wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 15, 2020 – 10:25 pm BST

PDB ID : 1GHB

Title : A SECOND ACTIVE SITE IN CHYMOTRYPSIN? THE X-RAY CRYSTAL STRUCTURE OF N-ACETYL-D-TRYPTOPHAN BOUND TO GAMMA-CHYMOTRYPSIN

Authors : Yennawar, H.P.; Yennawar, N.H.; Farber, G.K.

Deposited on : 1994-04-06

Resolution : 2.00 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 ⓘ 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)
Xtriage (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 (i)

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.
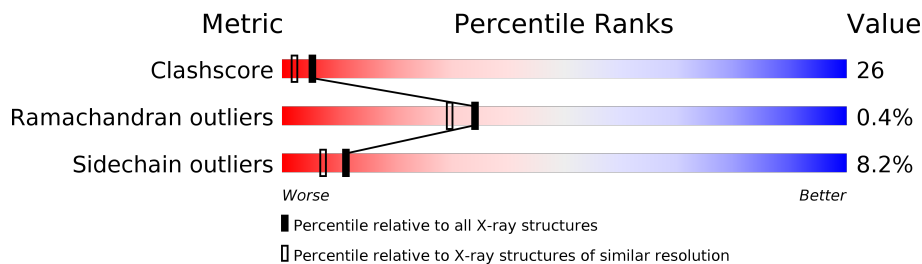
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 >=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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	13	54% 31% 15%
2	F	131	69% 27% •
3	G	97	58% 35% 5% •
4	P	3	67% 33%

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	ACE	E	600	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	TRP	E	601	X	-	X	-
7	HEX	E	651	-	-	X	-
7	HEX	E	653	-	-	X	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 2123 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 GAMMA-CHYMOTRYPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	E	11	Total	C	N	O	S	0	0	1
			69	45	12	11	1			

- Molecule 2 is a protein called GAMMA-CHYMOTRYPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	131	Total	C	N	O	S	0	0	0
			979	618	162	195	4			

- Molecule 3 is a protein called GAMMA-CHYMOTRYPSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	95	Total	C	N	O	S	0	0	0
			689	429	120	133	7			

- Molecule 4 is a protein called PRO-GLY-ALA.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	P	3	Total	C	N	O	0	0	0
			17	10	3	4			

- Molecule 5 is ACETYL GROUP (three-letter code: ACE) (formula: C_2H_4O).

ACE
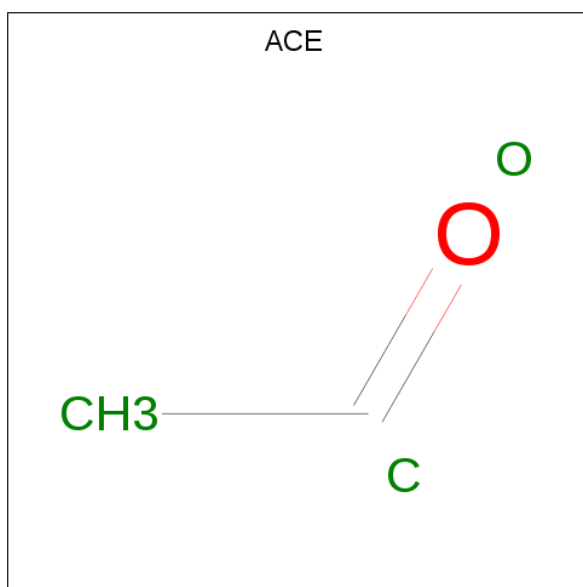
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	E	1	3	2	1	0	0

- Molecule 6 is TRYPTOPHAN (three-letter code: TRP) (formula: $C_{11}H_{12}N_2O_2$).

TRP
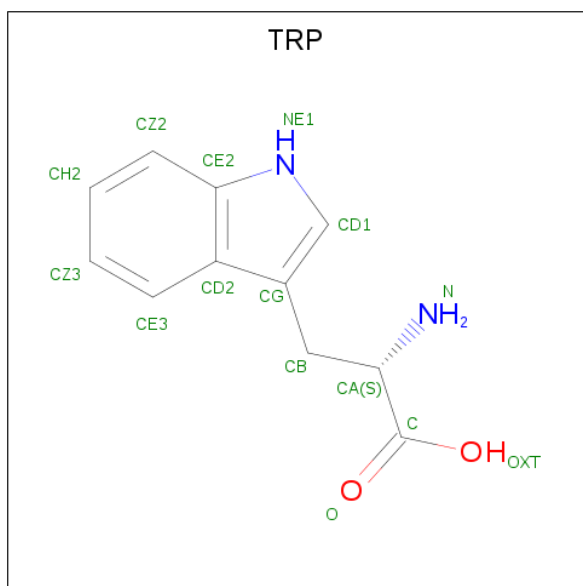
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	E	1	15	11	2	2	0	0

- Molecule 7 is HEXANE (three-letter code: HEX) (formula: C_6H_{14}).

HEX

C5

C6

C3

C4

C1

C2
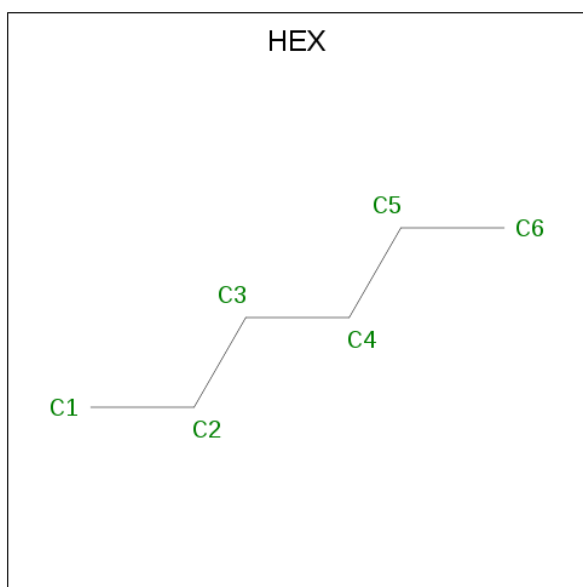
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	E	1	Total 6	C 6	0	0
7	E	1	Total 6	C 6	0	0
7	F	1	Total 6	C 6	0	0
7	F	1	Total 6	C 6	0	0
7	G	1	Total 6	C 6	0	0

- Molecule 8 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).

IPA

C3

C2

C1

OH

O2
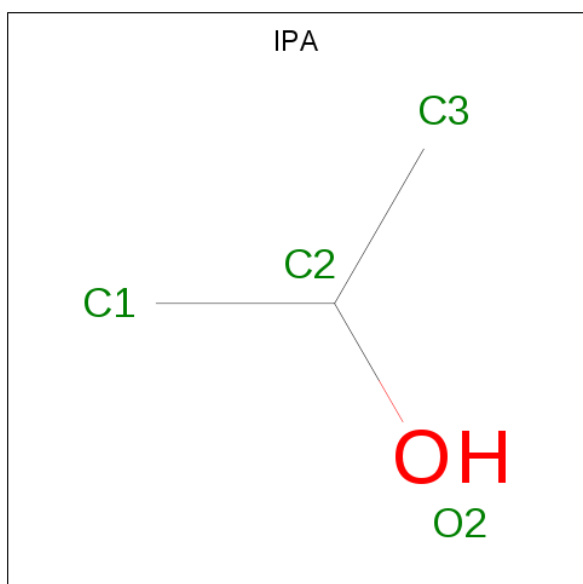
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	F	1	Total 4	C 3	O 1	0	0
8	G	1	Total 4	C 3	O 1	0	0
8	G	1	Total 4	C 3	O 1	0	0

- Molecule 9 is water.

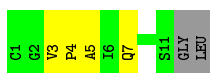| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	E	25	Total 25	O 25	0	0
9	F	174	Total 174	O 174	0	0
9	G	109	Total 109	O 109	0	0
9	P	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

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: GAMMA-CHYMOTRYPSIN

Chain E:

54% 31% 15%

C1 G2 V3 P4 A5 I6 Q7 S11 GLY LEU

- Molecule 2: GAMMA-CHYMOTRYPSIN

Chain F:

69% 27% •

I16 V17 N18 G19 E20 S26 W27 P28 W29 Q34 D35 F39 H40 F41 C42 I47 N48 E49 N50 V53 A56 F71 Q81 K82 L83 A86 K90 K93 Y94 N95 S96 L97 T98 I99 I103 T104 L105 L108 S109 Q116 T117 V118 S119 A120 V121

P124 S125 A126 S127 D128 C136 T144 R145 Y146
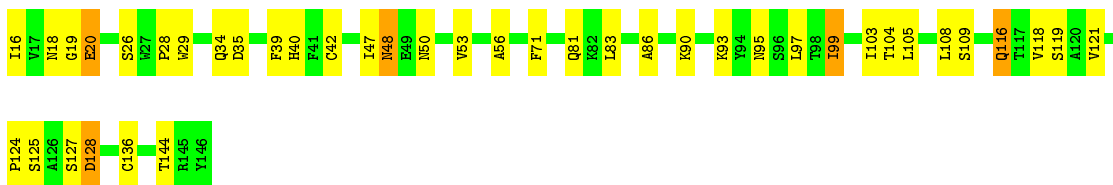
- Molecule 3: GAMMA-CHYMOTRYPSIN

Chain G:

58% 35% 5% •

ALA ASN T151 P152 D153 R154 A158 L162 N165 T166 N167 K170 Y171 W172 G173 T174 K175 I176 I181 G184 G187 S190 C191 M192 G193 G197 V200 C201 T208 L209 V210 G211 I212 V213 S214 W215 G216 S217 S218 T219 S223 G226 A229 R230 V231 T232

A233 L234 V238 Q239 Q240 M245
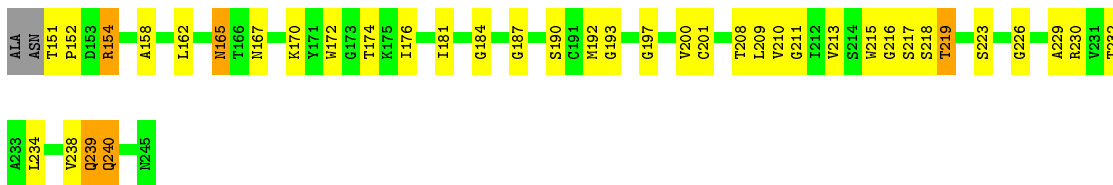
- Molecule 4: PRO-GLY-ALA

Chain P:

67% 33%

P571 G572 A573

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	69.60Å 69.60Å 97.51Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-2.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.152 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2123	wwPDB-VP
Average B, all atoms (Å^2)	8.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: IPA, HEX, ACE

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	E	0.78	0/70	0.95	0/97
2	F	0.59	0/999	0.76	0/1361
3	G	0.58	0/702	0.79	0/955
4	P	0.77	0/17	2.36	1/20 (5.0%)
All	All	0.60	0/1788	0.81	1/2433 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	P	571	PRO	CA-N-CD	-7.21	101.40	111.50

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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	E	69	0	76	7	0
2	F	979	0	951	53	0
3	G	689	0	687	39	0
4	P	17	0	15	0	0
5	E	3	0	3	3	0
6	E	15	0	10	14	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	E	12	0	28	15	0
7	F	12	0	28	3	0
7	G	6	0	14	1	0
8	F	4	0	8	3	0
8	G	8	0	16	4	0
9	E	25	0	0	2	0
9	F	174	0	0	2	0
9	G	109	0	0	1	0
9	P	1	0	0	0	0
All	All	2123	0	1836	96	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 26.

The worst 5 of 96 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)
2:F:104:THR:HG21	8:F:703:IPA:H11	1.34	1.10
2:F:90:LYS:HA	2:F:104:THR:HG22	1.35	1.08
6:E:601:TRP:N	6:E:601:TRP:HE3	1.73	0.85
6:E:601:TRP:CZ2	3:G:192:MET:SD	2.70	0.85
2:F:20:GLU:HB3	7:F:654:HEX:H52	1.63	0.80

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	E	9/13 (69%)	9 (100%)	0	0	100	100
2	F	129/131 (98%)	121 (94%)	7 (5%)	1 (1%)	19	13

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	G	93/97 (96%)	88 (95%)	5 (5%)	0	100	100
4	P	1/3 (33%)	0	1 (100%)	0	100	100
All	All	232/244 (95%)	218 (94%)	13 (6%)	1 (0%)	34	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	99	ILE

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	E	8/10 (80%)	8 (100%)	0	100	100
2	F	109/109 (100%)	104 (95%)	5 (5%)	27	23
3	G	76/77 (99%)	66 (87%)	10 (13%)	4	2
4	P	1/1 (100%)	1 (100%)	0	100	100
All	All	194/197 (98%)	179 (92%)	15 (8%)	11	8

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	154	ARG
3	G	165	ASN
3	G	232	THR
3	G	151	THR
3	G	223	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	F	116	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	240	GLN
3	G	165	ASN
2	F	48	ASN
3	G	239	GLN

5.3.3   RNA   ⓘ

There are no RNA molecules in this entry.

5.4   Non-standard residues in protein, DNA, RNA chains   ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5   Carbohydrates   ⓘ

There are no carbohydrates in this entry.

5.6   Ligand geometry   ⓘ

10 ligands are modelled in this entry.

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	$\#|Z| > 2$
7	HEX	F	654	-	5,5,5	0.76	0	4,4,4	0.15	0
5	ACE	E	600	6	1,2,2	0.79	0	1,1,1	0.72	0
8	IPA	F	703	-	3,3,3	0.51	0	3,3,3	0.25	0
7	HEX	E	651	-	5,5,5	0.20	0	4,4,4	0.33	0
8	IPA	G	702	-	3,3,3	0.69	0	3,3,3	0.26	0
8	IPA	G	701	-	3,3,3	0.42	0	3,3,3	0.20	0
7	HEX	F	655	-	5,5,5	0.26	0	4,4,4	0.37	0
7	HEX	G	652	-	5,5,5	0.36	0	4,4,4	0.33	0
7	HEX	E	653	-	5,5,5	0.41	0	4,4,4	0.25	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
7	HEX	F	654	-	-	0/3/3/3	-
7	HEX	E	651	-	-	0/3/3/3	-
7	HEX	F	655	-	-	0/3/3/3	-
7	HEX	G	652	-	-	0/3/3/3	-
6	TRP	E	601	5	1/1/2/2	-	-
7	HEX	E	653	-	-	0/3/3/3	-

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
6	E	601	TRP	CA

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 42 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	654	HEX	3	0
5	E	600	ACE	3	0
8	F	703	IPA	3	0
7	E	651	HEX	9	0
8	G	702	IPA	1	0
8	G	701	IPA	3	0
7	G	652	HEX	1	0
6	E	601	TRP	14	0
7	E	653	HEX	6	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)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.