



# Full wwPDB X-ray Structure Validation Report ⓘ

Apr 15, 2024 – 12:00 PM EDT

PDB ID : 1PIP  
Title : CRYSTAL STRUCTURE OF PAPAIN-SUCCINYL-GLN-VAL-VAL-ALA-ALA-P-NITROANILIDE COMPLEX AT 1.7 ANGSTROMS RESOLUTION: NONCOVALENT BINDING MODE OF A COMMON SEQUENCE OF ENDOGENOUS THIOL PROTEASE INHIBITORS  
Authors : Yamamoto, A.; Tomoo, K.; Doi, M.; Ohishi, H.; Inoue, M.; Ishida, T.; Yamamoto, D.; Tsuboi, S.; Okamoto, H.; Okada, Y.  
Deposited on : 1992-10-03  
Resolution : 1.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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) : 2.36.1

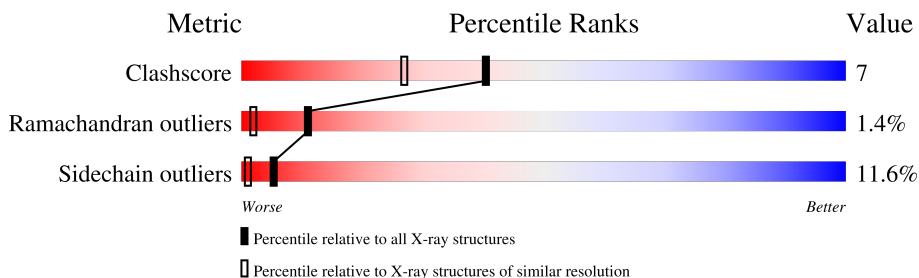
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.70 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	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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	212	
2	B	6	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 1705 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 Papain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	212	1655	1050	293	305	7	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	GLN	GLU	conflict	UNP P00784
A	118	GLN	GLU	conflict	UNP P00784
A	135	GLN	GLU	conflict	UNP P00784

- Molecule 2 is a protein called SUCCINYL-GLN-VAL-VAL-ALA-ALA-P-NITROANILIDE.

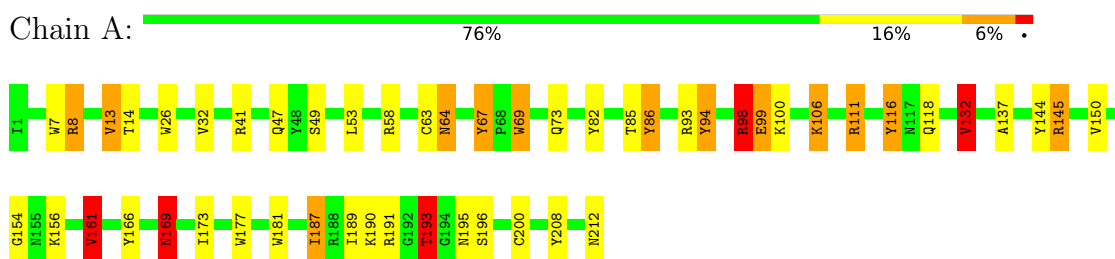
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	6	50	31	8	11	0	0	0

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

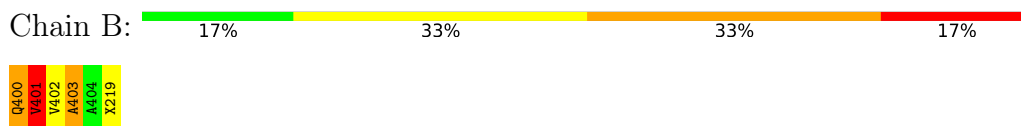
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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: Papain



- Molecule 2: SUCCINYL-GLN-VAL-VAL-ALA-ALA-P-NITROANILIDE



## 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, $\alpha$ , $\beta$ , $\gamma$	43.09Å 102.30Å 49.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 1.70	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-1.70)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.196 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1705	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	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: NIT, ZKO

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	0.90	0/1699	1.63	37/2306 (1.6%)
2	B	0.90	0/23	3.17	3/31 (9.7%)
All	All	0.90	0/1722	1.66	40/2337 (1.7%)

There are no bond length outliers.

All (40) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	98	ARG	NE-CZ-NH1	8.96	124.78	120.30
1	A	69	TRP	CD1-CG-CD2	8.33	112.96	106.30
2	B	401	VAL	CA-CB-CG2	-8.20	98.60	110.90
1	A	7	TRP	CD1-CG-CD2	7.57	112.35	106.30
1	A	111	ARG	NE-CZ-NH2	-7.45	116.57	120.30
1	A	116	TYR	CB-CG-CD2	-7.09	116.75	121.00
1	A	181	TRP	CD1-CG-CD2	6.89	111.81	106.30
1	A	69	TRP	CE2-CD2-CG	-6.89	101.79	107.30
1	A	7	TRP	CE2-CD2-CG	-6.84	101.83	107.30
1	A	177	TRP	CA-C-N	6.80	129.81	116.20
1	A	98	ARG	NE-CZ-NH2	-6.80	116.90	120.30
1	A	145	ARG	NE-CZ-NH1	6.79	123.69	120.30
1	A	177	TRP	O-C-N	-6.70	111.82	123.20
1	A	94	TYR	CA-CB-CG	6.66	126.05	113.40
1	A	64	ASN	CA-CB-CG	6.66	128.05	113.40
1	A	177	TRP	CD1-CG-CD2	6.60	111.58	106.30
2	B	403	ALA	N-CA-C	6.42	128.33	111.00
1	A	8	ARG	NE-CZ-NH2	-6.40	117.10	120.30
1	A	161	VAL	CG1-CB-CG2	6.29	120.97	110.90
1	A	177	TRP	CE2-CD2-CG	-6.20	102.34	107.30
1	A	132	VAL	N-CA-CB	-6.08	98.13	111.50
1	A	26	TRP	CD1-CG-CD2	6.04	111.13	106.30

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	63	CYS	O-C-N	-6.01	113.08	122.70
1	A	181	TRP	CE2-CD2-CG	-5.99	102.51	107.30
1	A	26	TRP	CE2-CD2-CG	-5.86	102.61	107.30
1	A	208	TYR	CB-CG-CD2	-5.83	117.50	121.00
1	A	161	VAL	N-CA-CB	-5.76	98.83	111.50
2	B	401	VAL	CA-CB-CG1	5.67	119.40	110.90
1	A	7	TRP	CG-CD1-NE1	-5.66	104.44	110.10
1	A	67	TYR	CB-CG-CD1	-5.64	117.61	121.00
1	A	193	THR	N-CA-CB	-5.60	99.65	110.30
1	A	166	TYR	CB-CG-CD2	-5.56	117.66	121.00
1	A	181	TRP	CG-CD1-NE1	-5.56	104.54	110.10
1	A	93	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	A	69	TRP	CG-CD1-NE1	-5.29	104.81	110.10
1	A	106	LYS	CA-CB-CG	5.20	124.84	113.40
1	A	169	ASN	CB-CG-ND2	5.13	129.01	116.70
1	A	145	ARG	NE-CZ-NH2	-5.10	117.75	120.30
1	A	94	TYR	CB-CG-CD1	5.08	124.05	121.00
1	A	49	SER	N-CA-CB	-5.01	102.99	110.50

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	A	1655	0	1596	23	2
2	B	50	0	32	6	2
All	All	1705	0	1628	24	2

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 7.

All (24) 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:137:ALA:HA	2:B:219:NIT:ON1	1.49	1.09
1:A:144:TYR:CE2	1:A:187:ILE:HG12	2.33	0.64
1:A:173:ILE:HD12	1:A:189:ILE:HG13	1.85	0.58
1:A:156:LYS:HE3	2:B:400:ZKO:C7	2.34	0.57
1:A:132:VAL:HG12	1:A:161:VAL:HG12	1.88	0.56
1:A:82:TYR:HB2	1:A:85:THR:HG23	1.90	0.52
2:B:400:ZKO:O	2:B:401:VAL:HB	2.09	0.52
1:A:150:VAL:HG12	1:A:190:LYS:HG3	1.92	0.51
1:A:137:ALA:HB2	2:B:401:VAL:HG21	1.93	0.51
1:A:118:GLN:NE2	1:A:191:ARG:HH11	2.09	0.50
1:A:137:ALA:CB	2:B:401:VAL:HG21	2.42	0.50
1:A:98:ARG:HH11	1:A:98:ARG:HB2	1.79	0.47
1:A:137:ALA:CA	2:B:219:NIT:ON1	2.42	0.47
1:A:161:VAL:HG13	1:A:173:ILE:HG23	1.97	0.46
1:A:116:TYR:HB3	1:A:196:SER:OG	2.15	0.46
1:A:169:ASN:H	1:A:169:ASN:HD22	1.63	0.45
1:A:154:GLY:O	1:A:200:CYS:HA	2.18	0.43
1:A:14:THR:HG21	1:A:47:GLN:HG2	2.02	0.41
1:A:144:TYR:CZ	1:A:187:ILE:HG12	2.54	0.41
1:A:67:TYR:HB3	1:A:69:TRP:NE1	2.36	0.41
1:A:99:GLU:HG2	1:A:100:LYS:HG3	2.02	0.41
1:A:193:THR:HG22	1:A:195:ASN:H	1.86	0.41
1:A:86:TYR:CD2	1:A:100:LYS:HD2	2.56	0.40
1:A:8:ARG:HA	1:A:13:VAL:HG13	2.03	0.40

All (2) 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:73:GLN:CG	2:B:400:ZKO:O4[2_765]	1.26	0.94
1:A:73:GLN:CG	2:B:400:ZKO:C8[2_765]	2.06	0.14

## 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	210/212 (99%)	203 (97%)	7 (3%)	0	100	100
2	B	3/6 (50%)	0	0	3 (100%)	0	0
All	All	213/218 (98%)	203 (95%)	7 (3%)	3 (1%)	11	2

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	401	VAL
2	B	403	ALA
2	B	402	VAL

### 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	170/170 (100%)	151 (89%)	19 (11%)	6	1
2	B	2/2 (100%)	1 (50%)	1 (50%)	0	0
All	All	172/172 (100%)	152 (88%)	20 (12%)	5	1

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	VAL
1	A	32	VAL
1	A	41	ARG
1	A	53	LEU
1	A	58	ARG
1	A	64	ASN
1	A	86	TYR
1	A	94	TYR
1	A	98	ARG
1	A	99	GLU
1	A	106	LYS

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
1	A	111	ARG
1	A	132	VAL
1	A	145	ARG
1	A	161	VAL
1	A	169	ASN
1	A	187	ILE
1	A	193	THR
1	A	212	ASN
2	B	401	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	64	ASN
1	A	169	ASN
1	A	212	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

1 non-standard protein/DNA/RNA residue is 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
2	ZKO	B	400	2	14,15,16	1.34	3 (21%)	16,18,20	4.16	9 (56%)

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
2	ZKO	B	400	2	-	2/15/16/18	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	400	ZKO	C6-C5	2.86	1.56	1.51
2	B	400	ZKO	O3-C8	-2.51	1.22	1.30
2	B	400	ZKO	C7-C8	2.01	1.55	1.50

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	400	ZKO	C6-C5-N	-12.33	94.45	115.83
2	B	400	ZKO	C-CA-N	-6.03	98.85	109.73
2	B	400	ZKO	CA-N-C5	-4.68	114.52	123.15
2	B	400	ZKO	C7-C6-C5	4.46	120.38	112.56
2	B	400	ZKO	O-C-CA	-3.87	114.65	124.78
2	B	400	ZKO	C6-C7-C8	2.82	119.67	113.60
2	B	400	ZKO	CG-CD-NE2	2.52	124.35	116.51
2	B	400	ZKO	O2-C5-C6	-2.21	117.97	122.02
2	B	400	ZKO	OE1-CD-NE2	-2.16	116.60	122.50

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	400	ZKO	O2-C5-N-CA
2	B	400	ZKO	C5-C6-C7-C8

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	400	ZKO	2	2

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 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.