



# Full wwPDB Geometry-Only Validation Report ⓘ

May 17, 2020 – 11:24 pm BST

PDB ID : 1GKT  
Title : Neutron Laue diffraction structure of endothiapepsin complexed with transition state analogue inhibitor H261  
Authors : Coates, L.; Erskine, P.T.; Wood, S.P.; Myles, D.A.A.; Cooper, J.B.  
Deposited on : 2001-08-20  
Resolution : 2.10 Å(reported)

This is a Full wwPDB Geometry-Only 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)  
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

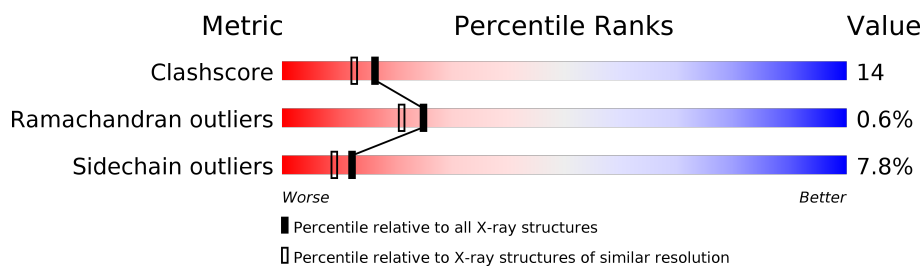
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*NEUTRON 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 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	329	
2	B	8	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5169 atoms, of which 2077 are hydrogens and 370 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 ENDOTHIAPEPSIN.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	329	Total	C	D	H	N	O	S	0	0	0
			4673	1514	285	2000	366	506	2			

- Molecule 2 is a protein called INHIBITOR, H261.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	8	Total	C	D	H	N	O	0	0	0
			159	55	3	77	13	11			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	248	Total	D	O	0	0
			330	82	248		
3	B	7	Total	O		0	0
			7	7			



Note EDS was not executed.

- Chain A:  68%  30%

S1	S4	T6	T7	T8	T9	A16	T19	T20	T21	T25	T26	L32	D33	F34	D35	S38	S39	D40	L41	L42	S45	S46	E52	O56	T57	K63	T66	S71	G72	A73	T74	T75	S76	T77	S78	T79	S95	L99	T100	V101	E107	S108	L109
K110	S115	F116	T117	E118	D119	D123	G124	L125	L126	A129	F130	S131	N134	S137	P138	K142	T143	F144	A150	L151	L152	D153	S154	T158	L161	G167	T168	Y169	N170	F171	G172	T176	K191	Q192	W195	T198	G201	Y202	S211	T212	D215		
A218	L224	L225	Y226	L227	P228	V232	Y235	S240	S246	F253	L259	P262	T263	G265	T266	G267	S268	A269	R270	I273	Y277	L278	D279	F280	G281	P282	L283	S284	S287	S288	G293	L294	Q295	I300	G301	I302	D307	K311	T322	K330			

- Chain B:  25% 63% 13%
- 

## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BOC, LOV, DOD, SUI

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	0.40	0/2432	0.70	0/3326
2	B	0.46	0/60	1.05	0/78
All	All	0.40	0/2492	0.71	0/3404

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
2	B	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
2	B	404	HIS	Peptide

### 4.2 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	2673	2000	2278	66	1
2	B	82	77	79	5	1
3	A	330	0	0	27	0
3	B	7	0	0	0	0
All	All	3092	2077	2357	69	1

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 14.

All (69) 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:56:GLN:NE2	1:A:119:ASP:O	1.98	0.96
1:A:101:VAL:HG13	3:A:2122:HOH:O	1.80	0.76
1:A:63:LYS:HD3	3:A:2083:HOH:O	1.83	0.72
1:A:282:PRO:O	2:B:400:BOC:H11	1.86	0.69
1:A:71:SER:HB3	3:A:2099:HOH:O	1.89	0.67
1:A:110:LYS:HE3	3:A:2131:HOH:O	1.90	0.67
1:A:46:SER:O	3:A:2069:HOH:O	2.12	0.67
1:A:26:PRO:HD2	3:A:2042:DOD:O	1.91	0.66
1:A:39:SER:OG	1:A:129:ALA:HB3	1.92	0.64
1:A:73:ALA:O	3:A:2101:HOH:O	2.16	0.64
1:A:138:PRO:HD2	3:A:2144:HOH:O	1.93	0.63
1:A:150:ALA:HB2	3:A:2153:DOD:O	1.94	0.62
1:A:153:ASP:OD1	3:A:2155:HOH:O	2.17	0.61
1:A:125:LEU:HD21	2:B:405:LOV:HD13	1.71	0.60
1:A:218:ALA:O	3:A:2205:HOH:O	2.20	0.60
1:A:277:TYR:O	3:A:2231:HOH:O	2.20	0.59
1:A:95:SER:HA	1:A:99:LEU:O	1.98	0.57
1:A:300:ILE:HG22	1:A:302:ILE:HG22	1.77	0.55
1:A:52:GLU:O	1:A:117:THR:HG23	2.02	0.55
1:A:269:ALA:O	1:A:270:ARG:HD3	2.04	0.53
1:A:75:TRP:CE2	1:A:107:GLU:HB3	2.37	0.53
1:A:115:SER:O	1:A:119:ASP:OD2	2.27	0.52
1:A:235:TYR:OH	1:A:278:ILE:HD13	2.04	0.52
1:A:95:SER:OG	1:A:100:THR:OG1	2.23	0.52
1:A:176:THR:HG21	3:A:2169:HOH:O	2.05	0.52
1:A:202:TYR:HA	1:A:263:THR:O	2.04	0.52
1:A:167:GLY:HA2	3:A:2157:HOH:O	2.06	0.51
1:A:228:PRO:O	1:A:232:VAL:HG23	2.05	0.51
2:B:400:BOC:C	2:B:402:PRO:HD3	2.35	0.51
1:A:215:ASP:OD1	3:A:2203:HOH:O	2.23	0.50

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:4:SER:OG	1:A:170:ASN:ND2	2.44	0.50
1:A:40:ASP:HB2	3:A:2062:HOH:O	2.06	0.50
1:A:129:ALA:O	3:A:2139:HOH:O	2.23	0.50
1:A:158:THR:OG1	1:A:170:ASN:HB2	2.07	0.49
1:A:35:ASP:OD2	1:A:38:SER:HB3	2.07	0.49
1:A:192:GLN:HG2	3:A:2188:HOH:O	2.07	0.48
1:A:40:ASP:N	3:A:2062:HOH:O	2.46	0.48
1:A:253:PHE:CZ	1:A:259:LEU:HD11	2.43	0.48
1:A:167:GLY:CA	3:A:2157:HOH:O	2.62	0.48
1:A:266:VAL:O	1:A:269:ALA:HB3	2.08	0.48
1:A:226:TYR:HA	1:A:295:GLN:O	2.09	0.47
1:A:45:SER:O	3:A:2066:HOH:O	2.25	0.47
1:A:100:THR:HB	3:A:2118:DOD:O	2.11	0.46
1:A:7:THR:HG21	1:A:161:LEU:HD23	1.87	0.46
1:A:25:THR:HG23	3:A:2037:HOH:O	2.10	0.46
1:A:21:VAL:HG21	1:A:32:LEU:HD12	1.87	0.46
1:A:154:SER:O	1:A:172:GLY:HA2	2.12	0.44
1:A:19:THR:HB	1:A:34:PHE:CE2	2.47	0.44
2:B:407:HIS:OXT	2:B:407:HIS:ND1	2.50	0.44
1:A:57:THR:N	1:A:123:ASP:OD1	2.50	0.44
1:A:235:TYR:CZ	1:A:278:ILE:HD13	2.48	0.44
1:A:307:ASP:O	1:A:311:LYS:HG2	2.12	0.44
1:A:40:ASP:OD1	1:A:134:ASN:HA	2.14	0.43
1:A:284:SER:HA	3:A:2235:HOH:O	2.13	0.43
1:A:322:THR:O	1:A:322:THR:HG22	2.12	0.43
1:A:5:ALA:O	1:A:168:THR:HG23	2.14	0.43
1:A:301:GLY:O	1:A:302:ILE:HB	2.14	0.42
1:A:77:ILE:HD13	1:A:79:TYR:OH	2.14	0.42
1:A:278:ILE:O	1:A:280:PHE:N	2.52	0.42
1:A:195:TRP:HB2	3:A:2205:HOH:O	2.14	0.42
1:A:151:SER:HA	3:A:2154:HOH:O	2.15	0.42
1:A:224:LEU:HD22	1:A:293:GLY:HA2	1.91	0.42
1:A:262:PHE:O	1:A:273:ILE:N	2.50	0.42
1:A:201:GLY:O	1:A:265:GLY:N	2.39	0.41
1:A:9:PRO:HA	1:A:16:ALA:O	2.16	0.41
1:A:42:TRP:HA	1:A:107:GLU:O	2.15	0.41
1:A:144:PHE:HB2	3:A:2125:HOH:O	2.16	0.41
1:A:131:SER:OG	1:A:142:LYS:HA	2.16	0.40
2:B:401:HIS:N	2:B:402:PRO:HD3	2.31	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:288:SER:HG	2:B:407:HIS:O[2_546]	1.53	0.07

### 4.3 Torsion angles [i](#)

#### 4.3.1 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	326/329 (99%)	306 (94%)	18 (6%)	2 (1%)	25	21
2	B	5/8 (62%)	4 (80%)	1 (20%)	0	100	100
All	All	331/337 (98%)	310 (94%)	19 (6%)	2 (1%)	25	21

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	302	ILE
1	A	279	ASP

#### 4.3.2 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	262/262 (100%)	242 (92%)	20 (8%)	13	10
2	B	6/6 (100%)	5 (83%)	1 (17%)	2	1
All	All	268/268 (100%)	247 (92%)	21 (8%)	12	9

All (21) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	1	SER
1	A	38	SER
1	A	41	LEU
1	A	66	THR
1	A	71	SER
1	A	78	SER
1	A	108	SER
1	A	115	SER
1	A	126	LEU
1	A	131	SER
1	A	137	SER
1	A	154	SER
1	A	191	LYS
1	A	198	THR
1	A	211	SER
1	A	212	THR
1	A	240	SER
1	A	246	SER
1	A	268	SER
1	A	287	SER
2	B	407	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	104	GLN
1	A	147	ASN
1	A	164	HIS
1	A	170	ASN

#### 4.3.3 RNA ⓘ

There are no RNA molecules in this entry.

### 4.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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
1	SUI	A	54	1	10,11,12	1.39	1 (10%)	11,15,17	1.87	3 (27%)
2	LOV	B	405	2	13,14,15	0.60	0	14,18,20	1.06	2 (14%)

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
1	SUI	A	54	1	-	0/2/19/20	0/1/1/1
2	LOV	B	405	2	-	2/18/18/20	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	54	SUI	CB-CG	2.99	1.55	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	54	SUI	C1-N2-CG	3.45	115.09	113.04
1	A	54	SUI	OD-CG-N2	3.04	127.28	123.92
1	A	54	SUI	O1-C1-N2	2.41	127.09	124.14
2	B	405	LOV	O-C-CA	-2.30	119.79	125.16
2	B	405	LOV	C1G-C1B-C1A	-2.12	111.26	115.82

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	405	LOV	OS-CS-CT-CA
2	B	405	LOV	C1A-CS-CT-CA

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	405	LOV	1	0

#### 4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

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

There are no ligands in this entry.

#### 4.7 Other polymers [i](#)

There are no such residues in this entry.

#### 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.