



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 06:27 PM GMT

PDB ID : 1AWF
Title : NOVEL COVALENT THROMBIN INHIBITOR FROM PLANT EXTRACT
Authors : Jhoti, H.; Cleasby, A.; Wonacott, A.
Deposited on : 1997-10-02
Resolution : 2.20 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
A user guide is available at
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

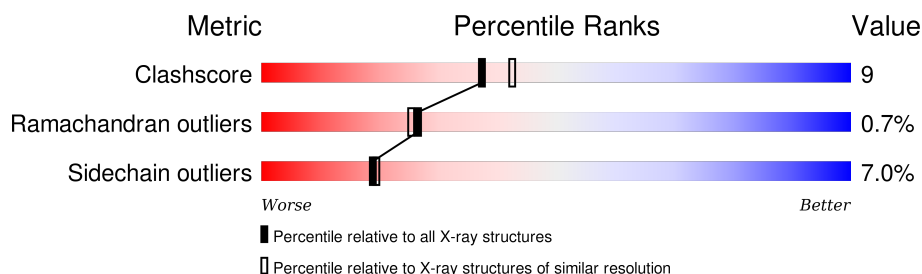
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	L	36	
2	H	259	
3	I	10	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 2682 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 ALPHA THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	27	Total	C	N	O	S	0	0	0
			222	140	36	45	1			

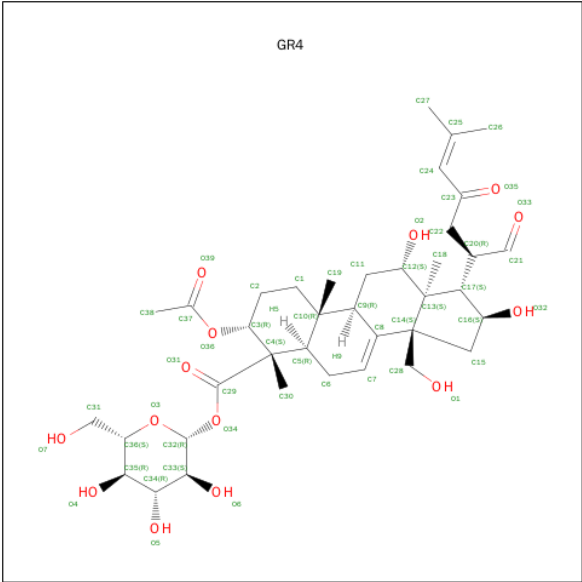
- Molecule 2 is a protein called ALPHA THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	258	Total	C	N	O	S	0	0	0
			2083	1329	369	371	14			

- Molecule 3 is a protein called HIRUGEN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	10	Total	C	N	O	S	0	0	0
			95	59	10	25	1			

- Molecule 4 is R3-ACETOXY-17-(1-FORMYL-5-METHYL-3-OXO-HEX-4-ENYL)-12,16-DI HYDROXY-14-HYDROXYMETHYL-4,10,13-TRIMETHYL-2,3,4,5,6,9,10,11,12,13,14,15,16,17-TETRADECAHYDRO-1H-CYCLOPENTA[A]PHENANTHRENE-4-CARBOXYLIC ACID IDOPYRANOSYL ESTER (three-letter code: GR4) (formula: C₃₈H₅₆O₁₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	H	1	Total	C	O	0	0
			52	38	14		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	185	Total	O	0	0
			185	185		
5	I	8	Total	O	0	0
			8	8		
5	L	37	Total	O	0	0
			37	37		

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 errors 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.

Note EDS was not executed.

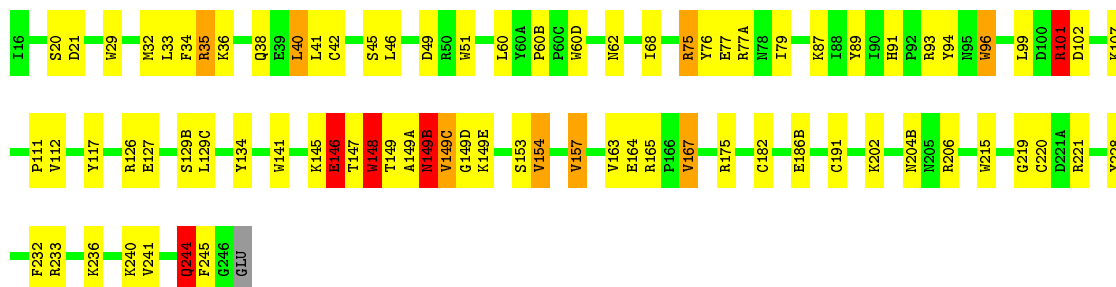
• Molecule 1: ALPHA THROMBIN

Chain L: 



• Molecule 2: ALPHA THROMBIN

Chain H: 



• Molecule 3: HIRUGEN

Chain I: 



4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	71.92Å 71.84Å 73.01Å 90.00° 100.39° 90.00°	Depositor
Resolution (Å)	15.00 – 2.20	Depositor
% Data completeness (in resolution range)	98.0 (15.00-2.20)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.196 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2682	wwPDB-VP
Average B, all atoms (Å ²)	27.0	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: GR4, TYS

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	L	0.89	0/224	1.50	1/298 (0.3%)
2	H	0.94	7/2138 (0.3%)	1.60	27/2891 (0.9%)
3	I	0.86	0/79	1.42	0/103
All	All	0.94	7/2441 (0.3%)	1.59	28/3292 (0.9%)

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	H	0	5

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	96	TRP	CD2-CE2	6.58	1.49	1.41
2	H	60(D)	TRP	CD2-CE2	6.03	1.48	1.41
2	H	51	TRP	CD2-CE2	5.97	1.48	1.41
2	H	141	TRP	CD2-CE2	5.96	1.48	1.41
2	H	148	TRP	CD2-CE2	5.77	1.48	1.41
2	H	29	TRP	CD2-CE2	5.29	1.47	1.41
2	H	215	TRP	CD2-CE2	5.14	1.47	1.41

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	75	ARG	NE-CZ-NH2	-11.37	114.62	120.30
2	H	165	ARG	NE-CZ-NH2	-8.81	115.89	120.30
2	H	206	ARG	NE-CZ-NH2	-8.52	116.04	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	94	TYR	CB-CG-CD2	-8.04	116.18	121.00
2	H	77(A)	ARG	NE-CZ-NH2	-7.27	116.67	120.30
2	H	134	TYR	CB-CG-CD2	-7.15	116.71	121.00
2	H	165	ARG	NE-CZ-NH1	7.02	123.81	120.30
2	H	157	VAL	CB-CA-C	-6.99	98.12	111.40
2	H	35	ARG	NE-CZ-NH2	-6.77	116.92	120.30
2	H	20	SER	CB-CA-C	-6.66	97.44	110.10
2	H	233	ARG	NE-CZ-NH1	6.47	123.54	120.30
2	H	93	ARG	NE-CZ-NH2	-6.37	117.12	120.30
2	H	101	ARG	NE-CZ-NH2	-6.35	117.13	120.30
2	H	77(A)	ARG	C-N-CA	-6.23	106.11	121.70
2	H	157	VAL	CA-CB-CG2	6.10	120.06	110.90
2	H	76	TYR	CB-CG-CD2	-5.87	117.48	121.00
2	H	40	LEU	CA-CB-CG	5.83	128.72	115.30
2	H	149	THR	N-CA-C	5.70	126.40	111.00
2	H	154	VAL	N-CA-CB	-5.68	99.01	111.50
2	H	149(B)	ASN	N-CA-C	5.52	125.89	111.00
1	L	1(B)	ALA	N-CA-CB	5.48	117.77	110.10
2	H	134	TYR	CB-CG-CD1	5.41	124.24	121.00
2	H	206	ARG	NE-CZ-NH1	5.37	122.98	120.30
2	H	186(B)	GLU	C-N-CA	-5.36	111.04	122.30
2	H	146	GLU	N-CA-C	5.36	125.46	111.00
2	H	153	SER	N-CA-CB	-5.35	102.47	110.50
2	H	228	TYR	CB-CG-CD2	-5.16	117.90	121.00
2	H	77(A)	ARG	N-CA-C	5.05	124.63	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	101	ARG	Sidechain
2	H	146	GLU	Peptide
2	H	148	TRP	Peptide
2	H	244	GLN	Peptide
2	H	245	PHE	Peptide

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	L	222	0	225	2	0
2	H	2083	0	2061	33	1
3	I	95	0	74	2	0
4	H	52	0	54	7	0
5	H	185	0	0	1	2
5	I	8	0	0	0	0
5	L	37	0	0	0	0
All	All	2682	0	2414	42	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 9.

All (42) 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:H:145:LYS:HD2	2:H:149(C):VAL:O	1.88	0.72
2:H:32:MET:HG3	2:H:40:LEU:HD23	1.77	0.66
2:H:68:ILE:HD12	2:H:112:VAL:HG11	1.79	0.63
2:H:164:GLU:HB2	2:H:167:VAL:HG13	1.83	0.60
2:H:240:LYS:O	2:H:244:GLN:HG3	2.03	0.59
2:H:145:LYS:HG3	2:H:149(B):ASN:HB3	1.85	0.58
2:H:146:GLU:O	2:H:149(B):ASN:HB2	2.05	0.57
4:H:1:GR4:H111	4:H:1:GR4:H281	1.87	0.57
2:H:191:CYS:HG	2:H:220:CYS:HG	0.90	0.55
1:L:14(G):LEU:HD21	2:H:202:LYS:HD3	1.88	0.55
2:H:126:ARG:NH1	2:H:232:PHE:HE2	2.05	0.55
2:H:49:ASP:OD2	2:H:111:PRO:HB3	2.07	0.54
2:H:35:ARG:O	2:H:38:GLN:HA	2.09	0.53
2:H:236:LYS:HB3	5:H:506:HOH:O	2.08	0.53
2:H:60(B):PRO:HG2	2:H:96:TRP:CE2	2.46	0.51
2:H:146:GLU:OE2	2:H:221:ARG:HD3	2.12	0.50
2:H:236:LYS:O	2:H:240:LYS:HG3	2.12	0.50
3:I:58:GLU:H	3:I:58:GLU:CD	2.16	0.49
2:H:148:TRP:H	2:H:148:TRP:HE3	1.61	0.48
2:H:75:ARG:CZ	2:H:77:GLU:HG2	2.43	0.48
2:H:34:PHE:CZ	2:H:38:GLN:HB3	2.49	0.48
2:H:148:TRP:O	4:H:1:GR4:H311	2.13	0.47
2:H:146:GLU:HA	2:H:147:THR:HG23	1.97	0.47
2:H:147:THR:HA	2:H:149(A):ALA:HA	1.97	0.45
4:H:1:GR4:H303	4:H:1:GR4:H192	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:36:LYS:HE2	2:H:62:ASN:O	2.16	0.45
2:H:148:TRP:HZ3	2:H:219:GLY:HA2	1.81	0.45
4:H:1:GR4:H281	4:H:1:GR4:C11	2.46	0.45
2:H:87:LYS:HB3	2:H:89:TYR:CE1	2.51	0.45
2:H:99:LEU:O	2:H:102:ASP:HB2	2.17	0.45
2:H:149(B):ASN:OD1	2:H:149(C):VAL:HG22	2.17	0.44
4:H:1:GR4:O36	4:H:1:GR4:H32	2.18	0.43
4:H:1:GR4:H183	4:H:1:GR4:H9	2.01	0.43
2:H:163:VAL:CG1	2:H:167:VAL:HG22	2.49	0.42
2:H:148:TRP:HH2	2:H:219:GLY:H	1.67	0.41
2:H:33:LEU:HB2	2:H:42:CYS:O	2.20	0.41
2:H:35:ARG:HD2	2:H:41:LEU:HD21	2.03	0.41
2:H:79:ILE:HG23	2:H:117:TYR:CD2	2.56	0.41
4:H:1:GR4:H191	4:H:1:GR4:C7	2.51	0.41
3:I:61:GLU:HA	3:I:64:LEU:HD12	2.02	0.41
2:H:91:HIS:CE1	2:H:101:ARG:HD3	2.56	0.40
1:L:3:LEU:HB2	1:L:9:LYS:HE3	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)
5:H:371:HOH:O	5:H:401:HOH:O[2_656]	2.14	0.06
2:H:175:ARG:NH2	5:H:523:HOH:O[2_556]	2.19	0.01

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	L	25/36 (69%)	22 (88%)	3 (12%)	0	100	100
2	H	256/259 (99%)	237 (93%)	17 (7%)	2 (1%)	24	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	I	7/10 (70%)	6 (86%)	1 (14%)	0	100	100
All	All	288/305 (94%)	265 (92%)	21 (7%)	2 (1%)	26	25

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	149(D)	GLY
2	H	149(C)	VAL

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	L	25/31 (81%)	25 (100%)	0	100	100
2	H	224/225 (100%)	206 (92%)	18 (8%)	15	15
3	I	9/9 (100%)	9 (100%)	0	100	100
All	All	258/265 (97%)	240 (93%)	18 (7%)	19	19

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	21	ASP
2	H	45	SER
2	H	46	LEU
2	H	60	LEU
2	H	107	LYS
2	H	127	GLU
2	H	129(B)	SER
2	H	129(C)	LEU
2	H	148	TRP
2	H	149(B)	ASN
2	H	149(E)	LYS
2	H	154	VAL
2	H	157	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	167	VAL
2	H	182	CYS
2	H	204(B)	ASN
2	H	241	VAL
2	H	244	GLN

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

Mol	Chain	Res	Type
2	H	62	ASN
2	H	204(B)	ASN
2	H	244	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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$
3	TYS	I	63	3	15,16,17	1.40	2 (13%)	16,22,24	1.18	1 (6%)

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	TYS	I	63	3	-	0/9/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	I	63	TYS	OH-S	-4.17	1.56	1.63
3	I	63	TYS	OH-CZ	2.52	1.46	1.42

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	63	TYS	CG-CB-CA	-2.59	108.37	114.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand 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	$\# Z > 2$
4	GR4	H	1	2	54,56,56	1.25	3 (5%)	70,89,89	1.64	18 (25%)

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
4	GR4	H	1	2	-	1/32/132/132	0/5/5/5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	1	GR4	O1-C28	-6.62	1.20	1.42
4	H	1	GR4	C10-C5	-2.40	1.52	1.56
4	H	1	GR4	C22-C20	2.48	1.57	1.54

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	H	1	GR4	O35-C23-C24	-3.55	116.94	123.23
4	H	1	GR4	C18-C13-C17	-2.78	107.70	113.12
4	H	1	GR4	C2-C3-C4	-2.67	108.95	112.84
4	H	1	GR4	C31-C36-C35	-2.39	107.12	113.02
4	H	1	GR4	C14-C8-C9	-2.23	114.19	119.27
4	H	1	GR4	C10-C9-C8	-2.12	109.12	112.52
4	H	1	GR4	C13-C14-C8	-2.11	108.82	110.60
4	H	1	GR4	O36-C37-O39	2.00	126.92	122.92
4	H	1	GR4	C18-C13-C12	2.25	112.89	109.09
4	H	1	GR4	C34-C35-C36	2.53	114.60	110.20
4	H	1	GR4	C32-O34-C29	2.63	121.14	116.97
4	H	1	GR4	O3-C36-C31	2.70	113.19	106.36
4	H	1	GR4	C6-C5-C4	2.84	116.52	112.75
4	H	1	GR4	C3-O36-C37	2.99	122.68	117.91
4	H	1	GR4	O6-C33-C32	3.20	117.03	110.02
4	H	1	GR4	C5-C4-C3	3.35	114.23	109.12
4	H	1	GR4	O34-C32-C33	3.53	118.09	107.48
4	H	1	GR4	O36-C3-C4	4.32	114.41	107.10

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	1	GR4	C29-O34-C32-C33

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1	GR4	7	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.