



wwPDB X-ray Structure Validation Summary Report ⓘ

May 30, 2020 – 02:18 pm BST

PDB ID : 1AWH
Title : NOVEL COVALENT THROMBIN INHIBITOR FROM PLANT EXTRACT
Authors : Jhoti, H.; Cleasby, A.; Wonacott, A.
Deposited on : 1997-10-02
Resolution : 3.00 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 ⓘ 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)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
buster-report : 1.1.7 (2018)
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

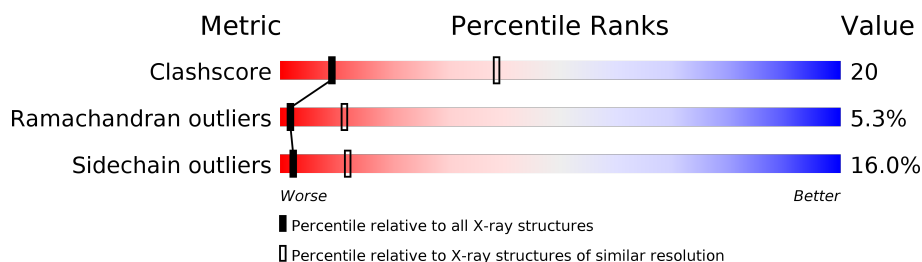
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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 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	36	
1	C	36	
2	B	259	
2	D	259	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4840 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.

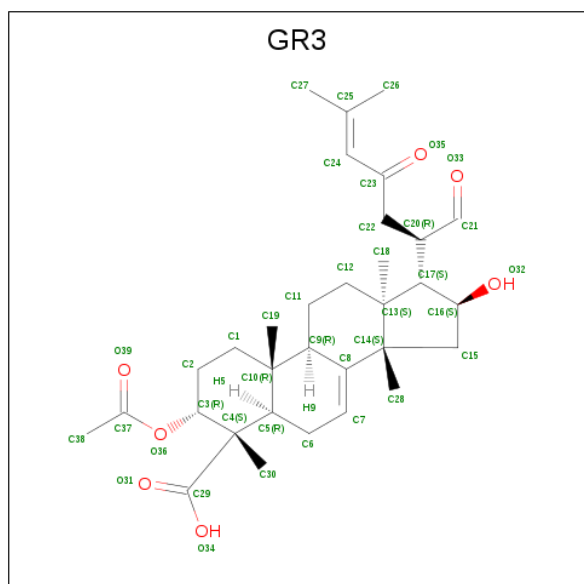
- Molecule 1 is a protein called ALPHA THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	36	Total	C	N	O	S	0	0	0
			287	177	48	61	1			
1	C	36	Total	C	N	O	S	0	0	0
			287	177	48	61	1			

- Molecule 2 is a protein called ALPHA THROMBIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	259	Total	C	N	O	S	0	0	0
			2093	1334	370	375	14			
2	D	259	Total	C	N	O	S	0	0	0
			2093	1334	370	375	14			

- Molecule 3 is 3-ACETOXY-17-(1-FORMYL-5-METHYL-3-OXO-HEX-4-ENYL)-16-HYDROXY-4,10,13,14-TETRAMETHYL-2,3,4,5,6,9,10,11,12,13,14,15,16,17-TETRADECAHYDRO-1H-CYCLOPENTA[A]PHENANTHRENE-4-CARBOXYLIC ACID (three-letter code: GR3) (formula: C₃₂H₄₆O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			39	32	7		
3	D	1	Total	C	O	0	0
			39	32	7		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	O	0	0
			1	1		
4	D	1	Total	O	0	0
			1	1		

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. 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: ALPHA THROMBIN

Chain A: 



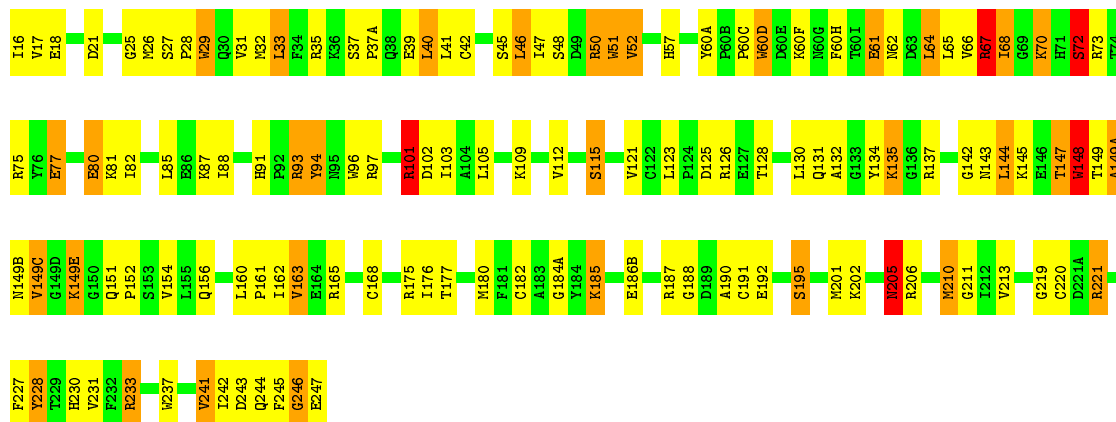
• Molecule 1: ALPHA THROMBIN

Chain C: 



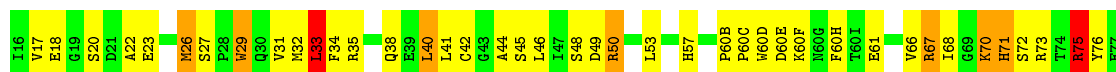
• Molecule 2: ALPHA THROMBIN

Chain B: 



• Molecule 2: ALPHA THROMBIN

Chain D: 



R77A	E146	G211	R126	R77A
R81	T147	E217	E127	R81
I82	W148	E217	E127	I82
L85	T149	D221A	T128	L85
E86	A149A	R221	A129	E86
R87	M149B	Y225	A129A	R87
I88	K149E	G226	S129B	I88
R91	G150	F227	L129C	R91
P92	Q151	H230	L130	P92
R93	P152	V231	Q131	R93
Y94	S153	F232	Y134	Y94
N95	V157	R233	K135	N95
R97	P161	L234	G136	R97
L99	I162	K235	R137	L99
D100	V163	K236	V138	D100
R101	E164	V237	T139	R101
D102	R165	I238	G140	D102
I103	P166	Q239	W141	I103
M106	V167	K240		M106
R107	C168	D243		R107
L108	K169	Q244		L108
K109	R173	F245		K109
K110	I174	G246		K110
P111	I176	E247		P111
V112	M180			V112
A113	F181			A113
F114	C182			F114
S115	K185			S115
D116	P186			D116
Y117	D186A			Y117
I118	E186B			I118
H119	R187			H119
P120	A190			P120
V121	C191			V121
R126	S195			R126
E127	P198			E127
T128	F199			T128
A129	V200			A129
S129B	W201			S129B
L129C	K202			L129C
L130	F204A			L130
Q131	N204B			Q131
Y134	N205			Y134
K135	R206			K135
G136	W207			G136
R137	Y208			R137
V138	Q209			V138
T139	M210			T139
G140				G140
W141				W141

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, α , β , γ	65.60 Å 102.80 Å 119.70 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	10.00 – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) (10.00-3.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.202 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4840	wwPDB-VP
Average B, all atoms (Å ²)	25.0	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: GR3

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	1.05	0/290	1.89	3/384 (0.8%)
1	C	1.10	0/290	1.83	6/384 (1.6%)
2	B	1.01	5/2148 (0.2%)	1.77	26/2903 (0.9%)
2	D	1.01	5/2148 (0.2%)	1.79	32/2903 (1.1%)
All	All	1.02	10/4876 (0.2%)	1.79	67/6574 (1.0%)

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	7
1	C	0	4
2	B	0	9
2	D	0	9
All	All	0	29

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	195	SER	CB-OG	-9.10	1.30	1.42
2	D	195	SER	CB-OG	-8.60	1.31	1.42
2	B	148	TRP	CD2-CE2	6.58	1.49	1.41
2	D	96	TRP	CD2-CE2	6.03	1.48	1.41
2	B	237	TRP	CD2-CE2	5.95	1.48	1.41

The worst 5 of 67 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	134	TYR	CB-CG-CD2	-13.41	112.96	121.00
2	B	134	TYR	CB-CG-CD1	12.84	128.70	121.00
1	A	15	ARG	NE-CZ-NH1	12.81	126.70	120.30
2	D	67	ARG	NE-CZ-NH1	8.03	124.31	120.30
2	B	221	ARG	NE-CZ-NH1	7.66	124.13	120.30

There are no chirality outliers.

5 of 29 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1(D)	GLY	Peptide
1	A	1(G)	PHE	Peptide
1	A	1(H)	THR	Peptide
1	A	14(J)	TYR	Sidechain,Peptide
1	A	14(L)	ASP	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	A	287	0	278	22	0
1	C	287	0	278	19	0
2	B	2093	0	2063	80	0
2	D	2093	0	2063	86	0
3	B	39	0	44	5	0
3	D	39	0	44	4	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
All	All	4840	0	4770	191	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.

The worst 5 of 191 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:B:60(F):LYS:HG3	2:B:60(H):PHE:HE1	1.33	0.94

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:14(G):LEU:HB3	1:A:15:ARG:NE	1.90	0.87
2:D:230:HIS:CD2	2:D:233:ARG:HG2	2.16	0.80
2:D:200:VAL:HG12	2:D:209:GLN:HA	1.62	0.80
2:D:190:ALA:HB3	3:D:1:GR3:H262	1.64	0.80

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	34/36 (94%)	16 (47%)	10 (29%)	8 (24%)	0	0
1	C	34/36 (94%)	20 (59%)	9 (26%)	5 (15%)	0	1
2	B	257/259 (99%)	220 (86%)	32 (12%)	5 (2%)	8	36
2	D	257/259 (99%)	210 (82%)	34 (13%)	13 (5%)	2	12
All	All	582/590 (99%)	466 (80%)	85 (15%)	31 (5%)	2	11

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1(B)	ALA
1	A	14(K)	ILE
1	A	14(L)	ASP
2	B	62	ASN
2	B	93	ARG

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	31/31 (100%)	23 (74%)	8 (26%)	0	2
1	C	31/31 (100%)	22 (71%)	9 (29%)	0	2
2	B	225/225 (100%)	192 (85%)	33 (15%)	3	15
2	D	225/225 (100%)	193 (86%)	32 (14%)	3	16
All	All	512/512 (100%)	430 (84%)	82 (16%)	2	12

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	241	VAL
1	C	12	LEU
2	D	210	MET
2	B	243	ASP
1	C	1(H)	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	131	GLN
2	D	230	HIS
2	D	151	GLN
2	D	71	HIS
2	D	143	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

2 ligands 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$
3	GR3	D	1	2	40,42,42	1.08	3 (7%)	56,68,68	1.59	8 (14%)
3	GR3	B	1	2	40,42,42	1.08	4 (10%)	56,68,68	1.40	9 (16%)

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	GR3	D	1	2	-	4/18/99/99	0/4/4/4
3	GR3	B	1	2	-	5/18/99/99	0/4/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1	GR3	C4-C29	2.81	1.55	1.50
3	D	1	GR3	C14-C8	-2.79	1.48	1.53
3	B	1	GR3	C4-C3	-2.30	1.51	1.56
3	B	1	GR3	C10-C5	-2.15	1.53	1.56
3	D	1	GR3	C4-C29	2.11	1.54	1.50

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	1	GR3	C15-C14-C8	4.80	120.99	117.19
3	D	1	GR3	C28-C14-C8	-4.63	101.65	107.83
3	B	1	GR3	C22-C20-C21	-4.40	98.83	109.51
3	D	1	GR3	C3-O36-C37	4.28	124.42	117.86
3	D	1	GR3	C22-C20-C21	-3.77	100.37	109.51

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

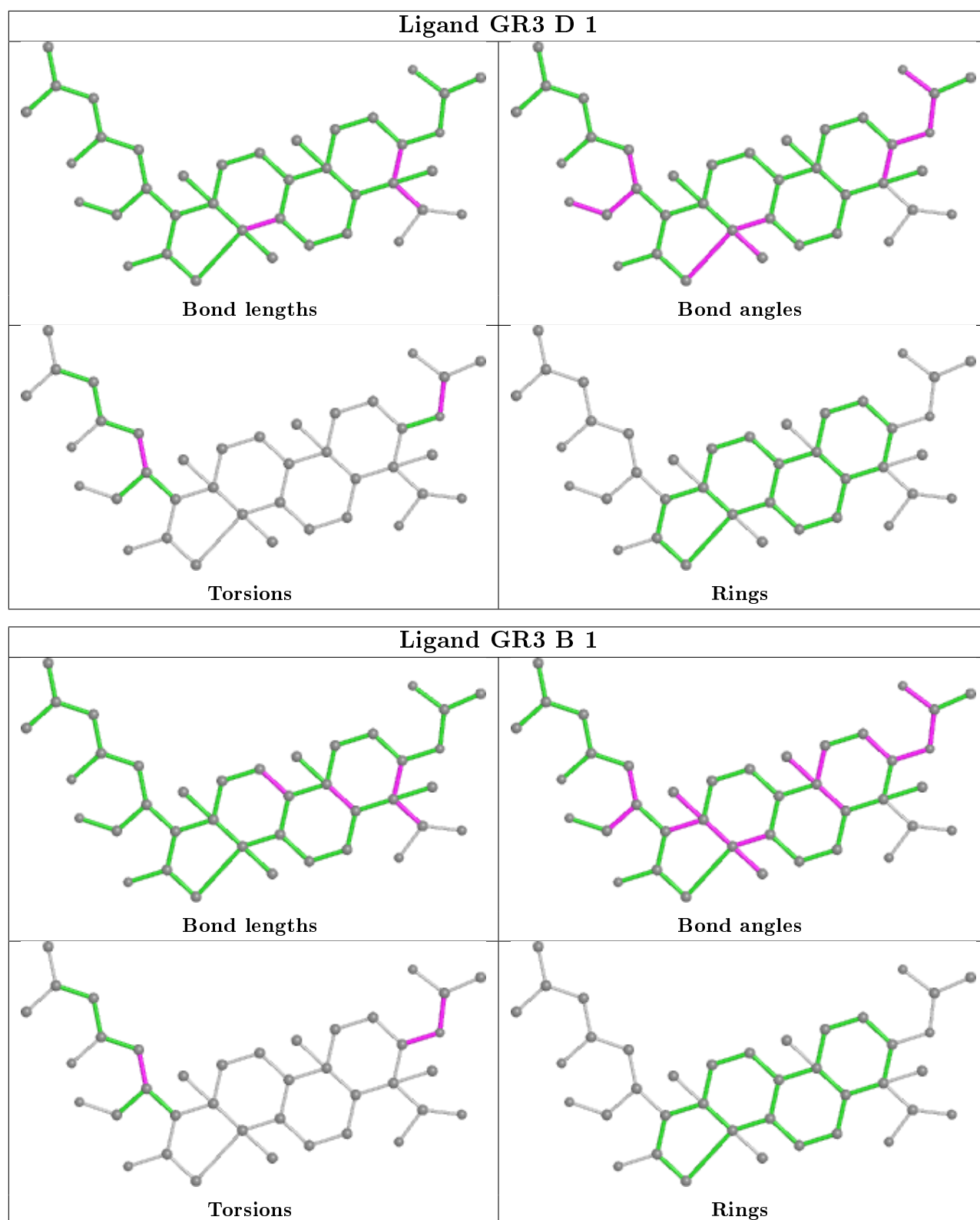
Mol	Chain	Res	Type	Atoms
3	D	1	GR3	C17-C20-C22-C23
3	D	1	GR3	C21-C20-C22-C23
3	B	1	GR3	C17-C20-C22-C23
3	B	1	GR3	C21-C20-C22-C23
3	B	1	GR3	C38-C37-O36-C3

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1	GR3	4	0
3	B	1	GR3	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 than 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.



5.7 Other polymers [i](#)

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 is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.