



# Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 07:31 pm BST

PDB ID : 1GBT  
Title : STRUCTURE OF AN ACYL-ENZYME INTERMEDIATE DURING  
CATALYSIS: (GUANIDINOBENZOYL) TRYPSIN  
Authors : Singer, P.T.; Sweet, R.M.  
Deposited on : 1991-09-17  
Resolution : 2.00 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) : 2.11

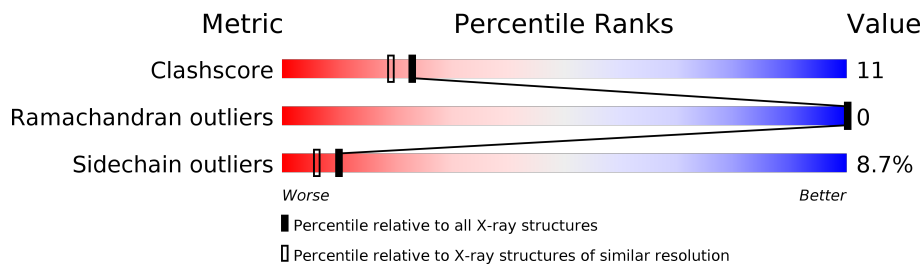
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 on 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	223	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 1761 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.

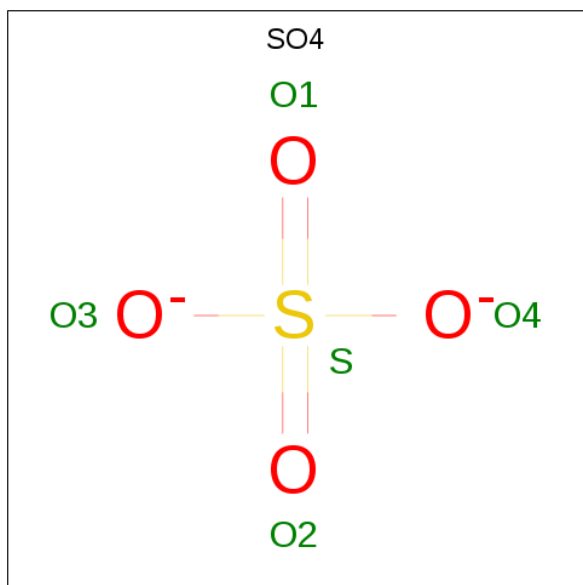
- Molecule 1 is a protein called BETA-TRYPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	1629	1012	279	324	14	0	0	0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

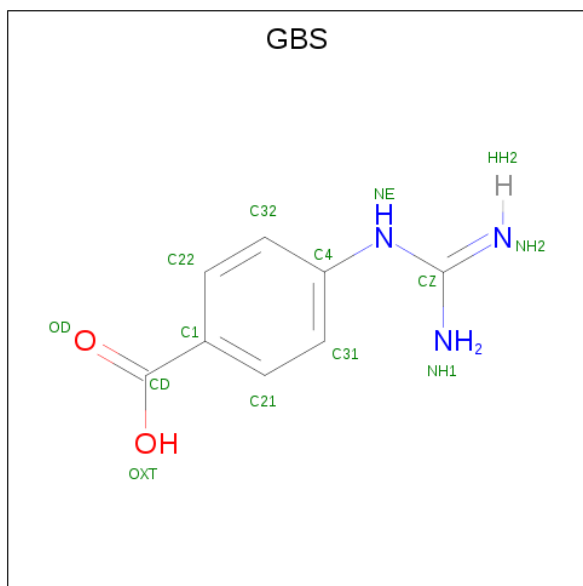
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		

- 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	S	0	0
			1	1		
3	A	1	Total	S	0	0
			1	1		

- Molecule 4 is 4-GUANIDINOBENZOIC ACID (three-letter code: GBS) (formula:  $C_8H_9N_3O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	12	8	3	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	117	117	117	0	0

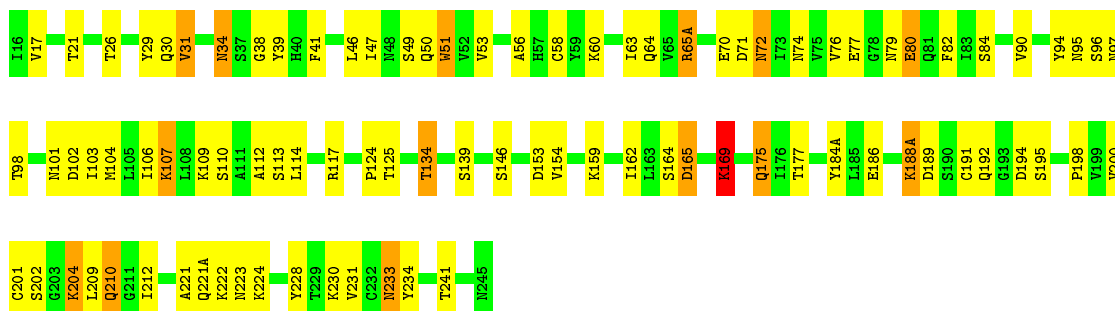
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: BETA-TRYPSIN

Chain A: 



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.74Å 63.54Å 68.93Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (7.00-2.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	TNT	Depositor
R, $R_{free}$	0.162 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1761	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	11.0	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: CA, GBS, SO4

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	1.07	6/1660 (0.4%)	1.99	50/2250 (2.2%)

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

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	195	SER	CB-OG	13.28	1.59	1.42
1	A	70	GLU	CD-OE1	6.90	1.33	1.25
1	A	198	PRO	N-CD	6.07	1.56	1.47
1	A	186	GLU	CD-OE2	5.57	1.31	1.25
1	A	77	GLU	CD-OE1	5.56	1.31	1.25
1	A	80	GLU	CD-OE2	5.13	1.31	1.25

All (50) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	98	THR	OG1-CB-CG2	-7.54	92.66	110.00
1	A	221	ALA	CB-CA-C	7.21	120.91	110.10
1	A	65(A)	ARG	NE-CZ-NH2	7.04	123.82	120.30
1	A	117	ARG	NE-CZ-NH1	-7.01	116.79	120.30
1	A	84	SER	CB-CA-C	-6.87	97.04	110.10
1	A	186	GLU	CB-CA-C	-6.87	96.66	110.40
1	A	71	ASP	CB-CG-OD1	-6.76	112.22	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	169	LYS	CA-CB-CG	-6.75	98.54	113.40
1	A	154	VAL	CG1-CB-CG2	6.67	121.56	110.90
1	A	58	CYS	CA-CB-SG	-6.62	102.08	114.00
1	A	112	ALA	N-CA-CB	-6.60	100.86	110.10
1	A	125	THR	CA-CB-CG2	-6.58	103.18	112.40
1	A	79	ASN	N-CA-CB	6.42	122.16	110.60
1	A	165	ASP	CB-CG-OD2	-6.39	112.55	118.30
1	A	221(A)	GLN	CB-CA-C	-6.34	97.71	110.40
1	A	233	ASN	CB-CA-C	-6.34	97.72	110.40
1	A	53	VAL	CG1-CB-CG2	-6.33	100.78	110.90
1	A	106	ILE	CA-CB-CG1	-6.25	99.13	111.00
1	A	195	SER	N-CA-CB	6.04	119.56	110.50
1	A	228	TYR	CD1-CE1-CZ	5.94	125.15	119.80
1	A	241	THR	CA-CB-CG2	-5.89	104.16	112.40
1	A	71	ASP	CB-CG-OD2	5.88	123.59	118.30
1	A	134	THR	CA-CB-CG2	-5.86	104.19	112.40
1	A	113	SER	CA-CB-OG	-5.83	95.45	111.20
1	A	31	VAL	CA-CB-CG2	5.77	119.56	110.90
1	A	177	THR	CA-CB-CG2	-5.71	104.40	112.40
1	A	21	THR	CA-CB-OG1	-5.68	97.06	109.00
1	A	204	LYS	N-CA-CB	5.67	120.81	110.60
1	A	169	LYS	N-CA-CB	5.62	120.71	110.60
1	A	139	SER	O-C-N	-5.54	113.77	123.20
1	A	103	ILE	O-C-N	-5.52	113.87	122.70
1	A	107	LYS	CA-CB-CG	-5.47	101.37	113.40
1	A	153	ASP	CB-CG-OD1	-5.47	113.38	118.30
1	A	50	GLN	CG-CD-OE1	-5.47	110.67	121.60
1	A	165	ASP	CB-CG-OD1	5.46	123.22	118.30
1	A	82	PHE	CG-CD2-CE2	-5.45	114.81	120.80
1	A	26	THR	CA-CB-CG2	-5.42	104.82	112.40
1	A	39	TYR	CG-CD1-CE1	5.41	125.62	121.30
1	A	49	SER	CA-CB-OG	-5.40	96.62	111.20
1	A	228	TYR	CB-CG-CD1	5.38	124.23	121.00
1	A	102	ASP	CB-CG-OD1	5.32	123.08	118.30
1	A	204	LYS	CD-CE-NZ	-5.28	99.56	111.70
1	A	184(A)	TYR	CB-CG-CD1	5.24	124.14	121.00
1	A	51	TRP	CB-CA-C	-5.12	100.16	110.40
1	A	82	PHE	CZ-CE2-CD2	5.11	126.23	120.10
1	A	165	ASP	N-CA-CB	-5.11	101.41	110.60
1	A	189	ASP	N-CA-CB	-5.08	101.47	110.60
1	A	53	VAL	CA-CB-CG1	-5.07	103.29	110.90
1	A	234	TYR	CG-CD2-CE2	5.05	125.34	121.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	50	GLN	CB-CA-C	-5.03	100.34	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	233	ASN	Mainchain

## 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	1629	0	1587	36	0
2	A	1	0	0	0	0
3	A	2	0	0	0	0
4	A	12	0	7	1	0
5	A	117	0	0	5	0
All	All	1761	0	1594	37	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 (37) 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:64:GLN:HE21	1:A:65(A):ARG:HE	1.01	0.99
1:A:64:GLN:HE22	1:A:65(A):ARG:HH21	1.17	0.89
1:A:64:GLN:NE2	1:A:65(A):ARG:HE	1.71	0.89
1:A:72:ASN:HD22	1:A:74:ASN:H	1.31	0.76
1:A:64:GLN:HE21	1:A:65(A):ARG:NE	1.84	0.65
1:A:72:ASN:ND2	1:A:74:ASN:H	1.95	0.64
1:A:51:TRP:CZ2	1:A:107:LYS:HD3	2.33	0.63
1:A:64:GLN:NE2	1:A:65(A):ARG:NE	2.46	0.62
1:A:64:GLN:HE22	1:A:65(A):ARG:NH2	1.95	0.60
1:A:34:ASN:ND2	1:A:38:GLY:H	2.01	0.59

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:64:GLN:NE2	1:A:65(A):ARG:HH21	1.95	0.56
1:A:134:THR:HB	1:A:162:ILE:HD12	1.89	0.54
1:A:46:LEU:HD21	1:A:114:LEU:HD21	1.91	0.53
1:A:56:ALA:HB1	1:A:90:VAL:HG13	1.93	0.50
1:A:94:TYR:HA	1:A:101:ASN:HB2	1.94	0.49
1:A:95:ASN:OD1	1:A:97:ASN:N	2.44	0.48
1:A:41:PHE:CE2	1:A:60:LYS:HE2	2.49	0.48
1:A:230:LYS:NZ	5:A:365:HOH:O	2.46	0.47
1:A:191:CYS:O	1:A:194:ASP:HB2	2.15	0.46
1:A:29:TYR:CZ	1:A:200:VAL:HG21	2.51	0.46
1:A:17:VAL:O	1:A:188(A):LYS:HA	2.16	0.46
1:A:41:PHE:HZ	1:A:63:ILE:HG12	1.81	0.46
1:A:47:ILE:HD13	1:A:47:ILE:HG21	1.71	0.44
1:A:109:LYS:HE2	1:A:109:LYS:HB3	1.68	0.44
1:A:210:GLN:NE2	5:A:277:HOH:O	2.49	0.44
1:A:76:VAL:HA	1:A:80:GLU:OE2	2.18	0.44
1:A:192:GLN:HG2	1:A:192:GLN:H	1.48	0.43
1:A:223:ASN:O	1:A:224:LYS:HE2	2.19	0.43
1:A:164:SER:HB2	5:A:349:HOH:O	2.18	0.43
4:A:704:GBS:HH11	4:A:704:GBS:C31	2.33	0.42
1:A:175:GLN:HB3	5:A:289:HOH:O	2.18	0.42
1:A:212:ILE:HG13	1:A:231:VAL:CG2	2.50	0.41
1:A:124:PRO:HD3	1:A:209:LEU:O	2.19	0.41
1:A:165:ASP:O	1:A:169:LYS:HD3	2.21	0.41
1:A:72:ASN:HD22	1:A:74:ASN:N	2.10	0.41
1:A:56:ALA:HA	1:A:104:MET:HB2	2.03	0.40
1:A:109:LYS:HE3	5:A:348: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	221/223 (99%)	214 (97%)	7 (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	184/184 (100%)	168 (91%)	16 (9%)	10	6

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	31	VAL
1	A	34	ASN
1	A	72	ASN
1	A	96	SER
1	A	110	SER
1	A	146	SER
1	A	159	LYS
1	A	169	LYS
1	A	175	GLN
1	A	188(A)	LYS
1	A	201	CYS
1	A	202	SER
1	A	204	LYS
1	A	210	GLN
1	A	222	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	34	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	64	GLN
1	A	72	ASN
1	A	74	ASN
1	A	210	GLN
1	A	240	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 1 is monoatomic and 2 are modelled with single atom - 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
4	GBS	A	704	1	12,12,13	1.55	2 (16%)	15,15,17	1.12	1 (6%)

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
4	GBS	A	704	1	-	0/6/6/8	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	704	GBS	CZ-NE	-4.00	1.33	1.37
4	A	704	GBS	OD-CD	2.54	1.30	1.21

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	704	GBS	OD-CD-C1	-2.38	116.88	124.59

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
4	A	704	GBS	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.