



# Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 07:59 pm BST

PDB ID : 1MPT  
Title : CRYSTAL STRUCTURE OF A NEW ALKALINE SERINE PROTEASE (M-PROTEASE) FROM BACILLUS SP. KSM-K16  
Authors : Yamane, T.; Kani, T.; Hatanaka, T.; Suzuki, A.; Ashida, T.; Kobayashi, T.; Ito, S.; Yamashita, O.  
Deposited on : 1994-04-13  
Resolution : 2.40 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

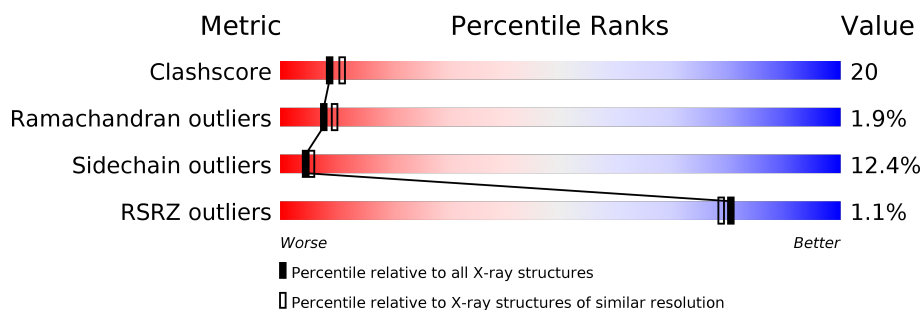
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	269	<div> <div></div> <div>57%</div> <div>34%</div> <div>7%</div> <div></div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 1928 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 M-PROTEASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	269	Total	C	N	O	S	0	0	0
			1882	1152	348	379	3			

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

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

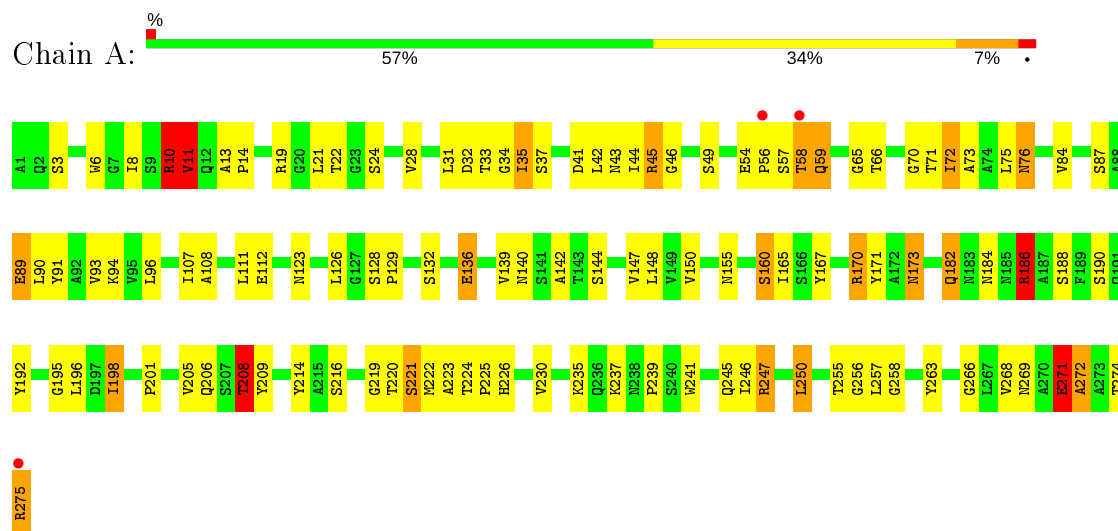
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	44	Total	O	0	0
			44	44		

### 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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

#### • Molecule 1: M-PROTEASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.82Å 57.79Å 54.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 2.40 19.76 – 2.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (7.00-2.40) 86.9 (19.76-2.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.14 (at 2.41Å)	Xtriage
Refinement program	PROLSQ, X-PLOR	Depositor
R, $R_{free}$	0.189 , (Not available) 0.193 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	14.4	Xtriage
Anisotropy	0.570	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 49.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	1928	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	8.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.31% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA

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	1.62	7/1915 (0.4%)	1.98	46/2613 (1.8%)

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	2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	160	SER	C-N	12.94	1.63	1.34
1	A	35	ILE	C-N	9.04	1.54	1.34
1	A	214	TYR	CG-CD2	-6.70	1.30	1.39
1	A	256	GLY	CA-C	6.31	1.61	1.51
1	A	167	TYR	CE2-CZ	6.10	1.46	1.38
1	A	132	SER	CB-OG	5.84	1.49	1.42
1	A	256	GLY	N-CA	5.13	1.53	1.46

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	170	ARG	NE-CZ-NH1	17.07	128.83	120.30
1	A	170	ARG	CD-NE-CZ	15.45	145.23	123.60
1	A	35	ILE	O-C-N	-15.01	98.69	122.70
1	A	247	ARG	NE-CZ-NH1	12.93	126.76	120.30
1	A	186	ARG	NE-CZ-NH2	-12.91	113.84	120.30
1	A	275	ARG	NE-CZ-NH1	-10.94	114.83	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	160	SER	O-C-N	-9.51	107.48	122.70
1	A	257	LEU	N-CA-C	-8.89	86.99	111.00
1	A	35	ILE	CA-C-N	8.57	136.06	117.20
1	A	41	ASP	CB-CG-OD1	8.47	125.92	118.30
1	A	247	ARG	NE-CZ-NH2	-8.40	116.10	120.30
1	A	142	ALA	CB-CA-C	8.23	122.45	110.10
1	A	170	ARG	NE-CZ-NH2	-8.21	116.20	120.30
1	A	10	ARG	NE-CZ-NH1	-8.17	116.22	120.30
1	A	160	SER	N-CA-CB	7.80	122.20	110.50
1	A	214	TYR	CB-CG-CD1	-7.46	116.52	121.00
1	A	45	ARG	NE-CZ-NH2	-7.37	116.62	120.30
1	A	271	GLU	CA-CB-CG	6.93	128.65	113.40
1	A	208	THR	N-CA-CB	-6.92	97.15	110.30
1	A	45	ARG	NE-CZ-NH1	6.83	123.72	120.30
1	A	275	ARG	NE-CZ-NH2	6.82	123.71	120.30
1	A	188	SER	N-CA-CB	6.48	120.22	110.50
1	A	32	ASP	CB-CG-OD2	6.32	123.99	118.30
1	A	73	ALA	N-CA-CB	-6.30	101.28	110.10
1	A	136	GLU	OE1-CD-OE2	6.18	130.72	123.30
1	A	59	GLN	CB-CA-C	6.01	122.42	110.40
1	A	43	ASN	N-CA-CB	5.95	121.32	110.60
1	A	19	ARG	NE-CZ-NH1	5.92	123.26	120.30
1	A	167	TYR	CB-CG-CD1	-5.88	117.47	121.00
1	A	11	VAL	CA-CB-CG1	5.80	119.59	110.90
1	A	275	ARG	CA-CB-CG	5.72	126.00	113.40
1	A	73	ALA	N-CA-C	5.70	126.39	111.00
1	A	93	VAL	CA-C-N	-5.69	104.68	117.20
1	A	271	GLU	CA-C-N	5.66	129.64	117.20
1	A	250	LEU	CB-CA-C	5.62	120.87	110.20
1	A	170	ARG	CA-CB-CG	5.58	125.68	113.40
1	A	136	GLU	CG-CD-OE2	-5.57	107.16	118.30
1	A	35	ILE	N-CA-CB	5.37	123.14	110.80
1	A	216	SER	N-CA-CB	5.28	118.42	110.50
1	A	186	ARG	CD-NE-CZ	5.24	130.94	123.60
1	A	271	GLU	N-CA-C	-5.22	96.90	111.00
1	A	186	ARG	CB-CA-C	5.15	120.69	110.40
1	A	263	TYR	CG-CD2-CE2	-5.10	117.22	121.30
1	A	195	GLY	O-C-N	5.10	130.85	122.70
1	A	209	TYR	CD1-CE1-CZ	-5.07	115.23	119.80
1	A	89	GLU	OE1-CD-OE2	-5.02	117.28	123.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	160	SER	Mainchain
1	A	186	ARG	Sidechain

## 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	1882	0	1841	76	0
2	A	2	0	0	0	0
3	A	44	0	0	0	0
All	All	1928	0	1841	76	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 20.

All (76) 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:182:GLN:H	1:A:182:GLN:HE21	1.13	0.94
1:A:57:SER:O	1:A:59:GLN:N	2.00	0.94
1:A:28:VAL:HG11	1:A:72:ILE:HD13	1.49	0.94
1:A:35:ILE:O	1:A:58:THR:O	1.91	0.88
1:A:37:SER:HA	1:A:58:THR:O	1.73	0.87
1:A:66:THR:O	1:A:208:THR:HG22	1.75	0.86
1:A:140:ASN:OD1	1:A:173:ASN:ND2	2.10	0.84
1:A:155:ASN:HD21	1:A:220:THR:H	1.27	0.83
1:A:35:ILE:O	1:A:58:THR:HG23	1.80	0.81
1:A:271:GLU:O	1:A:272:ALA:HB3	1.82	0.80
1:A:22:THR:HG23	1:A:87:SER:HB2	1.64	0.78
1:A:108:ALA:O	1:A:112:GLU:HG2	1.83	0.77
1:A:58:THR:HG22	1:A:59:GLN:OE1	1.85	0.75
1:A:230:VAL:HG12	1:A:250:LEU:HD11	1.70	0.72
1:A:57:SER:C	1:A:59:GLN:H	1.91	0.72
1:A:57:SER:C	1:A:59:GLN:N	2.44	0.69
1:A:155:ASN:HD22	1:A:220:THR:HG23	1.60	0.66

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:35:ILE:HG22	1:A:44:ILE:HD11	1.79	0.65
1:A:66:THR:HG22	1:A:208:THR:HG23	1.80	0.63
1:A:155:ASN:ND2	1:A:220:THR:H	1.96	0.62
1:A:56:PRO:O	1:A:94:LYS:HD3	2.00	0.62
1:A:54:GLU:C	1:A:56:PRO:N	2.55	0.60
1:A:28:VAL:HG11	1:A:72:ILE:CD1	2.26	0.60
1:A:37:SER:CA	1:A:58:THR:O	2.46	0.60
1:A:49:SER:HA	1:A:94:LYS:HB3	1.84	0.60
1:A:66:THR:O	1:A:208:THR:CG2	2.49	0.59
1:A:182:GLN:H	1:A:182:GLN:NE2	1.95	0.58
1:A:155:ASN:ND2	1:A:220:THR:HG23	2.18	0.57
1:A:235:LYS:HG2	1:A:246:ILE:HD11	1.87	0.56
1:A:57:SER:OG	1:A:58:THR:N	2.39	0.55
1:A:271:GLU:O	1:A:272:ALA:CB	2.45	0.55
1:A:165:ILE:CD1	1:A:196:LEU:HB2	2.38	0.53
1:A:46:GLY:O	1:A:91:TYR:HA	2.08	0.53
1:A:198:ILE:HG13	1:A:266:GLY:O	2.08	0.53
1:A:237:LYS:NZ	1:A:274:THR:O	2.34	0.53
1:A:22:THR:HG23	1:A:87:SER:CB	2.38	0.52
1:A:186:ARG:NH2	1:A:190:SER:O	2.28	0.51
1:A:13:ALA:HB3	1:A:14:PRO:HD3	1.92	0.51
1:A:35:ILE:C	1:A:58:THR:O	2.47	0.51
1:A:34:GLY:O	1:A:65:GLY:HA3	2.10	0.51
1:A:192:TYR:HA	1:A:196:LEU:HD22	1.93	0.51
1:A:35:ILE:O	1:A:58:THR:CG2	2.57	0.49
1:A:123:ASN:OD1	1:A:224:THR:HG22	2.12	0.49
1:A:6:TRP:CH2	1:A:201:PRO:HB3	2.48	0.49
1:A:45:ARG:HG3	1:A:89:GLU:HB3	1.95	0.48
1:A:136:GLU:HB3	1:A:171:TYR:CE2	2.48	0.48
1:A:35:ILE:CG2	1:A:44:ILE:HD11	2.42	0.48
1:A:22:THR:CG2	1:A:24:SER:HB2	2.44	0.48
1:A:136:GLU:HB3	1:A:171:TYR:CZ	2.49	0.47
1:A:70:GLY:HA3	1:A:208:THR:HB	1.96	0.47
1:A:13:ALA:N	1:A:14:PRO:CD	2.78	0.47
1:A:139:VAL:HG12	1:A:173:ASN:HB2	1.96	0.47
1:A:107:ILE:HD12	1:A:126:LEU:HD22	1.97	0.47
1:A:11:VAL:HG22	1:A:268:VAL:HG13	1.96	0.46
1:A:223:ALA:O	1:A:226:HIS:HB2	2.15	0.46
1:A:226:HIS:O	1:A:230:VAL:HG23	2.16	0.45
1:A:33:THR:HA	1:A:96:LEU:O	2.17	0.45
1:A:13:ALA:N	1:A:14:PRO:HD2	2.31	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:165:ILE:HD11	1:A:196:LEU:HB2	1.98	0.45
1:A:44:ILE:HG21	1:A:58:THR:OG1	2.17	0.45
1:A:71:THR:O	1:A:84:VAL:HG23	2.18	0.44
1:A:128:SER:HA	1:A:129:PRO:HD3	1.86	0.43
1:A:269:ASN:OD1	1:A:271:GLU:O	2.37	0.43
1:A:224:THR:N	1:A:225:PRO:HD2	2.34	0.43
1:A:107:ILE:HD12	1:A:126:LEU:CD2	2.48	0.43
1:A:57:SER:O	1:A:58:THR:C	2.56	0.43
1:A:237:LYS:NZ	1:A:275:ARG:C	2.72	0.42
1:A:31:LEU:N	1:A:31:LEU:HD12	2.34	0.42
1:A:10:ARG:HD3	1:A:10:ARG:HA	1.85	0.42
1:A:235:LYS:NZ	1:A:239:PRO:O	2.44	0.42
1:A:255:THR:O	1:A:266:GLY:HA3	2.20	0.41
1:A:219:GLY:C	1:A:221:SER:N	2.74	0.41
1:A:150:VAL:HG12	1:A:224:THR:HG23	2.03	0.41
1:A:275:ARG:HB2	1:A:275:ARG:HE	1.20	0.41
1:A:76:ASN:HD22	1:A:76:ASN:C	2.24	0.40
1:A:241:TRP:HA	1:A:245:GLN:OE1	2.22	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	265/269 (98%)	249 (94%)	11 (4%)	5 (2%)	<b>8</b> <b>10</b>

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	58	THR
1	A	184	ASN
1	A	258	GLY

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	272	ALA
1	A	198	ILE

### 5.3.2 Protein sidechains ⓘ

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	194/194 (100%)	170 (88%)	24 (12%)	<b>4</b> <b>5</b>

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	SER
1	A	8	ILE
1	A	10	ARG
1	A	11	VAL
1	A	21	LEU
1	A	42	LEU
1	A	72	ILE
1	A	75	LEU
1	A	76	ASN
1	A	90	LEU
1	A	111	LEU
1	A	144	SER
1	A	147	VAL
1	A	148	LEU
1	A	170	ARG
1	A	173	ASN
1	A	182	GLN
1	A	205	VAL
1	A	206	GLN
1	A	208	THR
1	A	221	SER
1	A	222	MET
1	A	247	ARG
1	A	271	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
1	A	76	ASN
1	A	155	ASN
1	A	182	GLN
1	A	184	ASN
1	A	204	ASN
1	A	249	HIS
1	A	261	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](#)

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 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	54:GLU	C	56:PRO	N	2.55
1	A	160:SER	C	165:ILE	N	1.63

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	269/269 (100%)	-0.50	3 (1%) 80 79	3, 8, 16, 27	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	58	THR	4.5
1	A	56	PRO	2.7
1	A	275	ARG	2.2

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	CA	A	277	1/1	0.97	0.20	26,26,26,26	0
2	CA	A	276	1/1	0.97	0.04	10,10,10,10	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.