	1:A:62:ARG:NH1	2.24	0.53
1:A:125:VAL:HG22	1:A:132:VAL:HB	1.92	0.51
1:A:68:LEU:HB3	4:A:417:HOH:O	2.10	0.50
1:A:149:GLN:NE2	1:A:165:VAL:HG11	2.30	0.47
1:A:156:ASP:OD2	1:A:160:ASN:HB2	2.14	0.47
1:A:139:GLN:O	1:A:145:VAL:HG22	2.16	0.46
1:A:96:VAL:HG21	1:A:137:TYR:CD1	2.52	0.45
1:A:4:ASP:N	4:A:332:HOH:O	2.50	0.45
1:A:7:LEU:HD11	1:A:122:PHE:CD2	2.52	0.44
1:A:49:VAL:HG12	1:A:162:ARG:HA	1.99	0.44
1:A:129:ASN:N	3:A:197:SO4:O1	2.48	0.44
1:A:100:PRO:HB2	1:A:101:GLN:NE2	2.33	0.43
1:A:140:HIS:ND1	1:A:140:HIS:N	2.66	0.43
1:A:91:SER:HB2	1:A:92:PRO:HD2	2.00	0.43
1:A:125:VAL:CG2	1:A:132:VAL:HB	2.49	0.42
1:A:155:ARG:HG3	1:A:161:ARG:HG2	2.00	0.42
1:A:90:ALA:HB2	1:A:109:GLN:O	2.18	0.42
1:A:21:TYR:CZ	1:A:177:LYS:HD3	2.54	0.42
1:A:45:PRO:O	1:A:86:PRO:HG2	2.18	0.42
1:A:85:PRO:HA	1:A:86:PRO:HD3	1.94	0.42
1:A:118:LEU:HD12	1:A:118:LEU:HA	1.95	0.41
1:A:7:LEU:HD21	1:A:122:PHE:HB2	2.03	0.41

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	173/186 (93%)	163 (94%)	8 (5%)	2 (1%)	13   8

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	ARG
1	A	142	GLU

### 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	152/162 (94%)	148 (97%)	4 (3%)	46   50

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	68	LEU
1	A	140	HIS
1	A	142	GLU
1	A	145	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	101	GLN
1	A	149	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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\# Z  > 2$
3	SO4	A	197	-	4,4,4	0.19	0	6,6,6	0.19	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	197	SO4	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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.