wwPDB X-ray Structure Validation Summary Report (i)

Mar 8, 2026 – 02:09 PM UTC

PDB ID	:	6TFI / pdb_00006tfi
Title	:	PXR IN COMPLEX WITH THROMBIN INHIBITOR COMPOUND 17
Authors	:	Hillig, R.C.; Puetter, V.
Deposited on	:	2019-11-14
Resolution	:	1.85 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 (i)) were used in the production of this report:

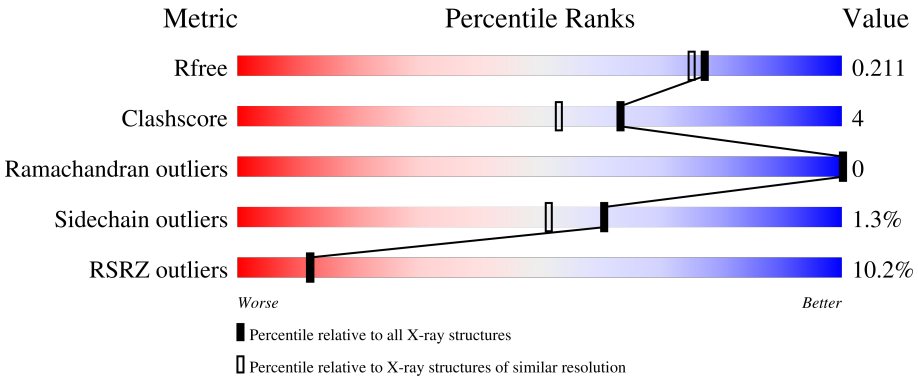
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 1.85 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}	180053	3428 (1.86-1.86)
Clashscore	190562	3579 (1.86-1.86)
Ramachandran outliers	187476	3553 (1.86-1.86)
Sidechain outliers	187428	3553 (1.86-1.86)
RSRZ outliers	180081	3429 (1.86-1.86)

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	344	8% 82% 5% 13%
1	B	344	10% 75% 11% 13%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5425 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 Nuclear receptor subfamily 1 group I member 2,Nuclear receptor coactivator 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	0	13	0
			2493	1598	429	444	22			
1	B	299	Total	C	N	O	S	0	9	0
			2483	1595	425	440	23			

There are 38 discrepancies between the modelled and reference sequences:

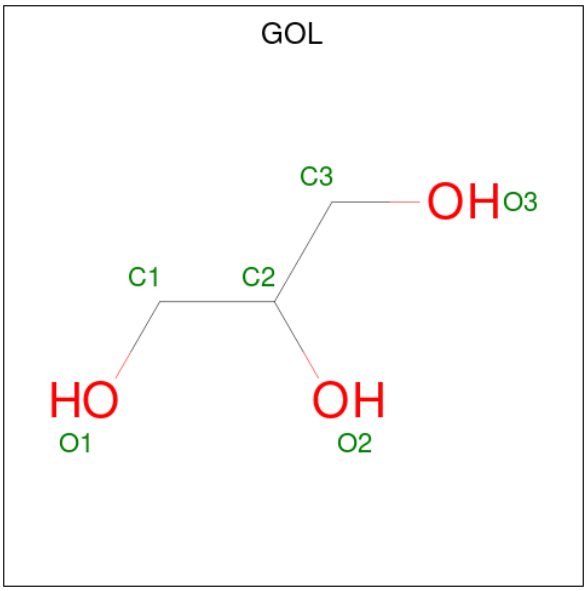
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	MET	-	initiating methionine	UNP O75469
A	120	LYS	-	expression tag	UNP O75469
A	121	LYS	-	expression tag	UNP O75469
A	122	GLY	-	expression tag	UNP O75469
A	123	HIS	-	expression tag	UNP O75469
A	124	HIS	-	expression tag	UNP O75469
A	125	HIS	-	expression tag	UNP O75469
A	126	HIS	-	expression tag	UNP O75469
A	127	HIS	-	expression tag	UNP O75469
A	128	HIS	-	expression tag	UNP O75469
A	129	GLY	LYS	engineered mutation	UNP O75469
A	435	GLY	-	linker	UNP O75469
A	436	GLY	-	linker	UNP O75469
A	437	SER	-	linker	UNP O75469
A	438	GLY	-	linker	UNP O75469
A	439	GLY	-	linker	UNP O75469
A	440	SER	-	linker	UNP O75469
A	441	SER	-	linker	UNP O75469
A	442	HIS	-	linker	UNP O75469
B	119	MET	-	initiating methionine	UNP O75469
B	120	LYS	-	expression tag	UNP O75469
B	121	LYS	-	expression tag	UNP O75469
B	122	GLY	-	expression tag	UNP O75469
B	123	HIS	-	expression tag	UNP O75469

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	124	HIS	-	expression tag	UNP O75469
B	125	HIS	-	expression tag	UNP O75469
B	126	HIS	-	expression tag	UNP O75469
B	127	HIS	-	expression tag	UNP O75469
B	128	HIS	-	expression tag	UNP O75469
B	129	GLY	LYS	engineered mutation	UNP O75469
B	435	GLY	-	linker	UNP O75469
B	436	GLY	-	linker	UNP O75469
B	437	SER	-	linker	UNP O75469
B	438	GLY	-	linker	UNP O75469
B	439	GLY	-	linker	UNP O75469
B	440	SER	-	linker	UNP O75469
B	441	SER	-	linker	UNP O75469
B	442	HIS	-	linker	UNP O75469

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



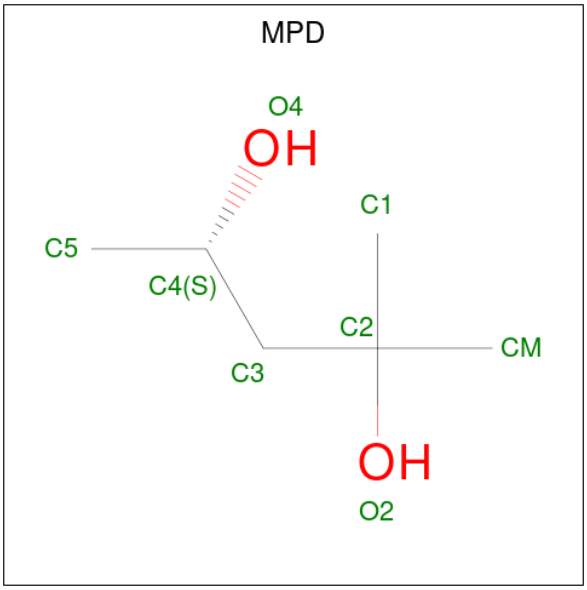
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0
2	B	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

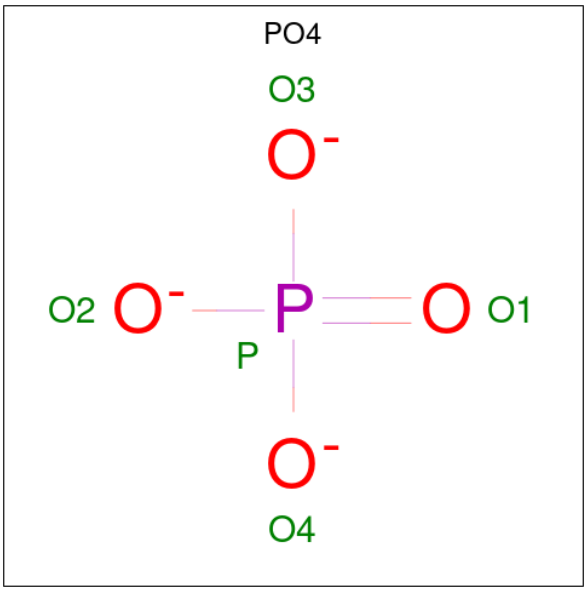
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C O 6 3 3	0	0

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: $C_6H_{14}O_2$).



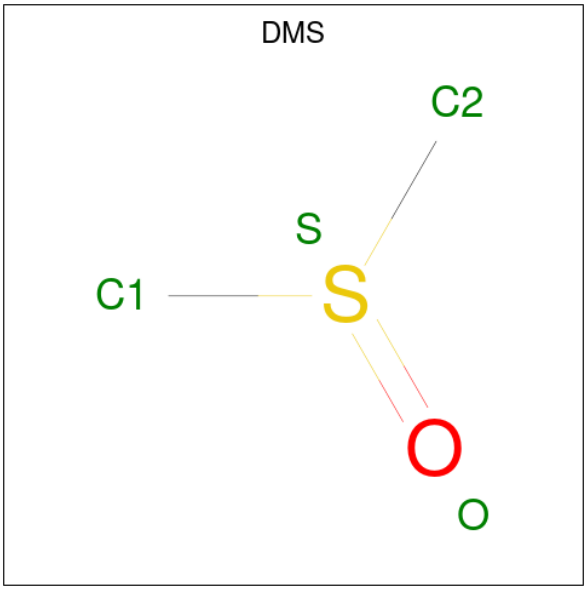
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 8 6 2	0	0
3	B	1	Total C O 8 6 2	0	0

- Molecule 4 is PHOSPHATE ION (CCD ID: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0

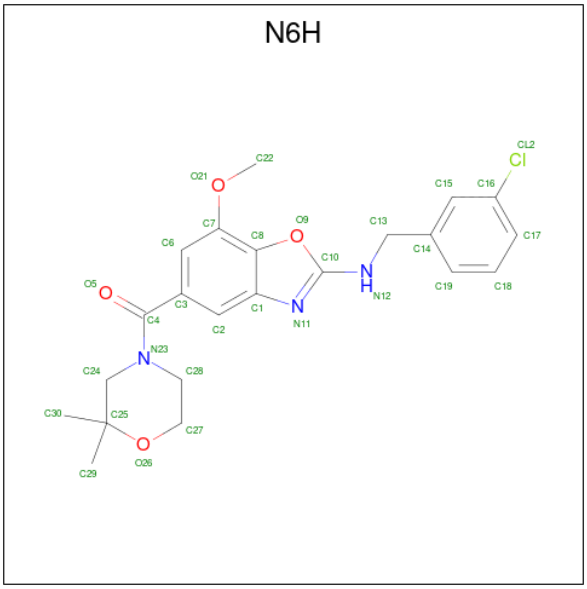
- Molecule 5 is DIMETHYL SULFOXIDE (CCD ID: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C O S 4 2 1 1	0	0
5	B	1	Total C O S 4 2 1 1	0	0

- Molecule 6 is [2-[(3-chlorophenyl)methylamino]-7-methoxy-1,3-benzoxazol-5-yl]-(2,2-dimeth

ylmorpholin-4-yl)methanone (CCD ID: N6H) (formula: $C_{22}H_{24}ClN_3O_4$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	Cl	N	O	0	0
			30	22	1	3	4		

- Molecule 7 is water.

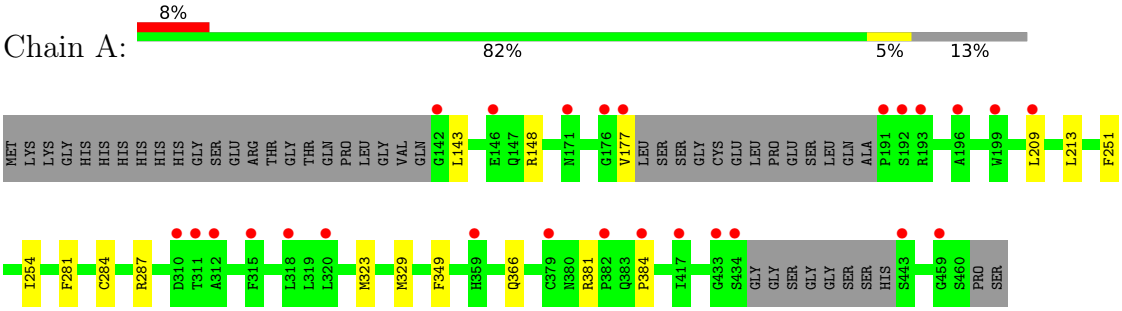
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	189	Total	O	0	0
			189	189		
7	B	171	Total	O	0	0
			171	171		

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.

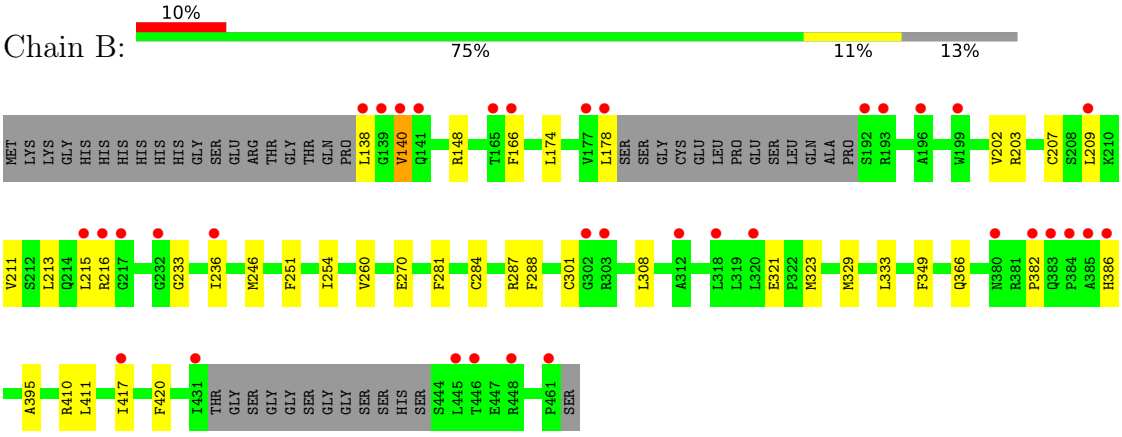
• Molecule 1: Nuclear receptor subfamily 1 group I member 2,Nuclear receptor coactivator 1

Chain A:



• Molecule 1: Nuclear receptor subfamily 1 group I member 2,Nuclear receptor coactivator 1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.36Å 88.96Å 105.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.48 – 1.85 44.48 – 1.85	Depositor EDS
% Data completeness (in resolution range)	97.9 (44.48-1.85) 98.0 (44.48-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	1.63 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.172 , 0.199 0.181 , 0.211	Depositor DCC
R_{free} test set	2101 reflections (3.09%)	wwPDB-VP
Wilson B-factor (Å^2)	32.5	Xtriage
Anisotropy	0.138	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 44.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5425	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, GOL, DMS, MPD, N6H

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	0.95	0/2575	1.30	1/3459 (0.0%)
1	B	0.95	0/2551	1.31	1/3432 (0.0%)
All	All	0.95	0/5126	1.31	2/6891 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	381	ARG	CB-CA-C	5.49	117.11	108.88
1	B	382	PRO	N-CA-C	5.03	121.11	114.68

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	2493	0	2532	12	0
1	B	2483	0	2510	30	0
2	A	12	0	16	1	0
2	B	18	0	24	0	0
3	A	8	0	14	2	0
3	B	8	0	14	2	0
4	A	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	B	8	0	12	0	0
6	B	30	0	0	4	0
7	A	189	0	0	0	0
7	B	171	0	0	0	0
All	All	5425	0	5122	41	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 4.

The worst 5 of 41 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:281:PHE:CZ	1:B:323:MET:HE1	2.13	0.83
1:B:329:MET:HE1	3:B:502:MPD:H12	1.63	0.80
1:A:281:PHE:CZ	1:A:323:MET:HE1	2.20	0.74
1:B:270[B]:GLU:H	1:B:270[B]:GLU:CD	1.99	0.70
1:B:140:VAL:HG22	1:B:148:ARG:HG2	1.75	0.68

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	305/344 (89%)	301 (99%)	4 (1%)	0	100	100
1	B	302/344 (88%)	296 (98%)	6 (2%)	0	100	100
All	All	607/688 (88%)	597 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	279/302 (92%)	278 (100%)	1 (0%)	84 82
1	B	276/302 (91%)	270 (98%)	6 (2%)	45 32
All	All	555/604 (92%)	548 (99%)	7 (1%)	61 51

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	178	LEU
1	B	216	ARG
1	B	308	LEU
1	B	260	VAL
1	B	140	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	406	GLN
1	B	407	HIS
1	B	418	HIS
1	A	407	HIS
1	B	168	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

11 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	$\#\|Z\| > 2$
5	DMS	B	501	-	3,3,3	0.20	0	3,3,3	0.10	0
2	GOL	A	502	-	5,5,5	0.08	0	5,5,5	0.20	0
4	PO4	A	504	-	4,4,4	0.79	0	6,6,6	0.45	0
2	GOL	B	506	-	5,5,5	0.10	0	5,5,5	0.27	0
6	N6H	B	503	-	32,33,33	1.00	3 (9%)	41,48,48	1.37	5 (12%)
5	DMS	B	504	-	3,3,3	0.17	0	3,3,3	0.14	0
2	GOL	A	501	-	5,5,5	0.10	0	5,5,5	0.29	0
3	MPD	B	502	-	7,7,7	0.12	0	9,10,10	0.37	0
3	MPD	A	503	-	7,7,7	0.13	0	9,10,10	0.33	0
2	GOL	B	507	-	5,5,5	0.09	0	5,5,5	0.22	0
2	GOL	B	505	-	5,5,5	0.09	0	5,5,5	0.25	0

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	GOL	A	502	-	-	2/4/4/4	-
2	GOL	B	506	-	-	0/4/4/4	-
6	N6H	B	503	-	-	1/14/27/27	0/4/4/4
2	GOL	A	501	-	-	2/4/4/4	-
3	MPD	B	502	-	-	0/5/5/5	-
3	MPD	A	503	-	-	0/5/5/5	-
2	GOL	B	507	-	-	2/4/4/4	-
2	GOL	B	505	-	-	2/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	B	503	N6H	C1-N11	-2.63	1.34	1.39
6	B	503	N6H	C8-C1	2.20	1.40	1.38
6	B	503	N6H	C3-C4	2.04	1.53	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	503	N6H	O9-C8-C7	4.70	131.03	124.75
6	B	503	N6H	C14-C13-N12	-2.67	107.45	113.07
6	B	503	N6H	C3-C4-N23	2.40	121.66	118.66
6	B	503	N6H	C25-C24-N23	-2.23	108.37	113.92
6	B	503	N6H	O9-C10-N11	-2.08	111.83	115.57

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	GOL	O1-C1-C2-C3
2	B	507	GOL	O1-C1-C2-C3
2	A	502	GOL	O1-C1-C2-C3
2	B	505	GOL	C1-C2-C3-O3
2	A	501	GOL	O1-C1-C2-O2

There are no ring outliers.

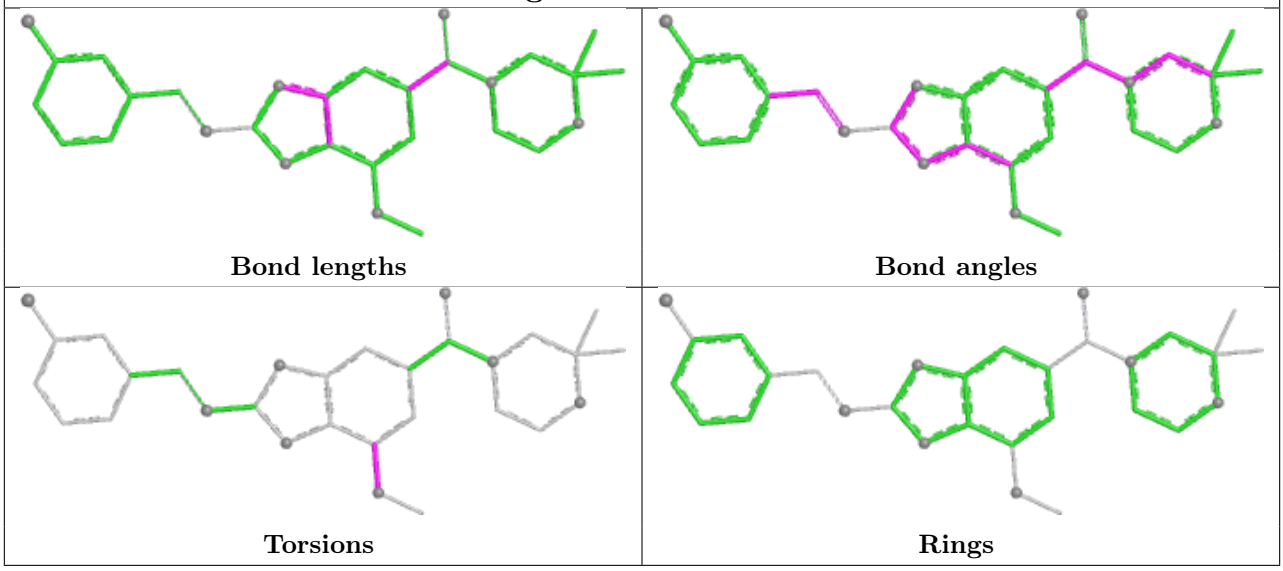
4 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	503	N6H	4	0
2	A	501	GOL	1	0
3	B	502	MPD	2	0
3	A	503	MPD	2	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 N6H B 503



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

6.1 Protein, DNA and RNA chains

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^{th} 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	298/344 (86%)	0.40	26 (8%)	16	16	12, 38, 91, 142	13 (4%)
1	B	299/344 (86%)	0.55	35 (11%)	9	8	14, 40, 82, 128	9 (3%)
All	All	597/688 (86%)	0.48	61 (10%)	12	12	12, 39, 85, 142	22 (3%)

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	384	PRO	6.5
1	A	177	VAL	6.3
1	B	178	LEU	6.0
1	A	209	LEU	4.8
1	B	166	PHE	4.5

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

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

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

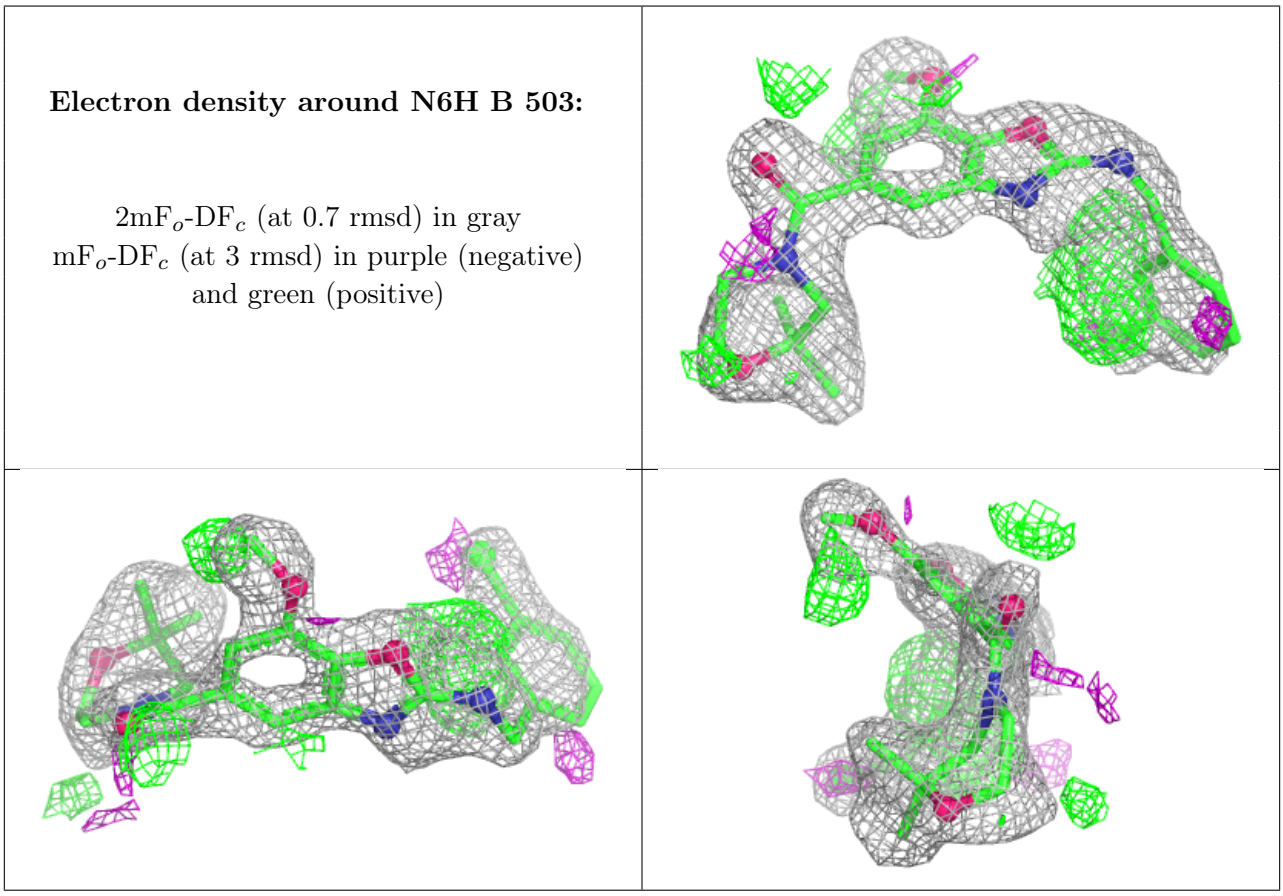
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	GOL	A	502	6/6	0.78	0.28	98,102,106,110	0
3	MPD	B	502	8/8	0.78	0.31	79,88,95,99	0
6	N6H	B	503	30/30	0.78	0.21	41,48,56,62	30
2	GOL	B	507	6/6	0.82	0.21	62,77,79,84	0
3	MPD	A	503	8/8	0.83	0.26	68,77,84,92	0
2	GOL	A	501	6/6	0.85	0.19	66,74,76,79	0
2	GOL	B	506	6/6	0.85	0.20	64,76,82,82	0
2	GOL	B	505	6/6	0.89	0.14	81,83,85,87	0
5	DMS	B	504	4/4	0.92	0.13	54,56,67,70	0
5	DMS	B	501	4/4	0.92	0.17	86,87,91,97	0
4	PO4	A	504	5/5	0.95	0.06	31,37,49,50	5

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 N6H B 503:

$2mF_o$-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

There are no such residues in this entry.