



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 2, 2024 – 04:00 pm GMT

PDB ID : 5JJM
Title : Crystal Structure of Homodimeric Androgen Receptor Ligand-Binding Domain bound to DHT and LxxLL peptide
Authors : Estebanez-Perpina, E.; Fuentes-Prior, P.
Deposited on : 2016-04-24
Resolution : 2.15 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

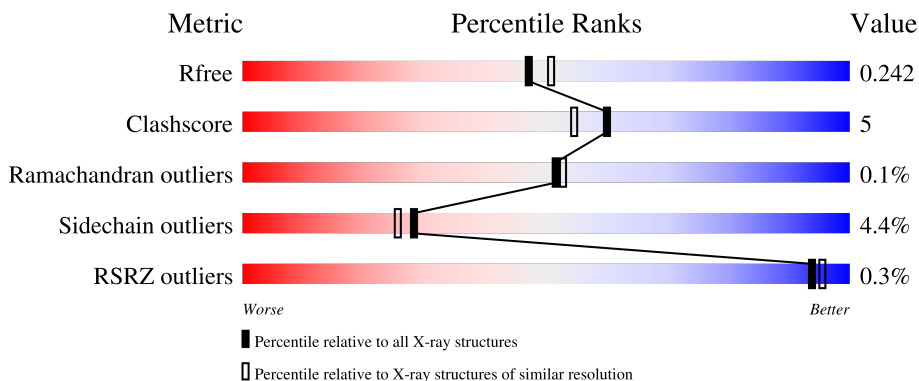
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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.




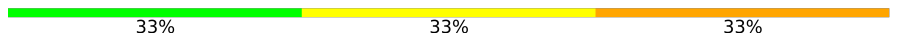





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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 $\leq 5\%$. 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	252	 84% 16%
1	C	252	 80% 18% ..
1	D	252	 83% 14% ..
2	F	3	 67% 33%
2	K	3	 33% 67% 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	3	
2	M	3	
3	B	252	
4	G	8	
4	H	8	
4	I	8	
4	J	8	

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
2	FME	L	503	-	X	-	-
9	GOL	A	1005	-	X	-	-

2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 8848 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 Androgen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	252	Total 2073	C 1340	N 356	O 358	S 19	0	3	0
1	C	250	Total 2052	C 1327	N 353	O 352	S 20	0	3	0
1	D	249	Total 2041	C 1320	N 351	O 350	S 20	0	2	0

- Molecule 2 is a protein called Unknown peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	M	3	Total 24	C 13	N 3	O 7	S 1	0	0	0
2	F	3	Total 24	C 13	N 3	O 7	S 1	0	0	0
2	K	3	Total 24	C 13	N 3	O 7	S 1	0	0	0
2	L	3	Total 24	C 13	N 3	O 7	S 1	0	0	0

- Molecule 3 is a protein called Androgen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	B	251	Total 2060	C 1329	N 354	O 356	S 21	0	3	0

- Molecule 4 is a protein called uba3-derived peptide.

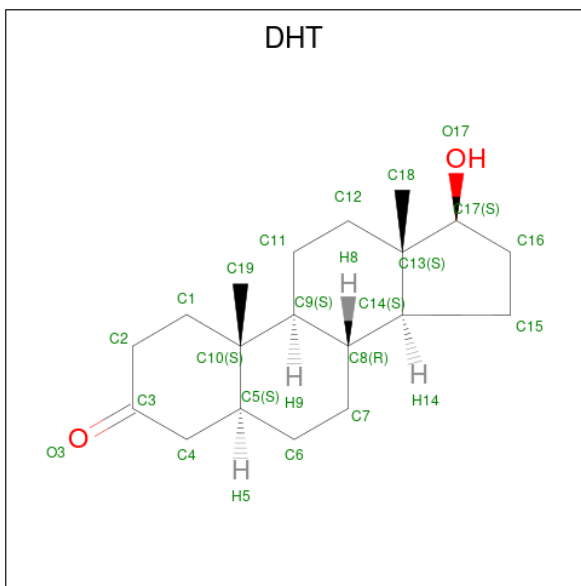
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	G	7	Total 58	C 39	N 8	O 11	0	0	0
4	H	8	Total 65	C 43	N 9	O 13	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace	
4	I	8	Total	C	N	O	0	0	0
			65	43	9	13			
4	J	8	Total	C	N	O	0	0	0
			65	43	9	13			

- Molecule 5 is 5-ALPHA-DIHYDROTESTOSTERONE (three-letter code: DHT) (formula: $C_{19}H_{30}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			21	19	2		
5	B	1	Total	C	O	0	0
			21	19	2		
5	C	1	Total	C	O	0	0
			21	19	2		
5	D	1	Total	C	O	0	0
			21	19	2		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

Continued on next page...

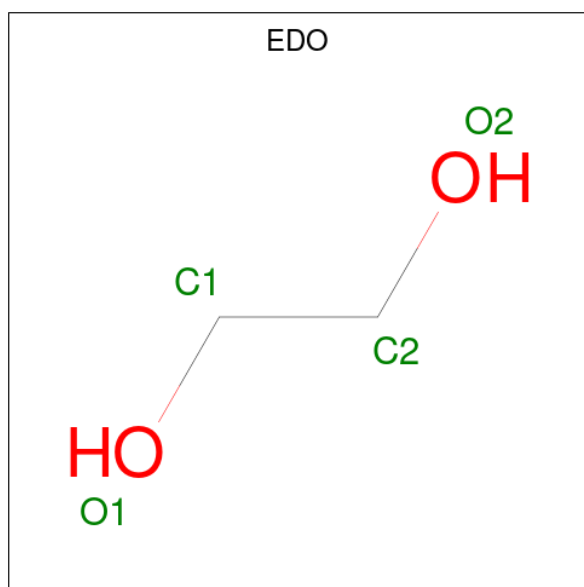
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total	Zn	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Cl	0	0
			1	1		
7	B	1	Total	Cl	0	0
			1	1		
7	D	1	Total	Cl	0	0
			1	1		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	C	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



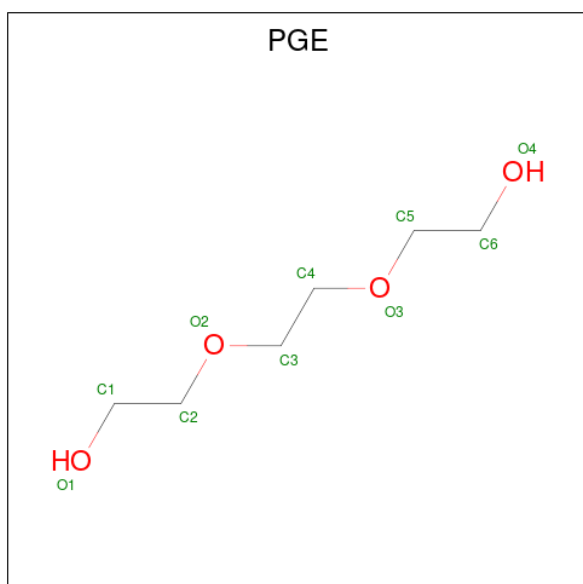
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			6	3	3		
9	C	1	Total	C	O	0	0
			6	3	3		

- Molecule 10 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



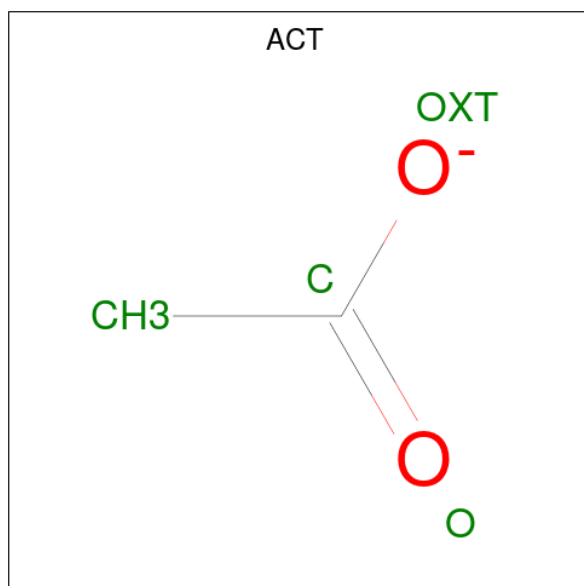
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	1	Total O S 5 4 1	0	0
10	B	1	Total O S 5 4 1	0	0
10	C	1	Total O S 5 4 1	0	0
10	C	1	Total O S 5 4 1	0	0
10	D	1	Total O S 5 4 1	0	0

- Molecule 11 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C₆H₁₄O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	C	1	Total	C	O	0	0
			10	6	4		

- Molecule 12 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
12	D	1	Total	C	O	0	0
			4	2	2		

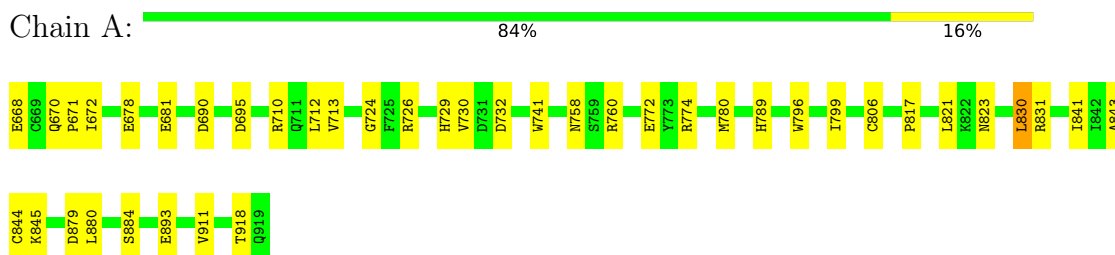
- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	26	Total	O	0	0
			26	26		
13	B	21	Total	O	0	0
			21	21		
13	C	28	Total	O	0	0
			28	28		
13	D	24	Total	O	0	0
			24	24		

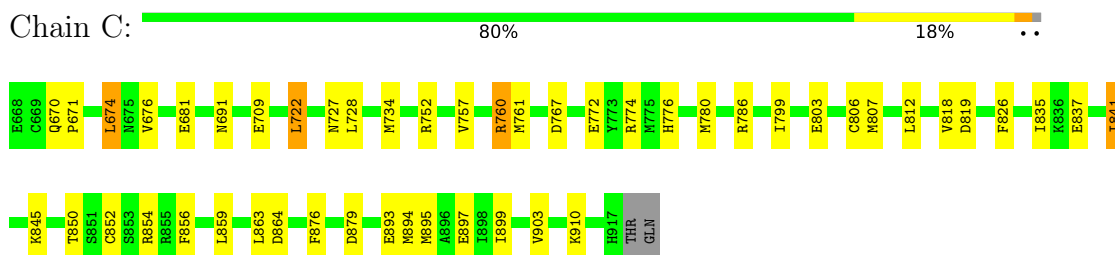
3 Residue-property plots [i](#)

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.

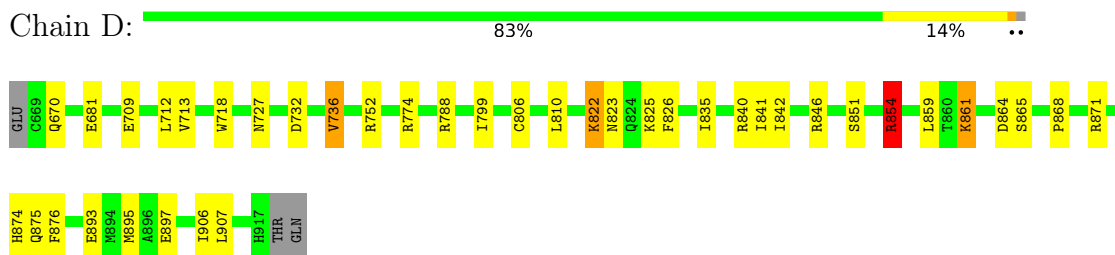
- Molecule 1: Androgen receptor



- Molecule 1: Androgen receptor



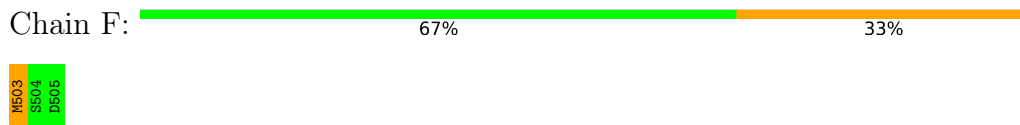
- Molecule 1: Androgen receptor



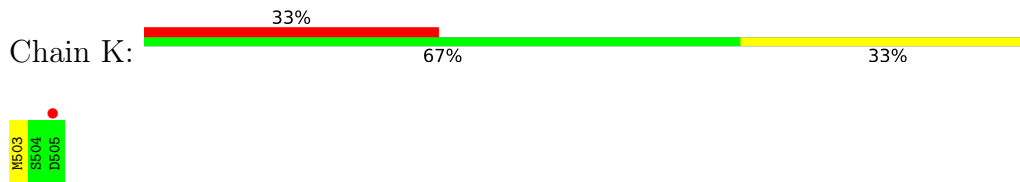
- Molecule 2: Unknown peptide



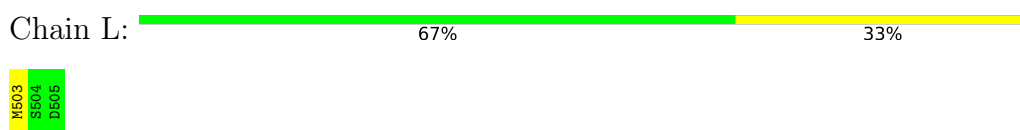
- Molecule 2: Unknown peptide



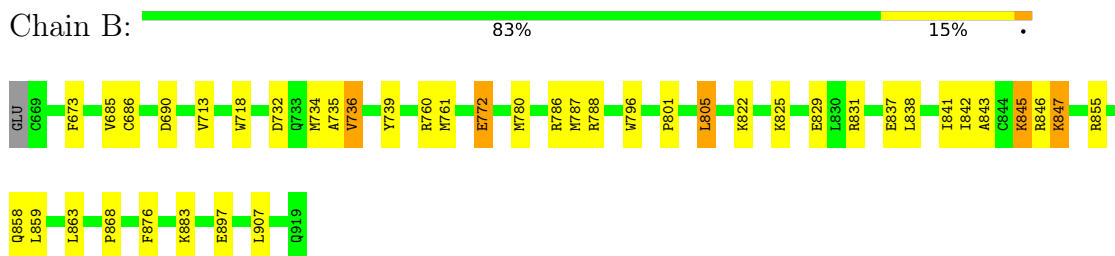
- Molecule 2: Unknown peptide



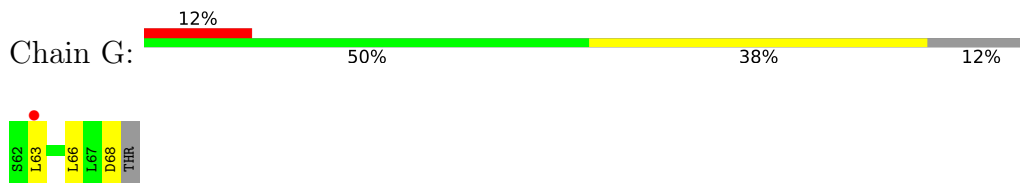
- Molecule 2: Unknown peptide



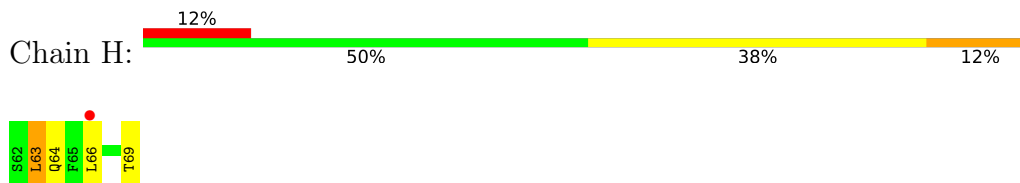
- Molecule 3: Androgen receptor



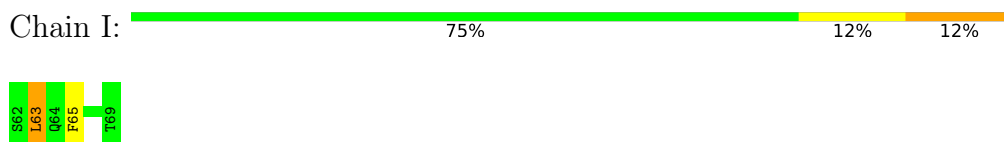
- Molecule 4: uba3-derived peptide



- Molecule 4: uba3-derived peptide

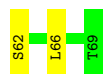


- Molecule 4: uba3-derived peptide



- Molecule 4: uba3-derived peptide

Chain J:  75% 25%



4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	91.09Å 91.01Å 157.23Å 90.00° 90.07° 90.00°	Depositor
Resolution (Å)	157.23 – 2.15 64.38 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.8 (157.23-2.15) 99.4 (64.38-2.15)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 2.14Å)	Xtrriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.201 , 0.243 0.200 , 0.242	Depositor DCC
R_{free} test set	3444 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	35.4	Xtrriage
Anisotropy	0.233	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 29.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.348 for k,h,-l 0.357 for -k,-h,-l 0.317 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8848	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.68% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: CL, PGE, ACT, SNC, GOL, DHT, FME, ZN, SO4, EDO

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	1.30	4/2135 (0.2%)	1.22	9/2884 (0.3%)
1	C	1.30	5/2121 (0.2%)	1.23	14/2863 (0.5%)
1	D	1.30	4/2102 (0.2%)	1.18	11/2838 (0.4%)
2	F	1.19	0/13	1.77	0/16
2	K	1.77	0/13	1.26	0/16
2	L	1.53	0/13	1.33	0/16
2	M	0.70	0/13	0.88	0/16
3	B	1.22	4/2117 (0.2%)	1.18	9/2855 (0.3%)
4	G	1.01	0/58	1.24	0/77
4	H	0.96	0/65	1.02	0/87
4	I	0.91	0/65	1.01	0/87
4	J	0.96	0/65	1.25	0/87
All	All	1.27	17/8780 (0.2%)	1.20	43/11842 (0.4%)

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
3	B	0	1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	852	CYS	C-N	14.04	1.66	1.34
1	A	741	TRP	CZ3-CH2	6.89	1.51	1.40
1	C	803	GLU	CD-OE1	6.27	1.32	1.25
1	A	681	GLU	CD-OE2	6.25	1.32	1.25
1	D	876	PHE	CG-CD2	5.88	1.47	1.38

The worst 5 of 43 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	805	LEU	CB-CG-CD1	11.01	129.72	111.00
1	C	864	ASP	CB-CG-OD1	9.96	127.27	118.30
1	A	831	ARG	NE-CZ-NH1	9.48	125.04	120.30
1	D	752	ARG	NE-CZ-NH1	9.27	124.93	120.30
1	A	726	ARG	NE-CZ-NH1	8.73	124.66	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	685	VAL	Mainchain

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	2073	0	2089	21	1
1	C	2052	0	2062	21	0
1	D	2041	0	2060	20	0
2	F	24	0	18	1	0
2	K	24	0	18	0	0
2	L	24	0	18	0	0
2	M	24	0	18	4	0
3	B	2060	0	2076	18	0
4	G	58	0	58	1	0
4	H	65	0	65	3	0
4	I	65	0	65	1	0
4	J	65	0	65	2	0
5	A	21	0	30	0	0
5	B	21	0	30	0	0
5	C	21	0	30	0	0
5	D	21	0	30	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	1	0	0	0	0
7	B	1	0	0	0	0
7	D	1	0	0	1	0
8	A	4	0	6	2	0
8	B	12	0	18	0	0
8	C	4	0	6	0	0
8	D	12	0	18	0	0
9	A	6	0	8	1	0
9	C	6	0	8	1	0
10	A	5	0	0	0	0
10	B	5	0	0	0	0
10	C	10	0	0	0	0
10	D	5	0	0	0	0
11	C	10	0	14	1	0
12	D	4	0	3	0	0
13	A	26	0	0	1	1
13	B	21	0	0	0	0
13	C	28	0	0	2	0
13	D	24	0	0	0	0
All	All	8848	0	8813	81	1

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 5.

The worst 5 of 81 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:806:CYS:SG	1:A:841:ILE:HD12	1.59	1.43
1:A:806:CYS:SG	1:A:841:ILE:CD1	2.28	1.21
1:A:843:ALA:HB2	2:M:503:FME:CE	2.12	0.80
1:A:710:ARG:HH12	8:A:1004:EDO:H22	1.50	0.75
1:A:772:GLU:HG3	1:C:676:VAL:HG21	1.73	0.68

All (1) 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)
1:A:760:ARG:NH1	13:A:1120:HOH:O[2_455]	2.13	0.07

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	253/252 (100%)	243 (96%)	9 (4%)	1 (0%)	34	29
1	C	251/252 (100%)	244 (97%)	7 (3%)	0	100	100
1	D	249/252 (99%)	240 (96%)	9 (4%)	0	100	100
2	F	1/3 (33%)	1 (100%)	0	0	100	100
2	K	1/3 (33%)	1 (100%)	0	0	100	100
2	L	1/3 (33%)	1 (100%)	0	0	100	100
2	M	1/3 (33%)	1 (100%)	0	0	100	100
3	B	251/252 (100%)	244 (97%)	7 (3%)	0	100	100
4	G	5/8 (62%)	5 (100%)	0	0	100	100
4	H	6/8 (75%)	5 (83%)	1 (17%)	0	100	100
4	I	6/8 (75%)	6 (100%)	0	0	100	100
4	J	6/8 (75%)	4 (67%)	2 (33%)	0	100	100
All	All	1031/1052 (98%)	995 (96%)	35 (3%)	1 (0%)	51	53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	695	ASP

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	233/230 (101%)	223 (96%)	10 (4%)	29	27
1	C	231/230 (100%)	221 (96%)	10 (4%)	29	27
1	D	229/230 (100%)	222 (97%)	7 (3%)	40	39
2	F	2/2 (100%)	2 (100%)	0	100	100
2	K	2/2 (100%)	2 (100%)	0	100	100
2	L	2/2 (100%)	2 (100%)	0	100	100
2	M	2/2 (100%)	1 (50%)	1 (50%)	0	0
3	B	231/229 (101%)	223 (96%)	8 (4%)	36	34
4	G	7/8 (88%)	5 (71%)	2 (29%)	0	0
4	H	8/8 (100%)	6 (75%)	2 (25%)	0	0
4	I	8/8 (100%)	6 (75%)	2 (25%)	0	0
4	J	8/8 (100%)	7 (88%)	1 (12%)	4	1
All	All	963/959 (100%)	920 (96%)	43 (4%)	28	24

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	910	LYS
1	D	861	LYS
1	D	670	GLN
1	D	825	LYS
4	G	68	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	874	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](#)

5 non-standard protein/DNA/RNA residues 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
2	FME	L	503	6,2	8,9,10	2.59	1 (12%)	7,9,11	4.87	6 (85%)
2	FME	K	503	6,2	8,9,10	1.15	1 (12%)	7,9,11	5.22	4 (57%)
2	FME	F	503	6,2	8,9,10	2.65	1 (12%)	7,9,11	3.60	1 (14%)
3	SNC	B	686	3	4,7,8	1.68	1 (25%)	1,7,9	0.29	0
2	FME	M	503	6,2	8,9,10	2.13	1 (12%)	7,9,11	5.92	4 (57%)

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	FME	L	503	6,2	-	5/7/9/11	-
2	FME	K	503	6,2	-	3/7/9/11	-
2	FME	F	503	6,2	-	1/7/9/11	-
3	SNC	B	686	3	-	0/0/6/8	-
2	FME	M	503	6,2	-	5/7/9/11	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	503	FME	CA-N	7.20	1.56	1.46
2	L	503	FME	CA-N	7.00	1.56	1.46
2	M	503	FME	CA-N	5.52	1.54	1.46
3	B	686	SNC	CB-CA	2.29	1.58	1.53
2	K	503	FME	O1-CN	2.08	1.29	1.22

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	503	FME	CA-N-CN	14.80	145.59	122.82
2	K	503	FME	CA-N-CN	12.38	141.85	122.82
2	L	503	FME	CA-N-CN	10.59	139.11	122.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	503	FME	CA-N-CN	8.93	136.55	122.82
2	L	503	FME	O1-CN-N	3.97	135.72	125.27

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	M	503	FME	CB-CA-N-CN
2	M	503	FME	N-CA-CB-CG
2	F	503	FME	CB-CA-N-CN
2	K	503	FME	O1-CN-N-CA
2	L	503	FME	CB-CA-N-CN

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	503	FME	1	0
2	M	503	FME	4	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 28 ligands modelled in this entry, 7 are monoatomic - leaving 21 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	$\# Z > 2$
8	EDO	D	1004	-	3,3,3	0.49	0	2,2,2	0.98	0
8	EDO	B	1004	-	3,3,3	1.02	0	2,2,2	0.19	0
5	DHT	B	1001	-	24,24,24	1.38	3 (12%)	39,39,39	2.05	12 (30%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
8	EDO	B	1006	-	3,3,3	0.80	0	2,2,2	0.86	0
10	SO4	C	1007	-	4,4,4	0.27	0	6,6,6	0.21	0
8	EDO	B	1005	-	3,3,3	1.03	0	2,2,2	0.92	0
10	SO4	A	1006	-	4,4,4	0.41	0	6,6,6	0.75	0
10	SO4	C	1006	-	4,4,4	0.36	0	6,6,6	0.80	0
5	DHT	C	1001	-	24,24,24	1.27	2 (8%)	39,39,39	2.25	15 (38%)
9	GOL	A	1005	-	5,5,5	0.54	0	5,5,5	2.26	3 (60%)
5	DHT	D	1001	-	24,24,24	1.56	3 (12%)	39,39,39	2.41	15 (38%)
12	ACT	D	1006	-	3,3,3	0.93	0	3,3,3	0.54	0
8	EDO	C	1003	-	3,3,3	0.72	0	2,2,2	0.14	0
8	EDO	D	1007	-	3,3,3	1.49	0	2,2,2	1.04	0
8	EDO	A	1004	-	3,3,3	0.49	0	2,2,2	0.32	0
10	SO4	B	1007	-	4,4,4	0.32	0	6,6,6	0.43	0
10	SO4	D	1008	-	4,4,4	0.17	0	6,6,6	1.09	0
8	EDO	D	1005	-	3,3,3	0.38	0	2,2,2	1.17	0
11	PGE	C	1005	-	9,9,9	0.88	0	8,8,8	1.03	0
9	GOL	C	1004	-	5,5,5	0.54	0	5,5,5	0.91	0
5	DHT	A	1001	-	24,24,24	1.73	3 (12%)	39,39,39	2.55	16 (41%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	EDO	B	1005	-	-	1/1/1/1	-
8	EDO	D	1004	-	-	1/1/1/1	-
8	EDO	B	1004	-	-	1/1/1/1	-
8	EDO	C	1003	-	-	1/1/1/1	-
8	EDO	D	1005	-	-	0/1/1/1	-
8	EDO	D	1007	-	-	0/1/1/1	-
5	DHT	B	1001	-	-	-	0/4/4/4
8	EDO	A	1004	-	-	1/1/1/1	-
11	PGE	C	1005	-	-	2/7/7/7	-
5	DHT	C	1001	-	-	-	0/4/4/4
9	GOL	C	1004	-	-	2/4/4/4	-
9	GOL	A	1005	-	-	4/4/4/4	-
5	DHT	D	1001	-	-	-	0/4/4/4
8	EDO	B	1006	-	-	0/1/1/1	-
5	DHT	A	1001	-	-	-	0/4/4/4

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	1001	DHT	O17-C17	5.76	1.53	1.43
5	C	1001	DHT	O17-C17	3.93	1.50	1.43
5	B	1001	DHT	O17-C17	3.56	1.49	1.43
5	D	1001	DHT	O17-C17	3.35	1.49	1.43
5	D	1001	DHT	C19-C10	3.17	1.59	1.54

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	D	1001	DHT	C2-C3-C4	6.22	125.17	115.89
5	A	1001	DHT	C14-C8-C9	-5.99	101.07	109.09
5	D	1001	DHT	O3-C3-C4	-5.28	115.26	121.96
5	A	1001	DHT	C7-C8-C9	5.14	116.87	110.49
5	A	1001	DHT	C15-C16-C17	5.10	111.11	105.75

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	A	1005	GOL	C1-C2-C3-O3
9	C	1004	GOL	O1-C1-C2-C3
9	A	1005	GOL	O2-C2-C3-O3
9	A	1005	GOL	O1-C1-C2-C3
9	C	1004	GOL	C1-C2-C3-O3

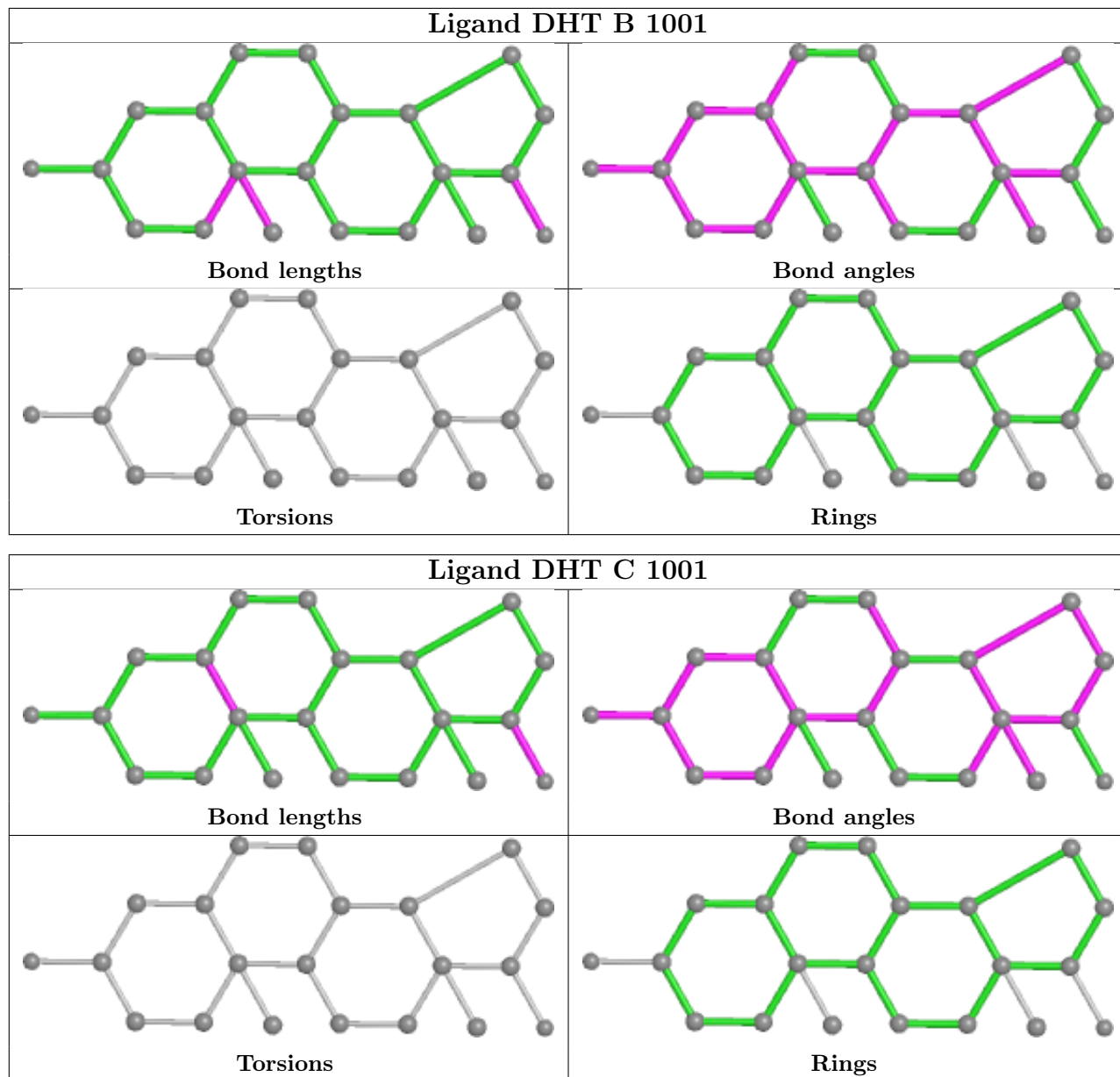
There are no ring outliers.

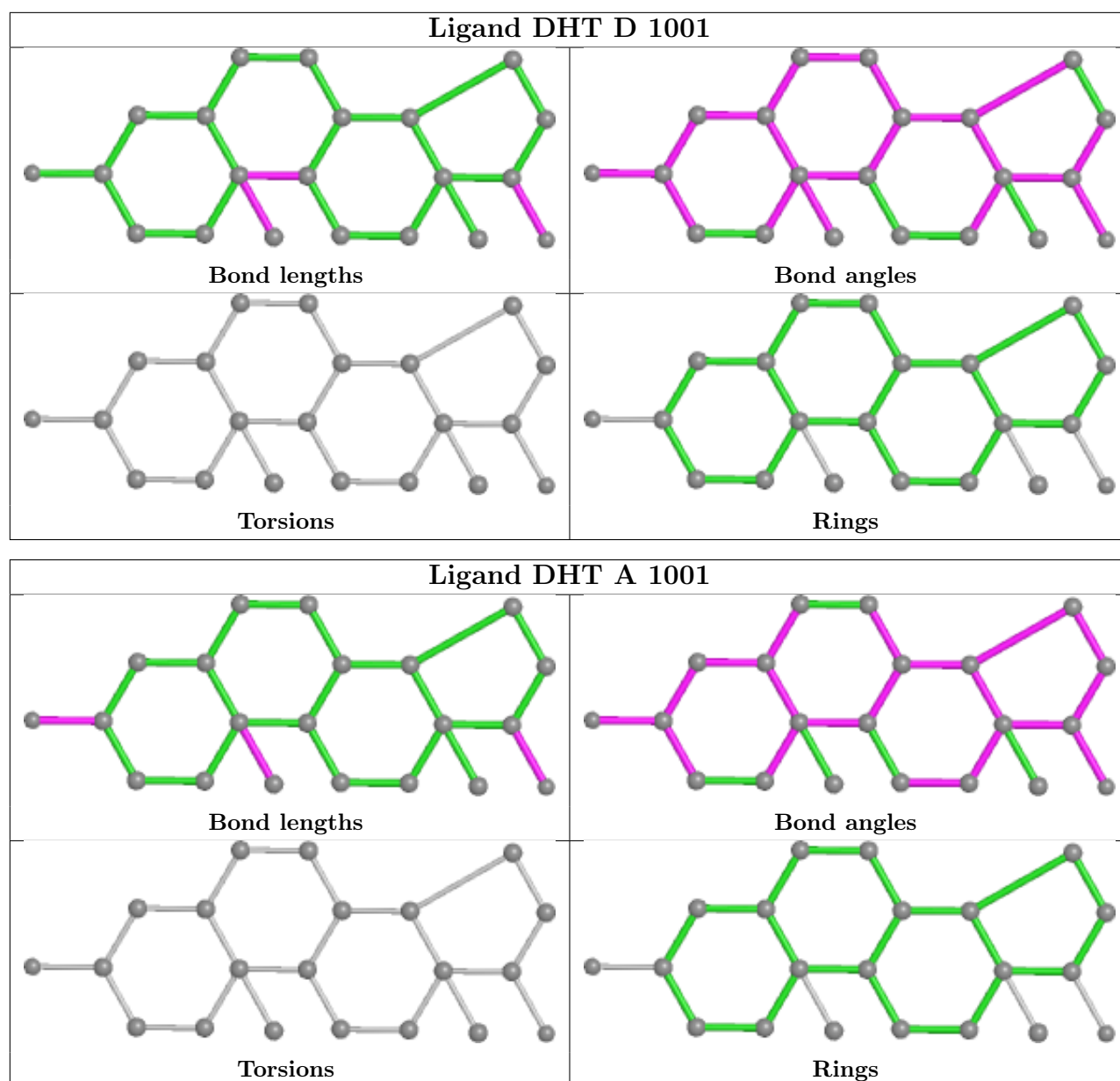
4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	A	1005	GOL	1	0
8	A	1004	EDO	2	0
11	C	1005	PGE	1	0
9	C	1004	GOL	1	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 [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	852:CYS	C	853:SER	N	1.66

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	252/252 (100%)	-0.23	0 100 100	24, 37, 59, 83	0
1	C	250/252 (99%)	-0.23	0 100 100	23, 37, 59, 81	0
1	D	249/252 (98%)	-0.23	0 100 100	22, 37, 58, 73	0
2	F	2/3 (66%)	-0.05	0 100 100	56, 56, 56, 69	0
2	K	2/3 (66%)	1.51	1 (50%) 0 0	63, 63, 63, 92	0
2	L	2/3 (66%)	0.75	0 100 100	61, 61, 61, 80	0
2	M	2/3 (66%)	0.55	0 100 100	67, 67, 67, 93	0
3	B	250/252 (99%)	-0.21	0 100 100	24, 38, 59, 76	0
4	G	7/8 (87%)	0.85	1 (14%) 2 3	63, 69, 76, 86	0
4	H	8/8 (100%)	0.52	1 (12%) 3 5	68, 75, 77, 81	0
4	I	8/8 (100%)	0.43	0 100 100	68, 76, 77, 90	0
4	J	8/8 (100%)	0.50	0 100 100	67, 80, 91, 92	0
All	All	1040/1052 (98%)	-0.20	3 (0%) 94 95	22, 37, 66, 93	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	K	505	ASP	2.8
4	H	66	LEU	2.3
4	G	63	LEU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains [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(\AA^2)	Q<0.9
3	SNC	B	686	8/9	0.92	0.10	34,39,51,57	2
2	FME	M	503	10/11	0.96	0.12	58,64,66,68	0
2	FME	F	503	10/11	0.97	0.12	53,55,63,65	0
2	FME	K	503	10/11	0.97	0.11	56,59,62,65	0
2	FME	L	503	10/11	0.97	0.10	48,51,58,60	0

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.

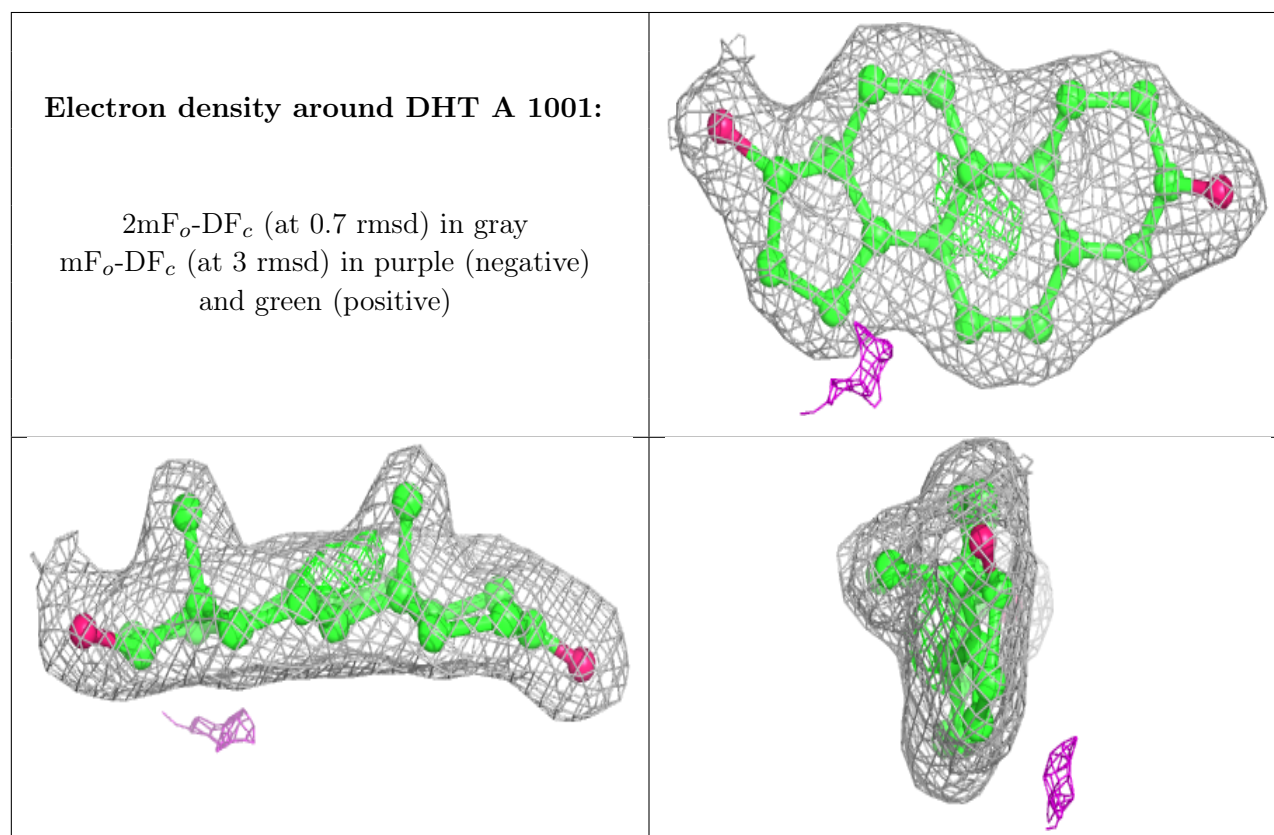
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
8	EDO	D	1005	4/4	0.91	0.15	59,64,65,69	0
11	PGE	C	1005	10/10	0.92	0.11	62,69,72,73	0
8	EDO	A	1004	4/4	0.93	0.10	60,65,66,71	0
8	EDO	C	1003	4/4	0.93	0.12	61,64,64,65	0
7	CL	A	1003	1/1	0.95	0.14	72,72,72,72	0
8	EDO	B	1004	4/4	0.96	0.12	38,43,43,46	0
8	EDO	B	1005	4/4	0.96	0.13	66,67,70,72	0
8	EDO	B	1006	4/4	0.96	0.16	54,60,60,63	0
5	DHT	A	1001	21/21	0.96	0.13	25,27,30,33	0
8	EDO	D	1004	4/4	0.96	0.12	53,57,58,64	0
7	CL	D	1003	1/1	0.96	0.08	74,74,74,74	0
8	EDO	D	1007	4/4	0.96	0.13	37,39,44,44	0
9	GOL	A	1005	6/6	0.96	0.12	41,48,50,54	0
10	SO4	A	1006	5/5	0.96	0.13	52,55,59,62	0
5	DHT	D	1001	21/21	0.96	0.11	27,28,29,30	0
10	SO4	C	1007	5/5	0.97	0.07	87,89,91,92	0
7	CL	B	1003	1/1	0.97	0.15	66,66,66,66	0
12	ACT	D	1006	4/4	0.97	0.13	62,63,67,68	0
5	DHT	B	1001	21/21	0.98	0.12	27,29,30,31	0
9	GOL	C	1004	6/6	0.98	0.10	39,46,49,53	0
5	DHT	C	1001	21/21	0.98	0.12	25,26,28,29	0
10	SO4	B	1007	5/5	0.99	0.12	50,50,54,56	0
10	SO4	C	1006	5/5	0.99	0.13	47,48,49,49	0
6	ZN	D	1002	1/1	0.99	0.10	49,49,49,49	0

Continued on next page...

Continued from previous page...

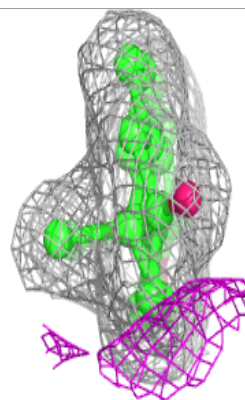
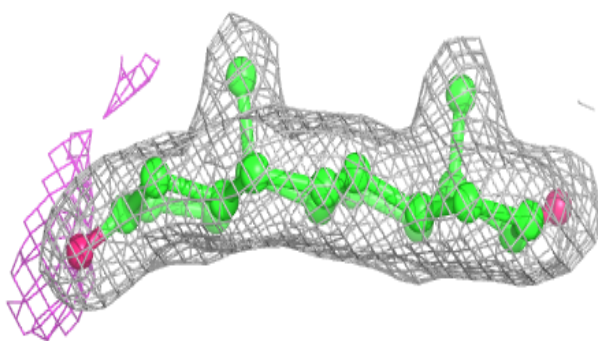
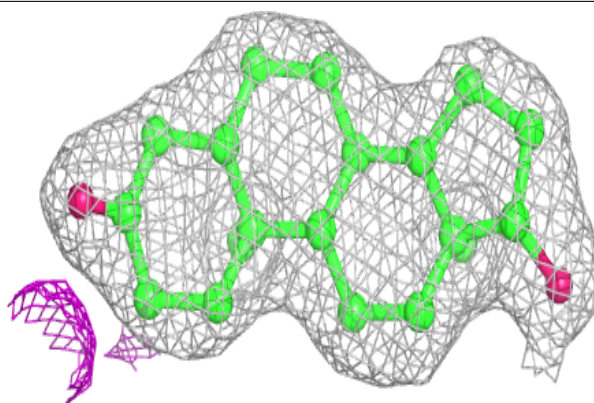
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
10	SO4	D	1008	5/5	0.99	0.12	53,56,58,60	0
6	ZN	B	1002	1/1	0.99	0.14	47,47,47,47	0
6	ZN	C	1002	1/1	0.99	0.09	53,53,53,53	0
6	ZN	A	1002	1/1	1.00	0.12	51,51,51,51	0

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.

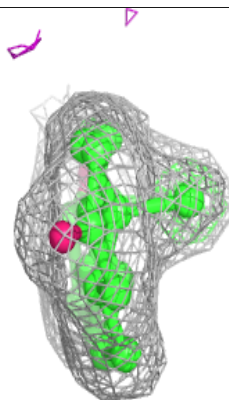
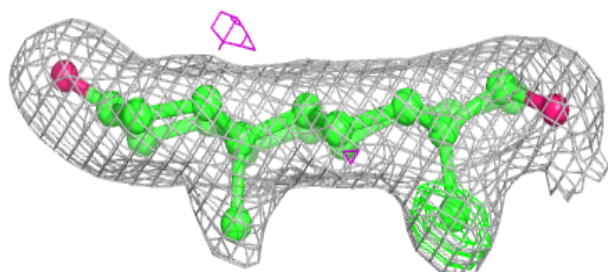
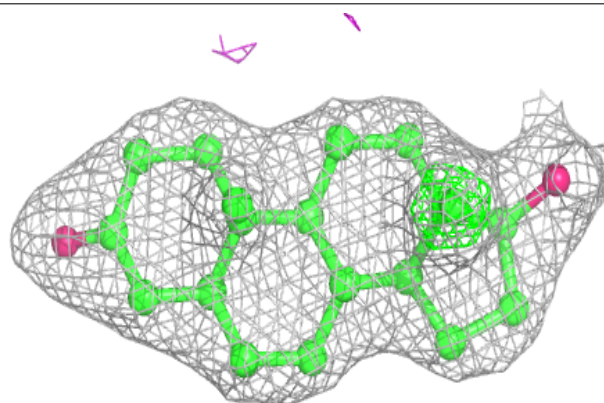


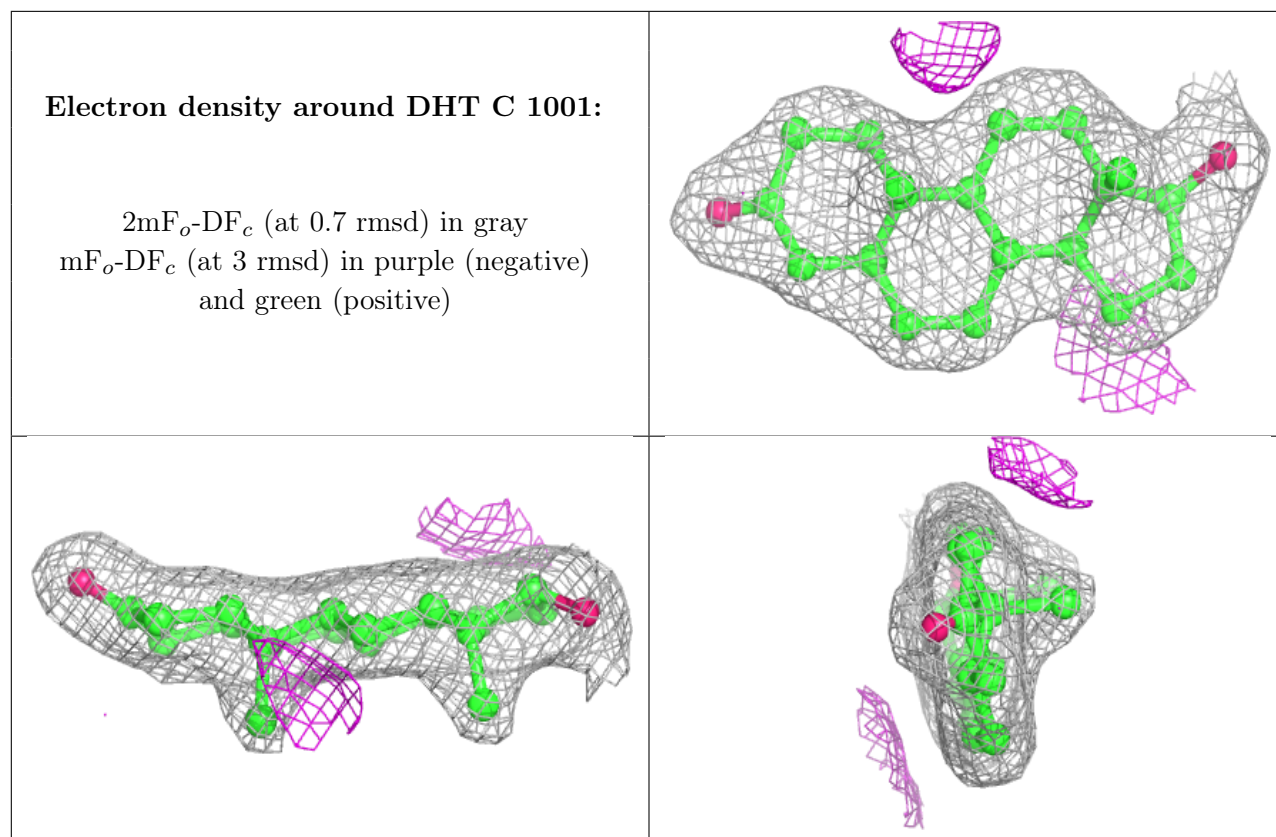
Electron density around DHT D 1001:

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

**Electron density around DHT B 1001:**

$2mF_o-DF_c$ (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.