



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 03:02 PM EDT

PDB ID : 7S6H
Title : Human PARP1 deltaV687-E688 bound to NAD⁺ analog EB-47 and to a DNA double strand break.
Authors : Rouleau-Turcotte, E.; Pascal, J.M.
Deposited on : 2021-09-14
Resolution : 3.10 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

<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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.35.1
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.35.1

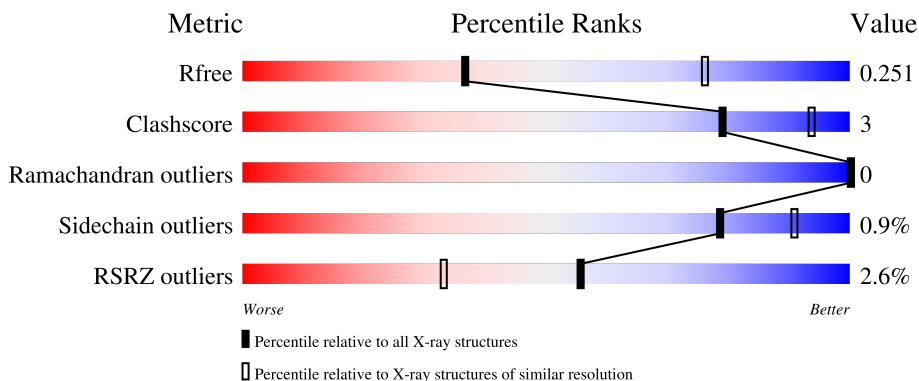
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	E	5	100%
1	M	5	80% 20%
2	F	5	100%
2	N	5	100%
3	B	504	85% 8% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	504	<p>% 86% 7% 7%</p>
4	A	276	<p>4% 75% 6% 19%</p>
4	C	276	<p>5% 74% 6% 20%</p>

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
5	EDO	B	1101	-	-	-	X
5	EDO	D	1101	-	-	-	X

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 22607 atoms, of which 11149 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 DNA chain called DNA (5'-D(*CP*GP*AP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				P
1	M	5	157	48	57	21	27	4	0	0	0
1	E	5	157	48	57	21	27	4	0	0	0

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*GP*TP*CP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				P
2	N	5	157	48	58	18	29	4	0	0	0
2	F	5	157	48	58	18	29	4	0	0	0

- Molecule 3 is a protein called Poly [ADP-ribose] polymerase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	B	469	7460	2383	3734	632	698	13	0	0	0
3	D	471	7477	2392	3737	635	700	13	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	517	MET	-	initiating methionine	UNP P09874
B	?	-	VAL	deletion	UNP P09874
B	?	-	GLU	deletion	UNP P09874
B	762	ALA	VAL	variant	UNP P09874
B	1015	LEU	-	expression tag	UNP P09874
B	1016	GLU	-	expression tag	UNP P09874
B	1017	HIS	-	expression tag	UNP P09874
B	1018	HIS	-	expression tag	UNP P09874

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1019	HIS	-	expression tag	UNP P09874
B	1020	HIS	-	expression tag	UNP P09874
B	1021	HIS	-	expression tag	UNP P09874
B	1022	HIS	-	expression tag	UNP P09874
D	517	MET	-	initiating methionine	UNP P09874
D	?	-	VAL	deletion	UNP P09874
D	?	-	GLU	deletion	UNP P09874
D	762	ALA	VAL	variant	UNP P09874
D	1015	LEU	-	expression tag	UNP P09874
D	1016	GLU	-	expression tag	UNP P09874
D	1017	HIS	-	expression tag	UNP P09874
D	1018	HIS	-	expression tag	UNP P09874
D	1019	HIS	-	expression tag	UNP P09874
D	1020	HIS	-	expression tag	UNP P09874
D	1021	HIS	-	expression tag	UNP P09874
D	1022	HIS	-	expression tag	UNP P09874

- Molecule 4 is a protein called Fusion of human PARP1 zinc fingers 1 and 3 (Zn1, Zn3).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
4	A	223	3456	1112	1702	302	327	13	0	0	0
4	C	221	3424	1103	1684	299	325	13	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

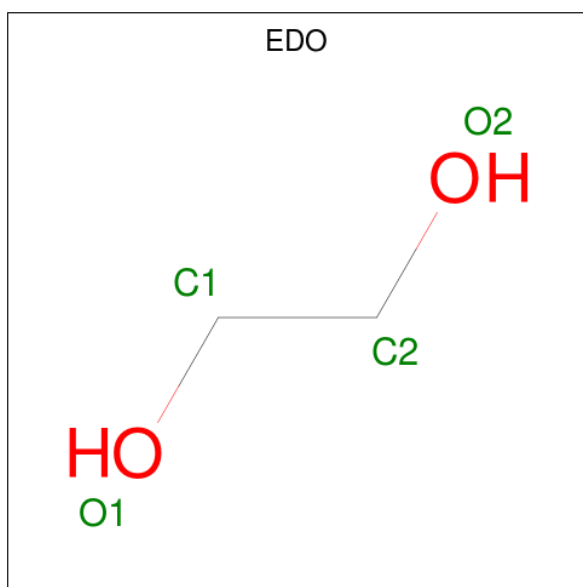
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P09874
A	-18	GLY	-	expression tag	UNP P09874
A	-17	SER	-	expression tag	UNP P09874
A	-16	SER	-	expression tag	UNP P09874
A	-15	HIS	-	expression tag	UNP P09874
A	-14	HIS	-	expression tag	UNP P09874
A	-13	HIS	-	expression tag	UNP P09874
A	-12	HIS	-	expression tag	UNP P09874
A	-11	HIS	-	expression tag	UNP P09874
A	-10	HIS	-	expression tag	UNP P09874
A	-9	SER	-	expression tag	UNP P09874
A	-8	SER	-	expression tag	UNP P09874
A	-7	GLY	-	expression tag	UNP P09874
A	-6	LEU	-	expression tag	UNP P09874

Continued on next page...

Continued from previous page...

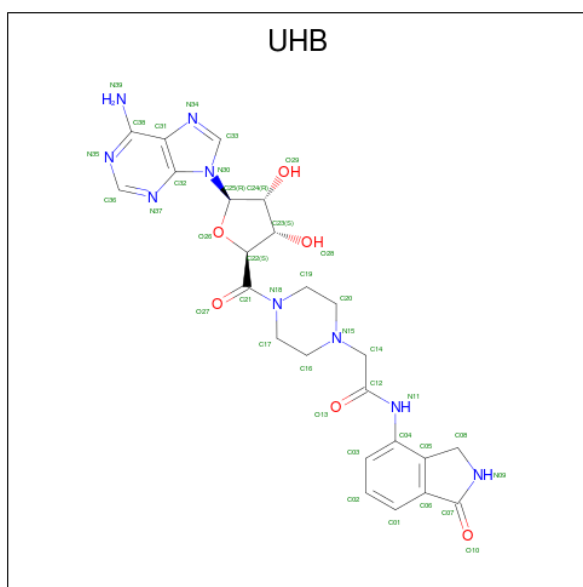
Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	VAL	-	expression tag	UNP P09874
A	-4	PRO	-	expression tag	UNP P09874
A	-3	ARG	-	expression tag	UNP P09874
A	-2	GLY	-	expression tag	UNP P09874
A	-1	SER	-	expression tag	UNP P09874
A	0	HIS	-	expression tag	UNP P09874
C	-19	MET	-	initiating methionine	UNP P09874
C	-18	GLY	-	expression tag	UNP P09874
C	-17	SER	-	expression tag	UNP P09874
C	-16	SER	-	expression tag	UNP P09874
C	-15	HIS	-	expression tag	UNP P09874
C	-14	HIS	-	expression tag	UNP P09874
C	-13	HIS	-	expression tag	UNP P09874
C	-12	HIS	-	expression tag	UNP P09874
C	-11	HIS	-	expression tag	UNP P09874
C	-10	HIS	-	expression tag	UNP P09874
C	-9	SER	-	expression tag	UNP P09874
C	-8	SER	-	expression tag	UNP P09874
C	-7	GLY	-	expression tag	UNP P09874
C	-6	LEU	-	expression tag	UNP P09874
C	-5	VAL	-	expression tag	UNP P09874
C	-4	PRO	-	expression tag	UNP P09874
C	-3	ARG	-	expression tag	UNP P09874
C	-2	GLY	-	expression tag	UNP P09874
C	-1	SER	-	expression tag	UNP P09874
C	0	HIS	-	expression tag	UNP P09874

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

- Molecule 6 is 2-[4-[(2S,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-bis(oxidanyl)oxolan-2-yl]carbon ylpiperazin-1-yl]-N-(1-oxidanylidene-2,3-dihydroisindol-4-yl)ethanamide (three-letter code: UHB) (formula: C₂₄H₂₇N₉O₆) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	H	N	O	0	0
			66	24	27	9	6		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
6	D	1	66	24	27	9	6	0	0

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	Zn	0	0
			2	2		
7	C	2	Total	Zn	0	0
			2	2		


- Molecule 8 is water.

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

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: DNA (5'-D(*CP*GP*AP*CP*G)-3')

Chain M: 



- Molecule 1: DNA (5'-D(*CP*GP*AP*CP*G)-3')

Chain E: 

There are no outlier residues recorded for this chain.

- Molecule 2: DNA (5'-D(*CP*GP*TP*CP*G)-3')

Chain N: 


There are no outlier residues recorded for this chain.

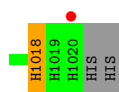
- Molecule 2: DNA (5'-D(*CP*GP*TP*CP*G)-3')

Chain F: 

There are no outlier residues recorded for this chain.

- Molecule 3: Poly [ADP-ribose] polymerase 1

Chain B: 



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	111.47Å 110.65Å 116.47Å 90.00° 115.06° 90.00°	Depositor
Resolution (Å)	48.52 – 3.10 48.52 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.52-3.10) 99.7 (48.52-3.10)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.10 (at 3.12Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.210 , 0.251 0.210 , 0.251	Depositor DCC
R_{free} test set	2150 reflections (4.62%)	wwPDB-VP
Wilson B-factor (Å ²)	124.3	Xtrriage
Anisotropy	0.135	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 89.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.025 for l,-k,h	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	22607	wwPDB-VP
Average B, all atoms (Å ²)	141.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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: ZN, EDO, UHB

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	E	0.44	0/112	0.62	0/171
1	M	0.46	0/112	0.63	0/171
2	F	0.46	0/110	0.82	0/168
2	N	0.45	0/110	0.83	0/168
3	B	0.25	0/3808	0.42	0/5139
3	D	0.25	0/3822	0.42	0/5159
4	A	0.24	0/1792	0.39	0/2413
4	C	0.24	0/1778	0.40	0/2395
All	All	0.26	0/11644	0.43	0/15784

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	100	57	57	0	0
1	M	100	57	57	1	0
2	F	99	58	58	0	0
2	N	99	58	58	0	0
3	B	3726	3734	3734	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	3740	3737	3737	18	0
4	A	1754	1702	1702	11	0
4	C	1740	1684	1684	10	0
5	B	4	4	6	0	0
5	D	4	4	6	0	0
6	B	39	27	27	1	0
6	D	39	27	27	1	0
7	A	2	0	0	0	0
7	C	2	0	0	0	0
8	B	4	0	0	0	0
8	C	2	0	0	0	0
8	D	4	0	0	0	0
All	All	11458	11149	11153	63	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 3.

All (63) 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)
3:B:831:LEU:HD12	3:B:1005:LEU:HD23	1.64	0.80
3:D:831:LEU:HD12	3:D:1005:LEU:HD23	1.63	0.79
3:B:618:TYR:O	3:B:622:THR:HG22	1.88	0.74
3:D:618:TYR:O	3:D:622:THR:HG22	1.88	0.73
3:D:622:THR:HG23	3:D:624:ASN:H	1.57	0.68
3:B:622:THR:HG23	3:B:624:ASN:H	1.57	0.68
4:C:36:ALA:HB2	4:C:51:TRP:CE3	2.32	0.64
4:A:36:ALA:HB2	4:A:51:TRP:CE3	2.33	0.63
3:D:673:ILE:HD11	3:D:794:TYR:HD1	1.66	0.60
3:D:1018:HIS:ND1	3:D:1018:HIS:O	2.34	0.60
3:B:673:ILE:HD11	3:B:794:TYR:HD1	1.66	0.59
3:B:1018:HIS:O	3:B:1018:HIS:ND1	2.35	0.59
3:D:606:SER:OG	3:D:609:ASP:OD1	2.21	0.59
3:B:606:SER:OG	3:B:609:ASP:OD1	2.20	0.58
4:C:76:GLU:N	4:C:76:GLU:OE2	2.37	0.58
4:A:76:GLU:N	4:A:76:GLU:OE2	2.37	0.57
4:A:285:ASP:OD2	4:A:355:ARG:NH2	2.38	0.56
4:C:285:ASP:OD2	4:C:355:ARG:NH2	2.37	0.56
4:A:60:VAL:HG12	4:A:60:VAL:O	2.07	0.55
4:C:60:VAL:HG12	4:C:60:VAL:O	2.06	0.55
3:D:962:ILE:HG13	3:D:971:LEU:HD21	1.91	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:1002:LEU:HD23	3:D:1003:LEU:N	2.26	0.51
4:A:339:PHE:O	4:A:343:SER:N	2.44	0.51
3:B:962:ILE:HG13	3:B:971:LEU:HD21	1.92	0.51
4:A:270:GLN:NE2	4:A:292:LEU:O	2.42	0.50
3:B:1002:LEU:HD23	3:B:1003:LEU:N	2.27	0.50
4:C:339:PHE:O	4:C:343:SER:N	2.44	0.48
3:B:660:GLY:O	3:B:661:THR:HG22	2.14	0.48
3:D:827:ASN:ND2	3:D:906:ASN:OD1	2.48	0.47
4:A:6:ASP:N	4:A:6:ASP:OD1	2.45	0.47
3:B:827:ASN:ND2	3:B:906:ASN:OD1	2.47	0.47
6:D:1102:UHB:O13	6:D:1102:UHB:C20	2.62	0.47
3:B:563:VAL:HG21	4:A:322:MET:SD	2.56	0.46
3:B:799:THR:HG22	3:B:841:ARG:HA	1.97	0.46
3:B:958:PRO:O	3:B:961:ASN:ND2	2.44	0.46
3:D:799:THR:HG22	3:D:841:ARG:HA	1.98	0.46
3:D:958:PRO:O	3:D:961:ASN:ND2	2.44	0.45
4:C:12:GLU:OE2	4:C:15:LYS:NZ	2.47	0.45
3:B:863:GLY:N	3:B:904:SER:OG	2.49	0.45
6:B:1102:UHB:H20	6:B:1102:UHB:O13	2.17	0.45
3:B:834:ILE:HD11	3:B:1006:LYS:HB2	1.98	0.44
4:C:270:GLN:NE2	4:C:292:LEU:O	2.42	0.44
3:D:834:ILE:HD11	3:D:1006:LYS:HB2	1.98	0.44
3:D:863:GLY:N	3:D:904:SER:OG	2.50	0.44
3:B:985:LEU:HD12	3:B:985:LEU:N	2.33	0.43
3:B:932:LEU:HD11	3:B:938:ILE:HD11	2.01	0.43
3:D:932:LEU:HD11	3:D:938:ILE:HD11	2.01	0.43
4:C:270:GLN:OE1	4:C:301:GLN:NE2	2.51	0.42
3:D:985:LEU:HD12	3:D:985:LEU:N	2.34	0.42
1:M:1:DC:H5''	3:B:621:LYS:HE3	2.01	0.42
3:D:821:THR:HB	3:D:900:MET:HA	2.01	0.42
3:D:962:ILE:HG22	3:D:963:SER:N	2.35	0.42
3:B:891:PHE:HB2	3:B:990:ILE:CD1	2.50	0.42
3:B:831:LEU:CD1	3:B:1005:LEU:HD23	2.43	0.42
3:B:962:ILE:HG22	3:B:963:SER:N	2.34	0.41
4:A:273:PRO:HB3	4:A:282:ARG:CZ	2.51	0.41
3:B:734:ASN:ND2	4:A:319:THR:HG22	2.36	0.41
3:B:821:THR:HB	3:B:900:MET:HA	2.01	0.41
4:C:273:PRO:HB3	4:C:282:ARG:CZ	2.50	0.41
4:A:36:ALA:HB2	4:A:51:TRP:CD2	2.56	0.41
3:D:891:PHE:HB2	3:D:990:ILE:CD1	2.51	0.41
4:C:36:ALA:HB2	4:C:51:TRP:CD2	2.55	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:712:ILE:O	3:B:716:VAL:HG23	2.21	0.40

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	
3	B	463/504 (92%)	446 (96%)	17 (4%)	0	100	100
3	D	465/504 (92%)	447 (96%)	18 (4%)	0	100	100
4	A	219/276 (79%)	212 (97%)	7 (3%)	0	100	100
4	C	217/276 (79%)	210 (97%)	7 (3%)	0	100	100
All	All	1364/1560 (87%)	1315 (96%)	49 (4%)	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	
3	B	408/441 (92%)	404 (99%)	4 (1%)	76	90
3	D	408/441 (92%)	403 (99%)	5 (1%)	71	88
4	A	188/243 (77%)	187 (100%)	1 (0%)	88	94
4	C	187/243 (77%)	186 (100%)	1 (0%)	88	94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1191/1368 (87%)	1180 (99%)	11 (1%)	78 91

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	603	GLN
3	B	817	TYR
3	B	982	THR
3	B	1018	HIS
4	A	330	ARG
4	C	330	ARG
3	D	603	GLN
3	D	809	GLU
3	D	817	TYR
3	D	982	THR
3	D	1018	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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
6	UHB	D	1102	-	41,44,44	0.47	0	50,65,65	0.72	2 (4%)
6	UHB	B	1102	-	41,44,44	0.46	0	50,65,65	0.77	2 (4%)
5	EDO	D	1101	-	3,3,3	0.34	0	2,2,2	0.32	0
5	EDO	B	1101	-	3,3,3	0.34	0	2,2,2	0.34	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
6	UHB	D	1102	-	-	12/16/55/55	1/6/6/6
6	UHB	B	1102	-	-	9/16/55/55	1/6/6/6
5	EDO	D	1101	-	-	1/1/1/1	-
5	EDO	B	1101	-	-	1/1/1/1	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	1102	UHB	C12-C14-N15	2.56	119.30	113.36
6	B	1102	UHB	C31-C38-N39	2.43	124.04	120.35
6	D	1102	UHB	C31-C38-N39	2.36	123.93	120.35
6	D	1102	UHB	C12-C14-N15	2.25	118.57	113.36

There are no chirality outliers.

All (23) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	1102	UHB	C14-C12-N11-C04
6	B	1102	UHB	C12-C14-N15-C20
6	B	1102	UHB	N18-C21-C22-C23
6	D	1102	UHB	C12-C14-N15-C20
6	D	1102	UHB	C22-C21-N18-C19
6	D	1102	UHB	N18-C21-C22-C23
6	D	1102	UHB	O27-C21-N18-C17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	D	1102	UHB	O27-C21-N18-C19
6	D	1102	UHB	C14-C12-N11-C04
6	B	1102	UHB	O13-C12-N11-C04
6	B	1102	UHB	C03-C04-N11-C12
6	D	1102	UHB	C05-C04-N11-C12
6	D	1102	UHB	O13-C12-N11-C04
6	B	1102	UHB	C05-C04-N11-C12
6	D	1102	UHB	C12-C14-N15-C16
6	D	1102	UHB	C03-C04-N11-C12
6	D	1102	UHB	C22-C21-N18-C17
6	B	1102	UHB	O27-C21-C22-C23
6	D	1102	UHB	O27-C21-C22-C23
6	B	1102	UHB	O13-C12-C14-N15
5	B	1101	EDO	O1-C1-C2-O2
5	D	1101	EDO	O1-C1-C2-O2
6	B	1102	UHB	C12-C14-N15-C16

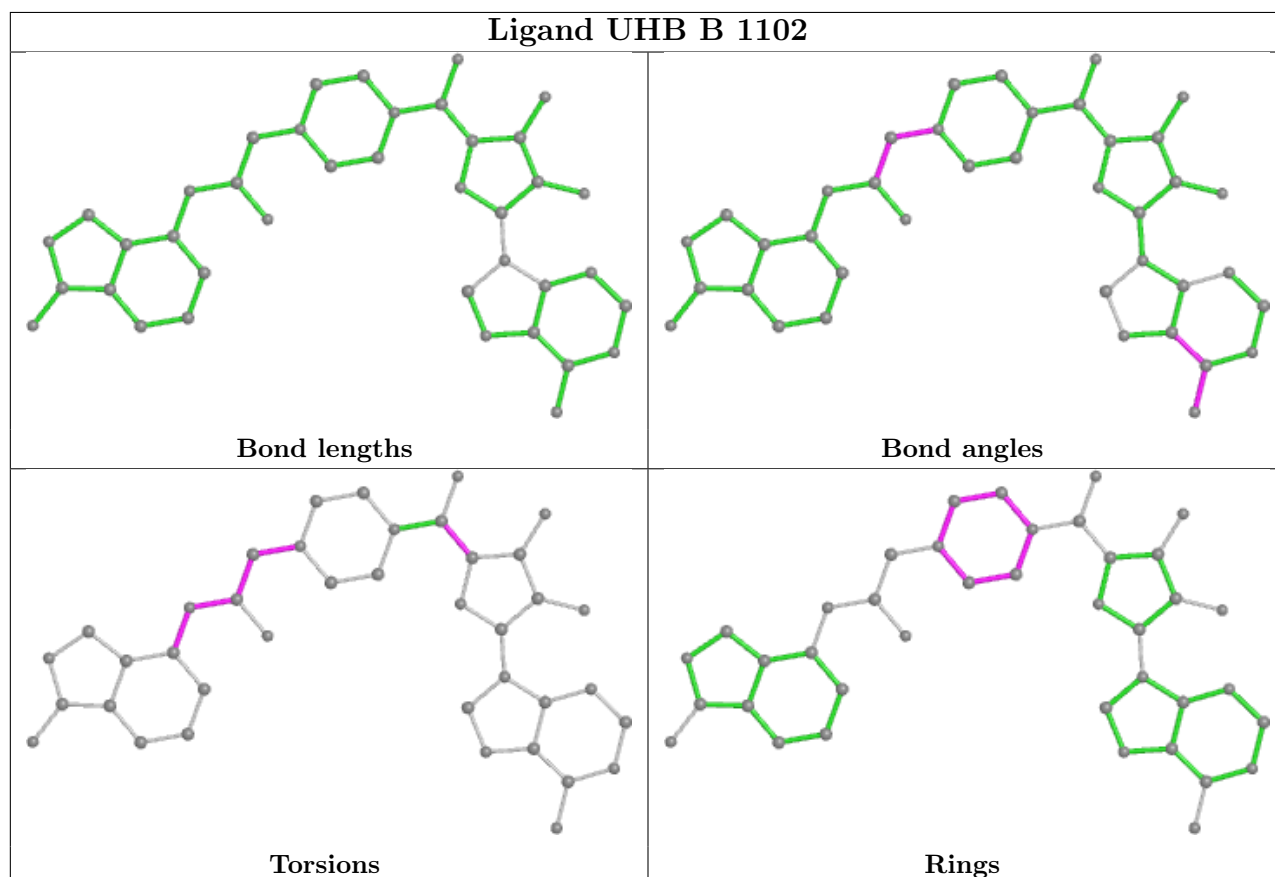
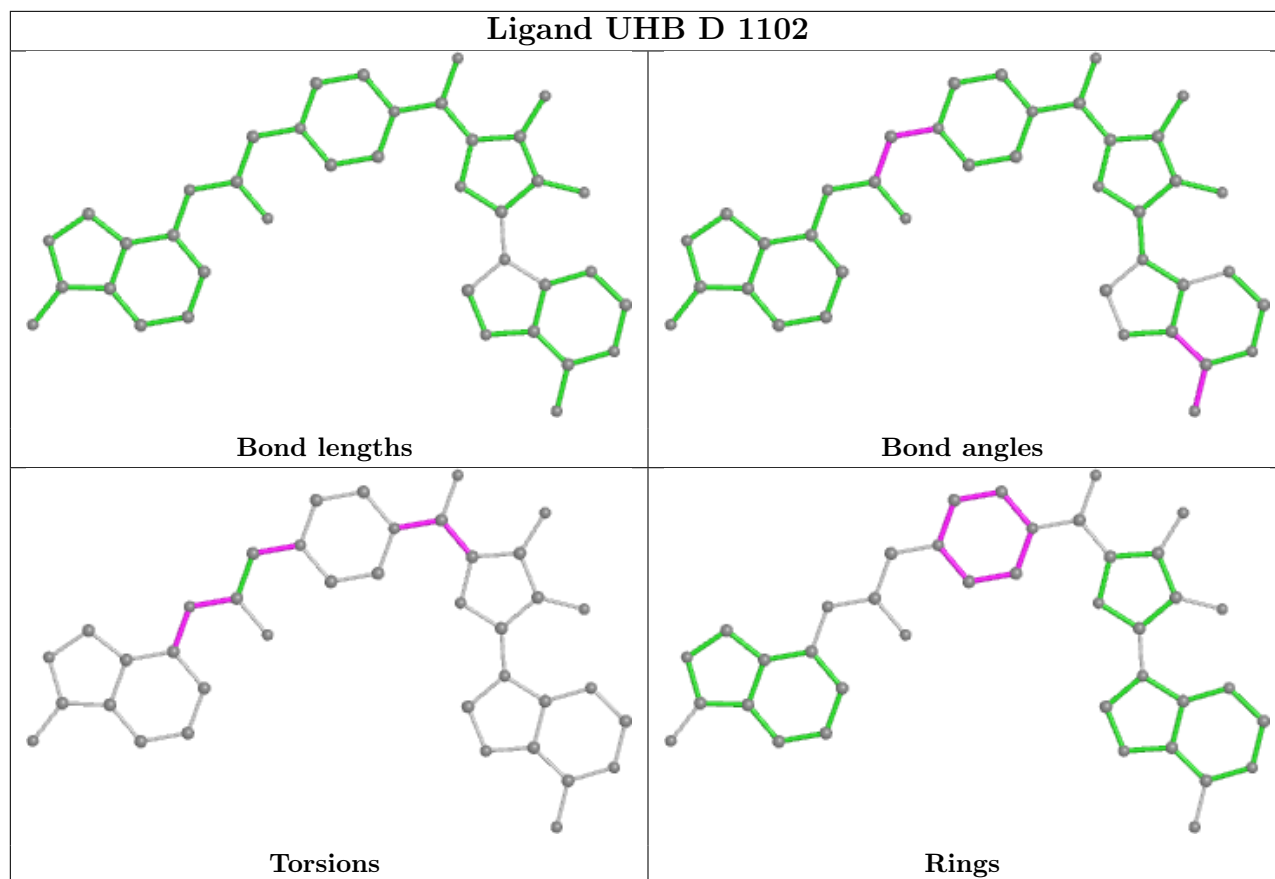
All (2) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	1102	UHB	C16-C17-C19-C20-N15-N18
6	D	1102	UHB	C16-C17-C19-C20-N15-N18

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	1102	UHB	1	0
6	B	1102	UHB	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](#)

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, 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	E	5/5 (100%)	0.20	0 100 100	94, 99, 112, 116	0
1	M	5/5 (100%)	0.84	0 100 100	102, 105, 117, 118	0
2	F	5/5 (100%)	-0.04	0 100 100	114, 116, 121, 123	0
2	N	5/5 (100%)	0.44	0 100 100	114, 116, 121, 124	0
3	B	469/504 (93%)	0.20	6 (1%) 77 59	93, 117, 166, 233	0
3	D	471/504 (93%)	0.16	6 (1%) 77 59	89, 118, 164, 203	0
4	A	223/276 (80%)	0.23	12 (5%) 25 12	100, 137, 195, 241	0
4	C	221/276 (80%)	0.31	13 (5%) 22 10	102, 136, 193, 237	0
All	All	1404/1580 (88%)	0.21	37 (2%) 56 33	89, 122, 179, 241	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	C	64	ILE	4.1
4	C	226	LYS	3.9
4	A	62	HIS	3.7
3	D	1020	HIS	3.4
3	B	661	THR	3.2
4	C	227	GLU	3.1
4	A	228	LYS	3.0
4	C	79	TRP	2.9
4	A	223	LYS	2.9
4	A	67	PRO	2.9
4	A	91	ALA	2.8
4	C	62	HIS	2.7
4	C	67	PRO	2.6
4	C	229	ASP	2.6
4	A	65	ARG	2.5
3	B	951	LEU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	A	302	LEU	2.5
3	D	679	VAL	2.4
4	A	252	LEU	2.4
4	C	228	LYS	2.3
4	A	64	ILE	2.3
4	C	233	LYS	2.3
3	B	756	ASP	2.3
4	C	224	SER	2.3
3	B	1020	HIS	2.3
3	B	531	ALA	2.1
4	A	227	GLU	2.1
4	A	343	SER	2.1
4	C	9	TYR	2.1
4	C	234	LEU	2.1
3	D	547	GLU	2.1
3	D	553	PHE	2.0
3	D	674	LYS	2.0
4	C	260	ASP	2.0
3	D	575	LEU	2.0
4	A	245	ILE	2.0
3	B	575	LEU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

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

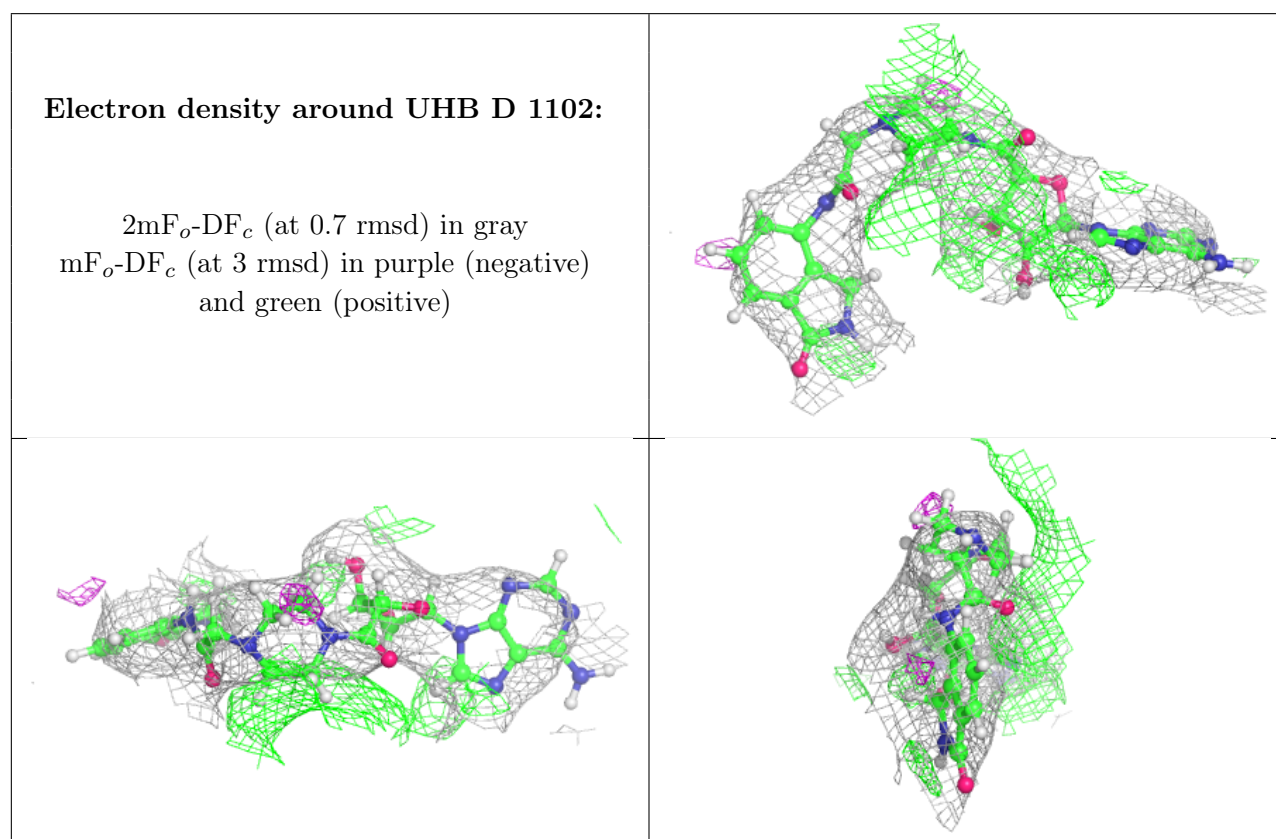
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	EDO	B	1101	4/4	0.72	0.46	101,129,155,155	0
5	EDO	D	1101	4/4	0.75	0.64	108,129,156,156	0

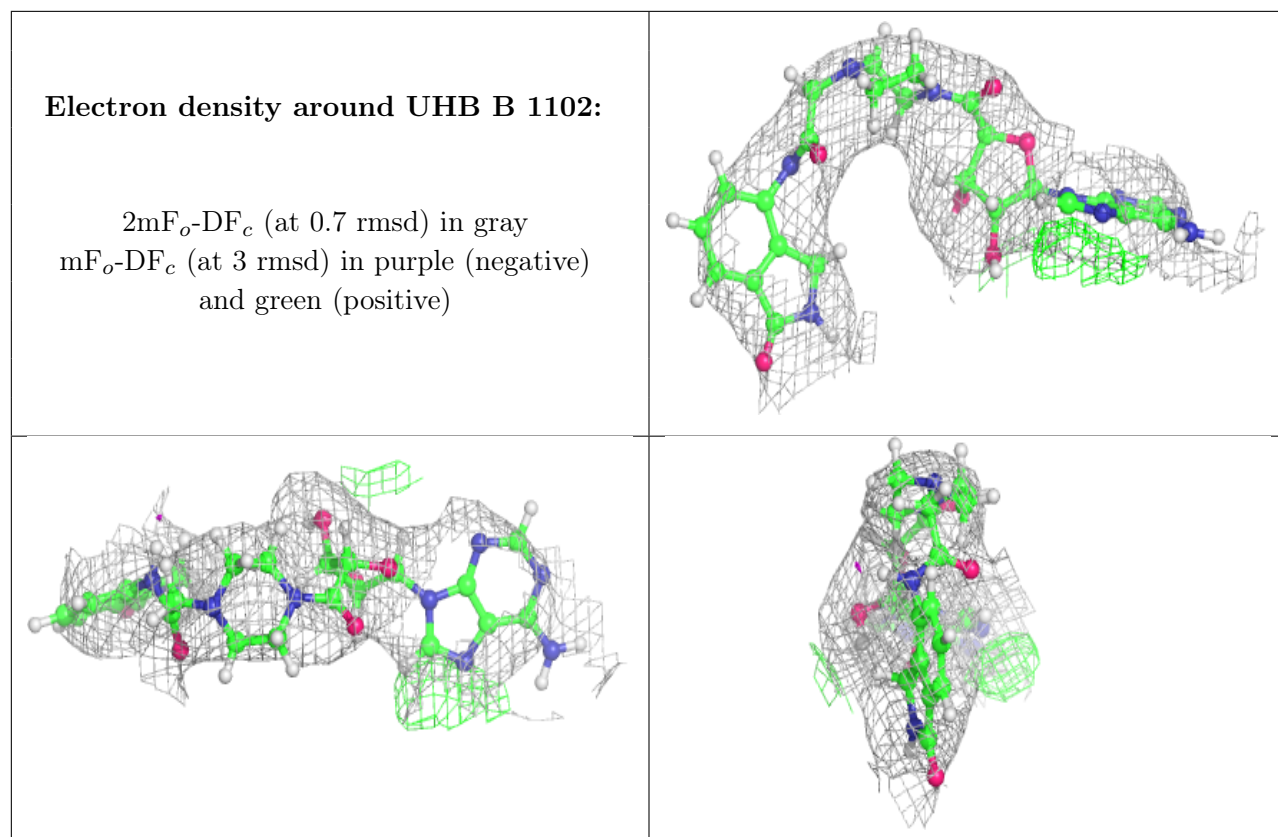
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
6	UHB	D	1102	39/39	0.90	0.31	74,113,160,178	0
6	UHB	B	1102	39/39	0.92	0.28	74,114,154,177	0
7	ZN	C	401	1/1	0.95	0.35	185,185,185,185	0
7	ZN	C	402	1/1	0.95	0.26	220,220,220,220	0
7	ZN	A	402	1/1	0.97	0.20	175,175,175,175	0
7	ZN	A	401	1/1	0.98	0.31	155,155,155,155	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.





6.5 Other polymers [i](#)

There are no such residues in this entry.