wwPDB X-ray Structure Validation Summary Report (i)

Aug 8, 2023 – 03:25 AM EDT

PDB ID	:	1PWU
Title	:	Crystal Structure of Anthrax Lethal Factor complexed with (3-(N-hydroxyca rboxamido)-2-isobutylpropanoyl-Trp-methylamide), a known small molecule inhibitor of matrix metalloproteases.
Authors	:	Wong, T.Y.; Schwarzenbacher, R.; Liddington, R.C.
Deposited on	:	2003-07-02
Resolution	:	2.70 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 (i) 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 (1)) were used in the production of this report:

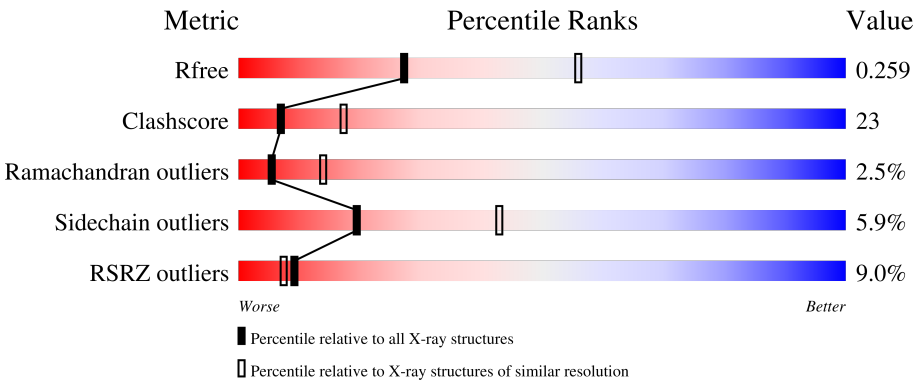
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.70 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))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 >=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% 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	776	9% 55% 36% 5% •
1	B	776	8% 57% 33% 5% 5%

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
3	GM6	A	1001	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 12238 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 Lethal factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	747	6134	3897	1032	1197	8	0	0	0
1	B	736	6046	3846	1014	1178	8	0	0	0

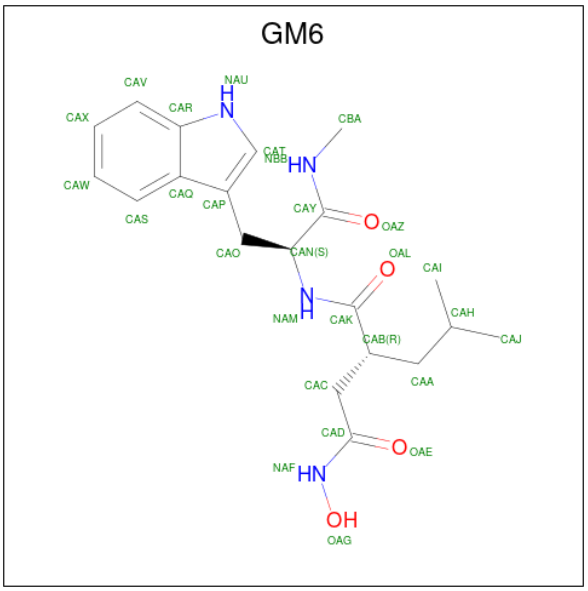
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	687	CYS	GLU	engineered mutation	UNP P15917
B	687	CYS	GLU	engineered mutation	UNP P15917

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is 3-(N-HYDROXYCARBOXAMIDO)-2-ISOBUTYLPROPANOYL-TRP-METHYLAMIDE (three-letter code: GM6) (formula: $C_{20}H_{28}N_4O_4$).



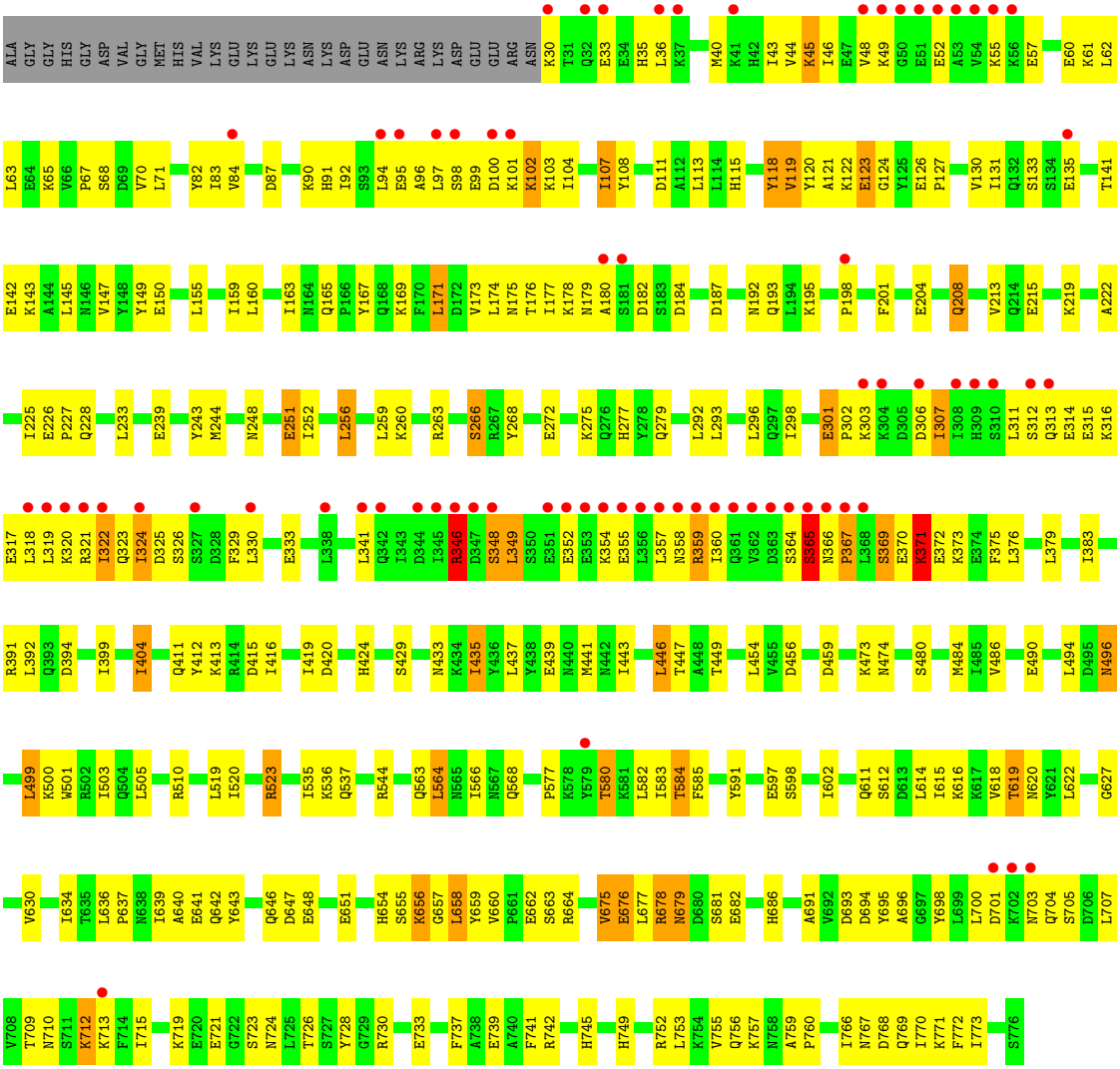
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	28	20	4	4	0	0
			Total	C	N	O		
3	B	1	28	20	4	4	0	0

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 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.

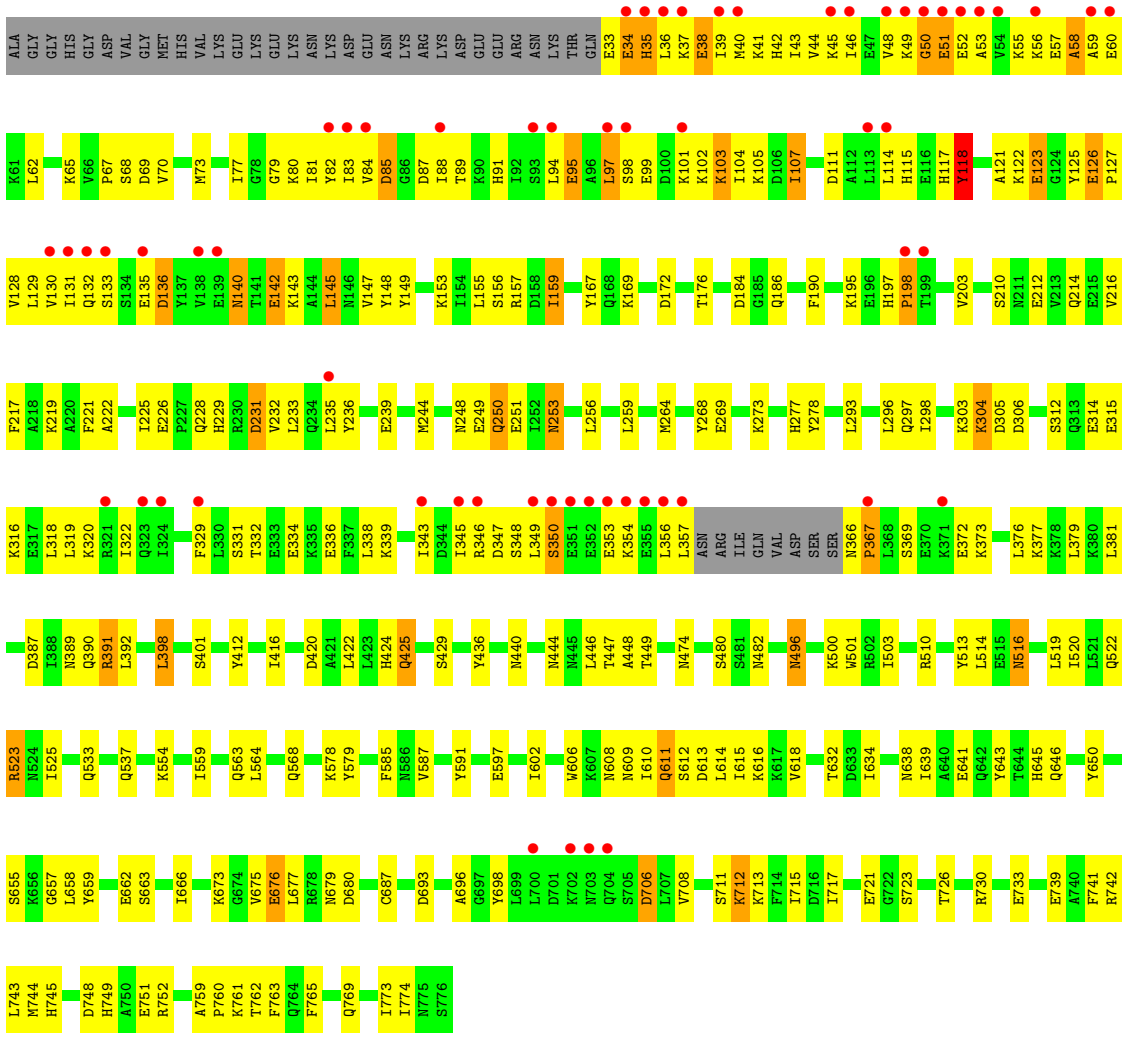
- Molecule 1: Lethal factor

Chain A:



- Molecule 1: Lethal factor

Chain B:


ALA GLY GLY HIS GLY ASP VAL GLY MET HIS VAL LYS GLU LYS GLU LYS ASN LYS ASP GLU ASN LYS ARG LYS ASP GLU GLU ARG ASN LYS THR GLN
E33 E34 H35 L36 K37 E38 I39 M40 K41 H42 I43 V44 K45 I46 E47 V48 K49 G50 E51 E52 A53 V54 K55 K56 E57 A58 A59 E60
K61 L62 K65 V66 P67 S68 D69 V70 M73 I77 G78 G79 K80 I81 Y82 I83 V84 D85 G86 D87 I88 T89 K90 H91 I92 S93 L94 E95 A96 L97 S98 E99 D100 K101 K102 K103 I104 K105 D106 I107 D111 A112 L113 L114 H115 E116 H117 Y118 A121 K122 E123 G124 Y125 E126 P127
V128 L129 V130 I131 Q132 S133 S134 E135 D136 Y137 V138 E139 N140 T141 E142 K143 A144 L145 N146 V147 Y148 Y149 K153 I154 L155 S156 R157 D158 I159 Y167 Q168 K169 D172 T176 D184 G185 Q186 F190 K195 E196 H197 P198 T199 V203 S210 N211 E212 V213 Q214 E215 V216
F217 A218 K219 A220 F221 A222 I225 E226 P227 Q228 H229 R230 D231 V232 L233 Q234 L235 Y236 E239 M244 N248 E249 Q250 E251 I252 N253 L256 L259 M264 Y268 E269 K273 H277 Y278 L293 L296 Q297 I298 K303 K304 D305 D306 S312 Q313 E314 E315
K316 E317 L318 L319 K320 R321 I322 Q323 I324 F329 L330 S331 T332 E333 E334 K335 E336 F337 L338 K339 I343 D344 I345 R346 D347 S348 L349 S350 E351 E352 E353 K354 E355 L356 L357 ASN ARG ILE GLN VAL ASP SER SER N366 P367 L368 S369 E370 K371 E372 K373 L376 K377 K378 L379 K380 L381
D387 I388 N389 Q390 R391 L392 L398 S401 Y412 I416 D420 A421 L422 L423 H424 Q425 S429 Y436 N440 N444 N445 L446 T447 A448 T449 N474 S480 S481 N482 N496 K500 W501 R502 I503 R510 Y513 L514 E515 N516 L519 I520 L521 Q522
R523 N524 I525 Q533 Q537 K554 I559 Q563 L564 Q568 K578 Y579 F585 N586 V587 Y591 E597 I602 W606 K607 N608 N609 I610 Q611 S612 D613 L614 I615 K616 K617 V618 T632 D633 I634 N638 I639 A640 E641 Q642 Y643 T644 H645 Q646 Y650
S655 K656 G657 L658 Y659 E662 S663 I666 K673 G674 V675 E676 L677 R678 N679 D680 C687 D693 A696 G697 Y698 L699 L700 D701 K702 N703 Q704 S705 D706 L707 V708 S711 K712 K713 F714 I715 D716 I717 E721 G722 S723 T726 R730 E733 E739 A740 F741 R742
L743 M744 H745 D748 H749 A750 E751 R752 A759 P760 K761 T762 F763 Q764 F765 Q769 I773 I774 N775 S776

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	96.70Å 137.40Å 98.30Å 90.00° 98.00° 90.00°	Depositor
Resolution (Å)	33.10 – 2.70 46.05 – 2.71	Depositor EDS
% Data completeness (in resolution range)	93.7 (33.10-2.70) 93.7 (46.05-2.71)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 2.69Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.235 , 0.270 0.226 , 0.259	Depositor DCC
R_{free} test set	3280 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	45.1	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 47.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	12238	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: ZN, GM6

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	0.38	0/6243	0.63	1/8411 (0.0%)
1	B	0.41	0/6154	0.65	3/8290 (0.0%)
All	All	0.39	0/12397	0.64	4/16701 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	118	TYR	CB-CG-CD2	-7.99	116.21	121.00
1	B	118	TYR	CB-CG-CD1	7.54	125.52	121.00
1	B	97	LEU	CB-CA-C	-5.88	99.02	110.20
1	A	365	SER	N-CA-C	5.26	125.20	111.00

There are no chirality outliers.

There are no planarity outliers.

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	6134	0	6123	283	0
1	B	6046	0	6033	283	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	28	0	28	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	28	0	28	5	0
All	All	12238	0	12212	563	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 23.

The worst 5 of 563 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:B:104:ILE:CD1	1:B:114:LEU:HD23	1.48	1.39
1:B:104:ILE:HD11	1:B:114:LEU:CD2	1.54	1.36
1:A:48:VAL:CG2	1:A:52:GLU:HG2	1.63	1.29
1:A:99:GLU:HA	1:A:102:LYS:HD2	1.11	1.10
1:B:366:ASN:HB3	1:B:367:PRO:HD3	1.26	1.09

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	745/776 (96%)	655 (88%)	66 (9%)	24 (3%)	4 9
1	B	732/776 (94%)	659 (90%)	60 (8%)	13 (2%)	8 21
All	All	1477/1552 (95%)	1314 (89%)	126 (8%)	37 (2%)	5 14

5 of 37 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	LYS
1	A	367	PRO
1	A	370	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	473	LYS
1	B	34	GLU

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	686/710 (97%)	640 (93%)	46 (7%)	16	37
1	B	675/710 (95%)	641 (95%)	34 (5%)	24	51
All	All	1361/1420 (96%)	1281 (94%)	80 (6%)	19	43

5 of 80 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	145	LEU
1	B	496	ASN
1	B	231	ASP
1	B	348	SER
1	B	676	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 68 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	608	ASN
1	B	611	GLN
1	B	767	ASN
1	A	679	ASN
1	A	652	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	$\#\|Z\| > 2$
3	GM6	A	1001	2	28,29,29	0.90	0	34,39,39	0.95	1 (2%)
3	GM6	B	2001	2	28,29,29	0.92	0	34,39,39	0.93	1 (2%)

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
3	GM6	A	1001	2	-	2/27/28/28	0/2/2/2
3	GM6	B	2001	2	-	2/27/28/28	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	2001	GM6	CBA-NBB-CAY	-2.73	117.49	122.22
3	A	1001	GM6	CBA-NBB-CAY	-2.71	117.53	122.22

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1001	GM6	CAN-CAY-NBB-CBA
3	A	1001	GM6	OAZ-CAY-NBB-CBA
3	B	2001	GM6	CAN-CAY-NBB-CBA
3	B	2001	GM6	OAZ-CAY-NBB-CBA

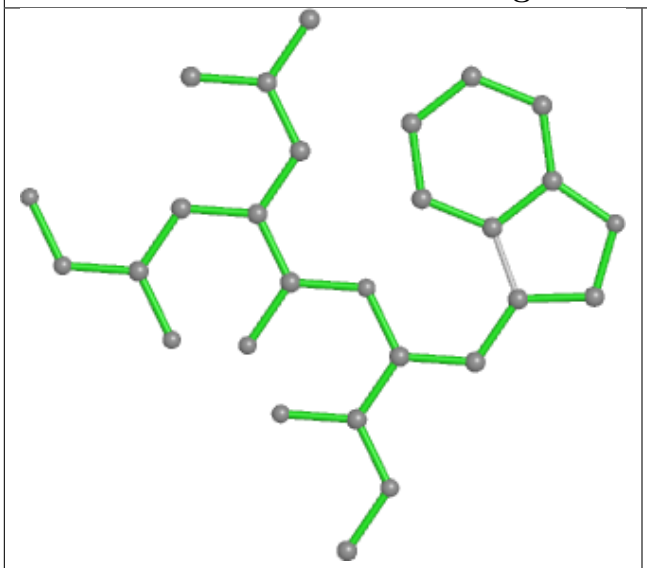
There are no ring outliers.

2 monomers are involved in 16 short contacts:

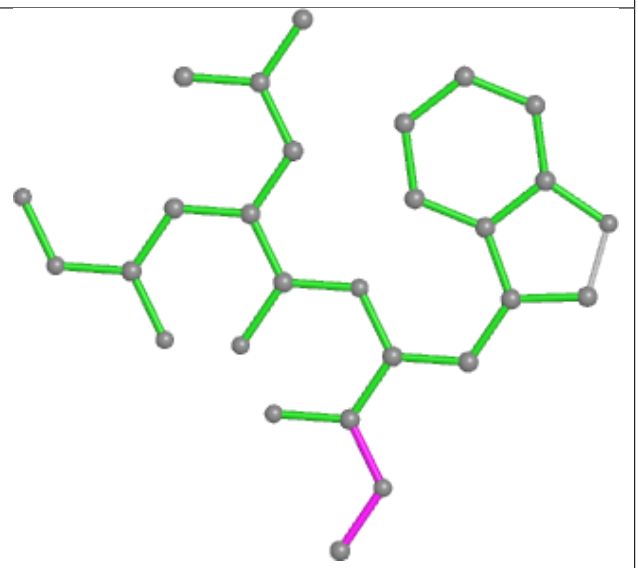
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	GM6	11	0
3	B	2001	GM6	5	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

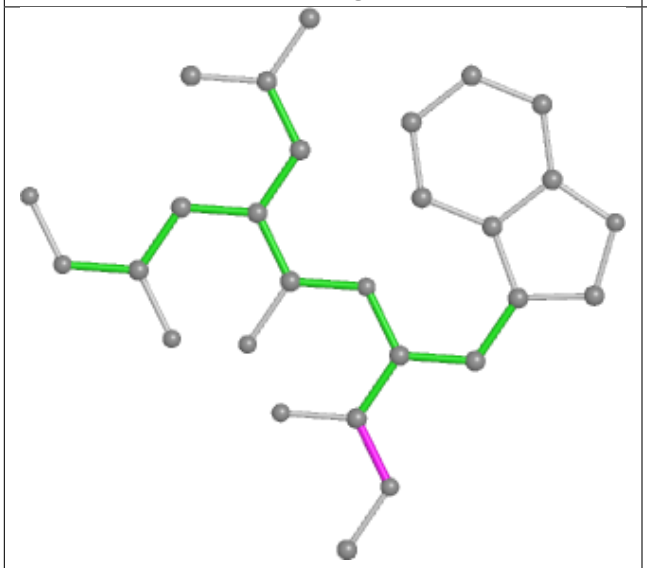
Ligand GM6 A 1001



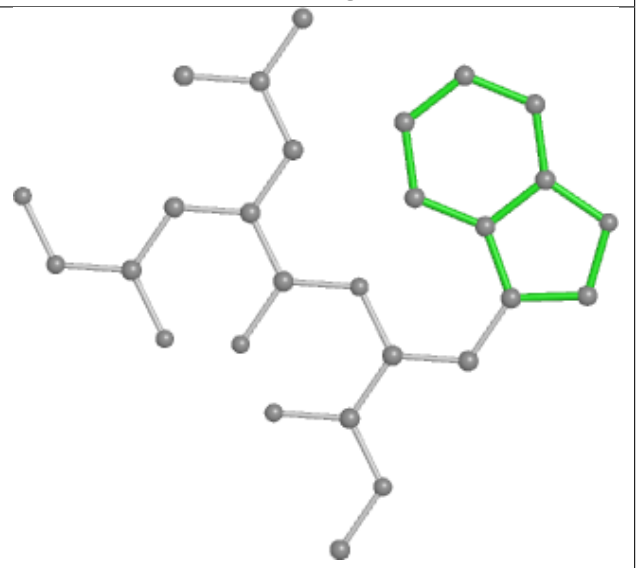
Bond lengths



Bond angles

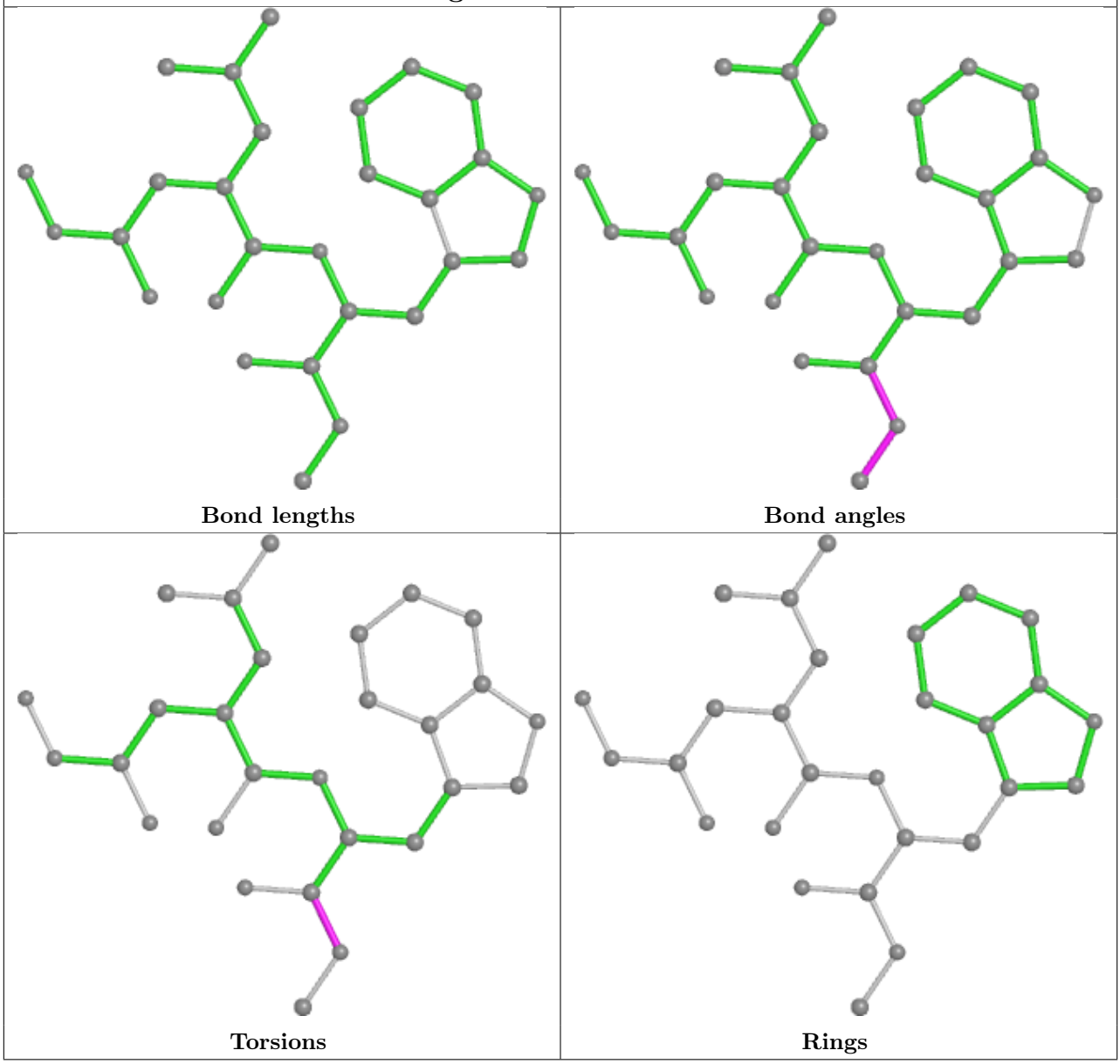


Torsions



Rings

Ligand GM6 B 2001



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)

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, 95th 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^2)	Q<0.9
1	A	747/776 (96%)	0.43	73 (9%)	7	5	21, 58, 126, 154	0
1	B	736/776 (94%)	0.25	61 (8%)	11	9	21, 51, 120, 152	0
All	All	1483/1552 (95%)	0.34	134 (9%)	9	7	21, 54, 123, 154	0

The worst 5 of 134 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	360	ILE	9.8
1	A	362	VAL	8.5
1	B	48	VAL	8.0
1	B	356	LEU	6.6
1	A	324	ILE	6.5

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 monosaccharides in this entry.

6.4 Ligands (i)

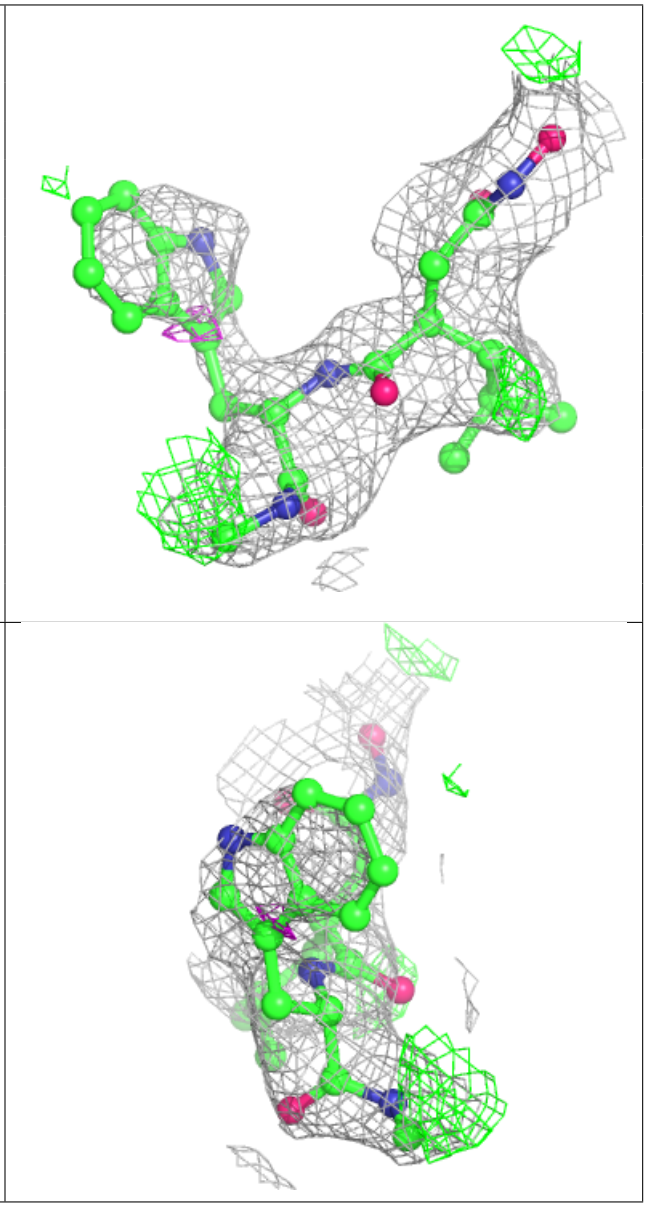
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GM6	A	1001	28/28	0.78	0.31	96,98,103,103	0
3	GM6	B	2001	28/28	0.86	0.30	95,98,102,103	0
2	ZN	A	9001	1/1	0.98	0.10	84,84,84,84	0
2	ZN	B	9002	1/1	0.99	0.09	75,75,75,75	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

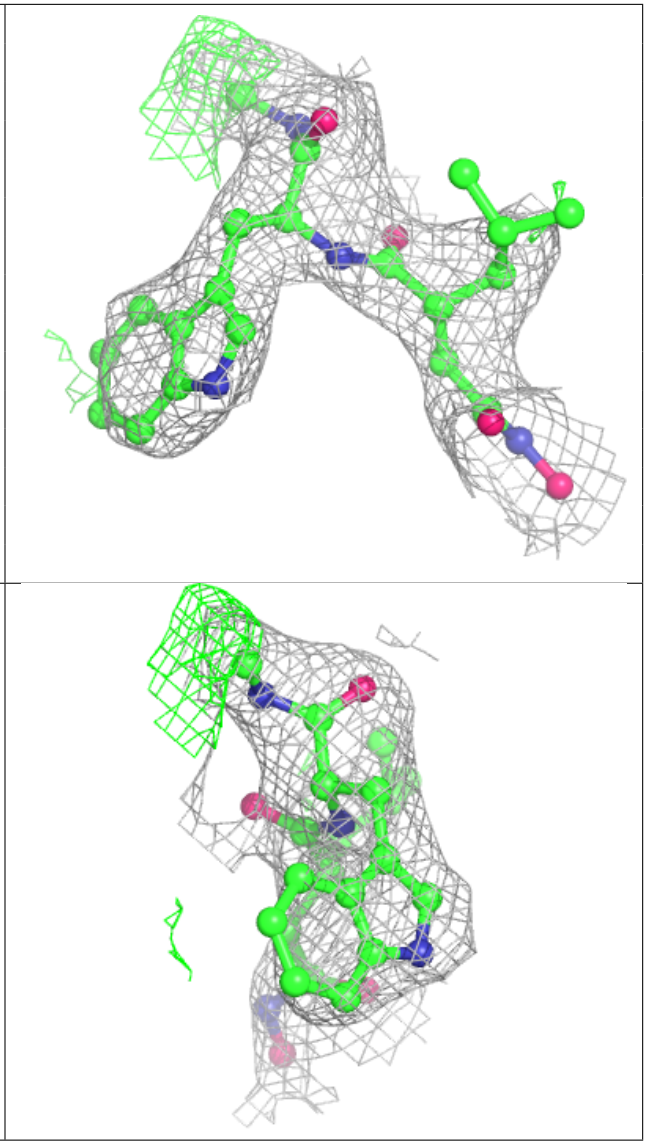
Electron density around GM6 A 1001:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around GM6 B 2001:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.