



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 8, 2026 – 03:19 PM UTC

PDB ID : 8CG9 / pdb_00008cg9
Title : Crystal structure of human DNA cross-link repair 1A in complex with a cyclic N-hydroxyurea inhibitor
Authors : Newman, J.A.; Yosaatmadja, Y.; Baddock, H.T.; Bielinski, M.; von Delft, F.; Bountra, C.; McHugh, P.J.; Schofield, C.J.; Gileadi, O.
Deposited on : 2023-02-03
Resolution : 1.68 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	:	4-5-2 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

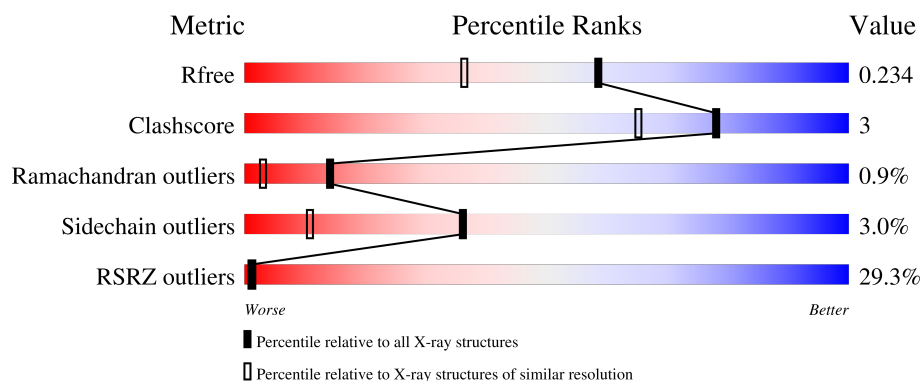
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 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	1054 (1.68-1.68)
Clashscore	190562	1078 (1.68-1.68)
Ramachandran outliers	187476	1068 (1.68-1.68)
Sidechain outliers	187428	1067 (1.68-1.68)
RSRZ outliers	180081	1055 (1.68-1.68)

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	345	28% 86% 10% .

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 2749 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 DNA cross-link repair 1A protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	331	Total	C	N	O	S	0	1	0
			2554	1658	420	460	16			

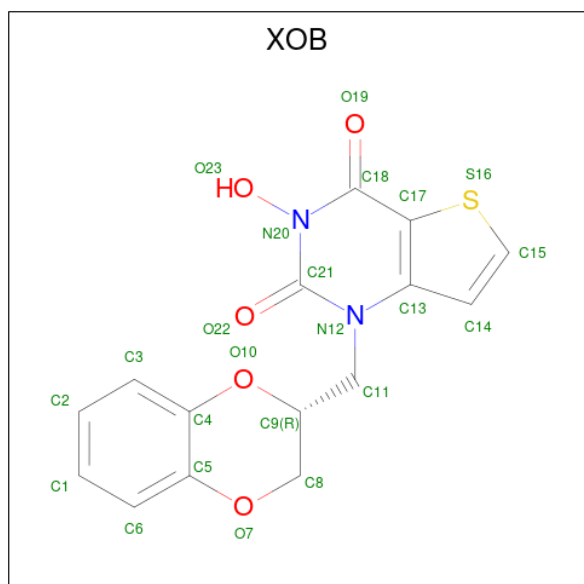
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	697	MET	ARG	conflict	UNP Q6PJP8

- Molecule 2 is NICKEL (II) ION (CCD ID: NI) (formula: Ni).

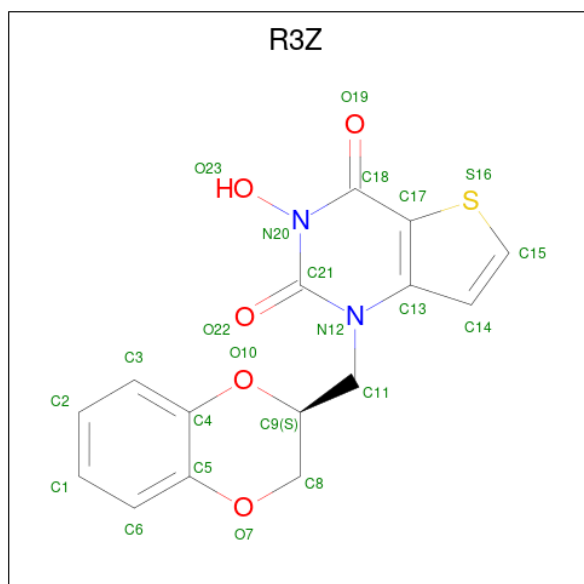
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ni	0	0
			1	1		

- Molecule 3 is 1-[[[(3 {R})-2,3-dihydro-1,4-benzodioxin-3-yl)methyl]-3-oxidanyl-thieno[3,2-d]pyrimidine-2,4-dione (CCD ID: XOB) (formula: C₁₅H₁₂N₂O₅S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			23	15	2	5	1		

- Molecule 4 is 1-[(2S)-2,3-dihydro-1,4-benzodioxin-2-ylmethyl]-3-hydroxythieno[3,2-d]pyrimidine-2,4(1H,3H)-dione (CCD ID: R3Z) (formula: C₁₅H₁₂N₂O₅S).

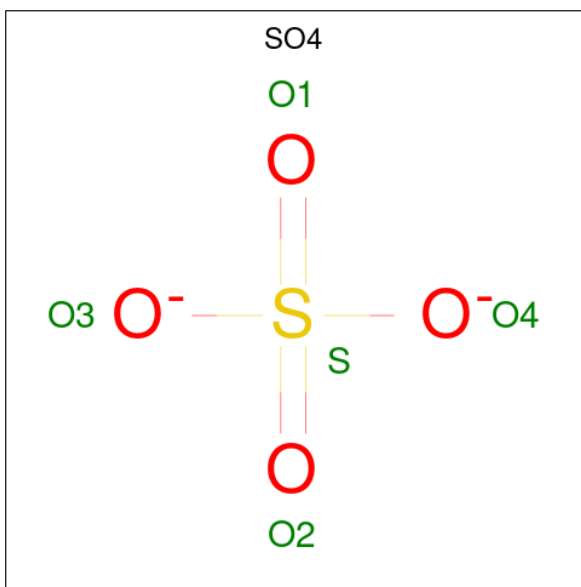


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			23	15	2	5	1		

- Molecule 5 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Zn	0	0
			1	1		

- Molecule 6 is SULFATE ION (CCD ID: SO4) (formula: O₄S).



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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	142	Total	O	0	0
			142	142		

- Molecule 1: DNA cross-link repair 1A protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	51.52Å 57.15Å 113.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.08 – 1.68 51.08 – 1.68	Depositor EDS
% Data completeness (in resolution range)	99.8 (51.08-1.68) 99.9 (51.08-1.68)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.58 (at 1.68Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.223 , 0.231 0.224 , 0.234	Depositor DCC
R_{free} test set	1898 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å ²)	26.2	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 42.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2749	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.63% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, R3Z, NI, XOB, ZN

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	0.17	0/2623	0.38	0/3574

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

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	2554	0	2478	14	0
2	A	1	0	0	0	0
3	A	23	0	0	0	0
4	A	23	0	0	0	0
5	A	1	0	0	0	0
6	A	5	0	0	0	0
7	A	142	0	0	0	1
All	All	2749	0	2478	14	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 3.

All (14) 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:859:ILE:HD11	1:A:889:ALA:HB1	1.77	0.67
1:A:962:THR:O	1:A:990:PRO:HA	2.07	0.55
1:A:734:HIS:CD2	1:A:880:SER:HB3	2.48	0.49
1:A:884:GLU:HA	1:A:887:PHE:CE2	2.49	0.47
1:A:773:LEU:HB2	1:A:788:LEU:HD22	1.96	0.46
1:A:702:PRO:HD3	1:A:719:TYR:CZ	2.49	0.46
1:A:814:GLY:O	1:A:838:ASP:HB2	2.17	0.45
1:A:843:SER:OG	1:A:845:GLU:OE1	2.33	0.44
1:A:801:LEU:HD13	1:A:811:LEU:HD13	2.01	0.42
1:A:746:THR:O	1:A:746:THR:OG1	2.37	0.42
1:A:712:PHE:CD1	1:A:727:ALA:HB3	2.55	0.42
1:A:1018:VAL:O	1:A:1024:ARG:HD2	2.20	0.42
1:A:949:CYS:O	1:A:952:LYS:HG2	2.20	0.41
1:A:859:ILE:HG23	1:A:893:VAL:HG11	2.03	0.40

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)
7:A:1254:HOH:O	7:A:1266:HOH:O[3_545]	2.17	0.03

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	328/345 (95%)	316 (96%)	9 (3%)	3 (1%)	14 3

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	829	ASP
1	A	964	TRP
1	A	953	TYR

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	272/307 (89%)	264 (97%)	8 (3%)	37 11

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	746	THR
1	A	768	GLN
1	A	816	PHE
1	A	879	TYR
1	A	934	MET
1	A	957	LEU
1	A	976	VAL
1	A	977	ILE

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\# Z > 2$
4	R3Z	A	1103	2,5	26,26,26	0.74	0	33,38,38	1.68	7 (21%)
6	SO4	A	1105	-	4,4,4	0.23	0	6,6,6	0.05	0
3	XOB	A	1102	5	26,26,26	0.82	1 (3%)	33,38,38	1.72	8 (24%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	R3Z	A	1103	2,5	-	0/4/13/13	0/4/4/4
3	XOB	A	1102	5	-	1/4/13/13	0/4/4/4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1102	XOB	C21-N20	-2.07	1.38	1.38

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1103	R3Z	O22-C21-N20	-5.25	121.11	123.72
3	A	1102	XOB	O22-C21-N20	-5.03	121.22	123.72
3	A	1102	XOB	C9-C11-N12	3.43	115.57	112.57
4	A	1103	R3Z	C21-N20-C18	-3.36	122.98	126.92
3	A	1102	XOB	C21-N20-C18	-3.20	123.17	126.92
4	A	1103	R3Z	O23-N20-C21	3.03	117.94	116.27
4	A	1103	R3Z	C14-C13-C17	-2.90	110.65	113.98
3	A	1102	XOB	C14-C13-C17	-2.77	110.80	113.98
3	A	1102	XOB	C11-N12-C13	2.70	122.29	119.95
4	A	1103	R3Z	C11-N12-C13	2.60	122.20	119.95
3	A	1102	XOB	O23-N20-C21	2.51	117.66	116.27
4	A	1103	R3Z	C13-C17-S16	2.40	112.50	110.82
3	A	1102	XOB	C13-C17-S16	2.24	112.38	110.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1103	R3Z	N12-C21-N20	2.20	119.25	115.85
3	A	1102	XOB	N12-C21-N20	2.15	119.18	115.85

There are no chirality outliers.

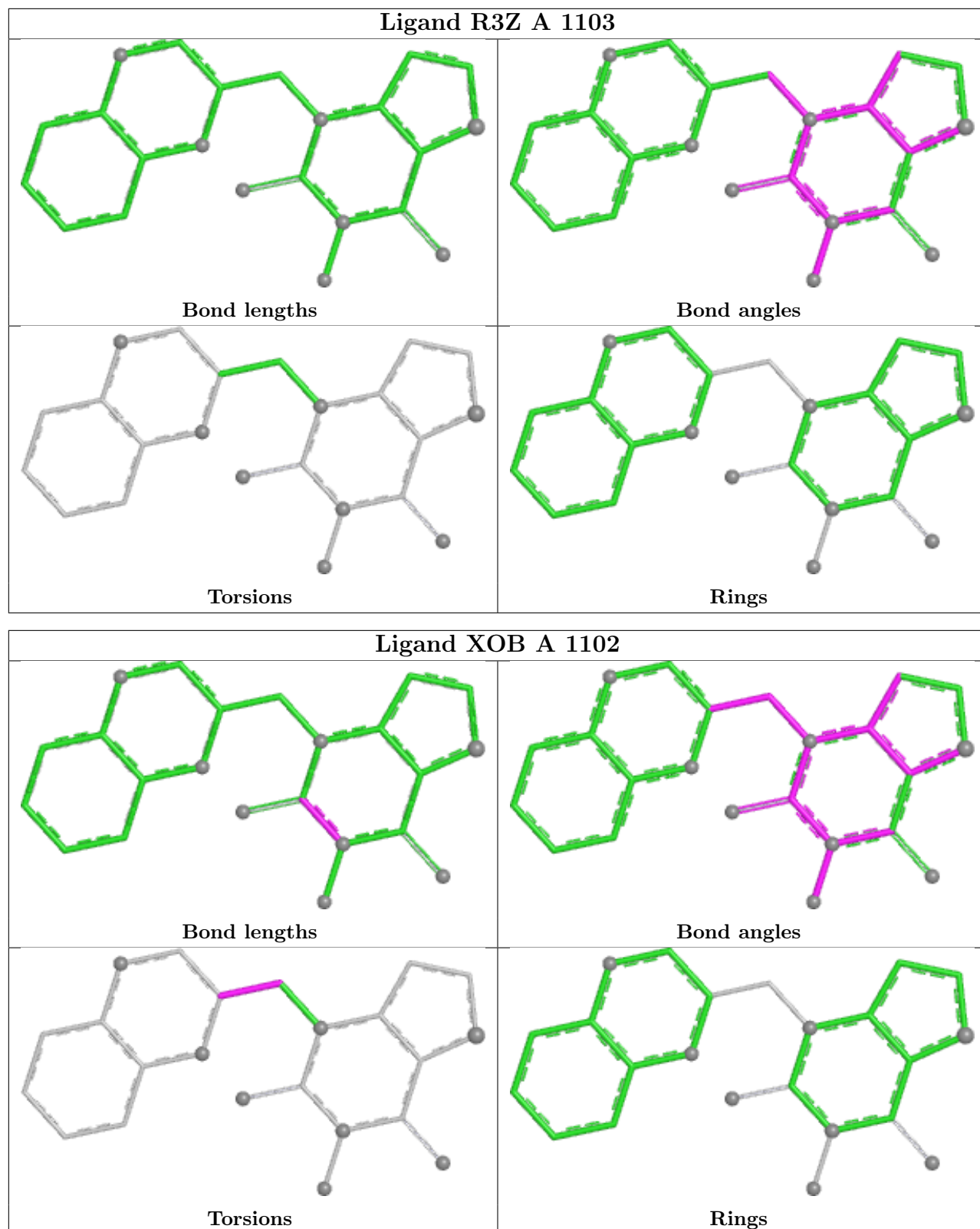
All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1102	XOB	N12-C11-C9-O10

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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	331/345 (95%)	1.39	97 (29%) 1 1	16, 38, 67, 85	1 (0%)

All (97) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	879	TYR	5.9
1	A	826	LEU	4.8
1	A	954	ASN	4.6
1	A	946	LEU	4.4
1	A	953	TYR	4.3
1	A	962	THR	4.2
1	A	912	ASN	4.1
1	A	964	TRP	4.0
1	A	942	LEU	4.0
1	A	977	ILE	3.9
1	A	925	CYS	3.9
1	A	939	PHE	3.9
1	A	982	GLY	3.9
1	A	747	PHE	3.8
1	A	963	GLY	3.8
1	A	949	CYS	3.8
1	A	950	GLY	3.7
1	A	951	GLY	3.7
1	A	928	LEU	3.7
1	A	980	THR	3.7
1	A	986	ILE	3.6
1	A	919	LEU	3.6
1	A	947	LYS	3.5
1	A	965	THR	3.5
1	A	916	ILE	3.5
1	A	956	ILE	3.4
1	A	745	PHE	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	703	PHE	3.4
1	A	700	THR	3.4
1	A	746	THR	3.4
1	A	872	ALA	3.3
1	A	780	ILE	3.3
1	A	957	LEU	3.3
1	A	983	ASN	3.3
1	A	721	VAL	3.3
1	A	866	VAL	3.2
1	A	976	VAL	3.2
1	A	741	LEU	3.1
1	A	944	SER	3.1
1	A	913	ILE	3.1
1	A	984	ILE	3.1
1	A	981	LYS	3.0
1	A	711	GLY	2.9
1	A	722	VAL	2.8
1	A	873	LEU	2.8
1	A	827	LEU	2.7
1	A	748	PRO	2.7
1	A	744	HIS	2.6
1	A	985	SER	2.6
1	A	868	LEU	2.6
1	A	935	MET	2.6
1	A	914	PRO	2.6
1	A	726	THR	2.5
1	A	847	THR	2.5
1	A	769	TYR	2.5
1	A	864	GLU	2.5
1	A	863	PHE	2.4
1	A	724	GLY	2.4
1	A	828	ALA	2.4
1	A	727	ALA	2.4
1	A	937	ILE	2.4
1	A	926	SER	2.3
1	A	876	CYS	2.3
1	A	917	ASN	2.3
1	A	720	GLY	2.3
1	A	961	PRO	2.3
1	A	865	ALA	2.3
1	A	911	LEU	2.3
1	A	764	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	987	TYR	2.3
1	A	941	GLY	2.3
1	A	878	THR	2.3
1	A	824	ARG	2.3
1	A	978	PRO	2.2
1	A	710	THR	2.2
1	A	915	GLU	2.2
1	A	742	SER	2.2
1	A	943	GLN	2.2
1	A	924	MET	2.2
1	A	871	HIS	2.2
1	A	719	TYR	2.2
1	A	749	VAL	2.1
1	A	979	GLN	2.1
1	A	932	LEU	2.1
1	A	880	SER	2.1
1	A	927	SER	2.1
1	A	805	PRO	2.1
1	A	867	THR	2.1
1	A	831	LYS	2.1
1	A	1039	GLY	2.1
1	A	781	VAL	2.1
1	A	861	THR	2.1
1	A	940	LYS	2.1
1	A	728	TYR	2.1
1	A	938	ASN	2.0
1	A	1040	TYR	2.0
1	A	765	VAL	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 oligosaccharides in this entry.

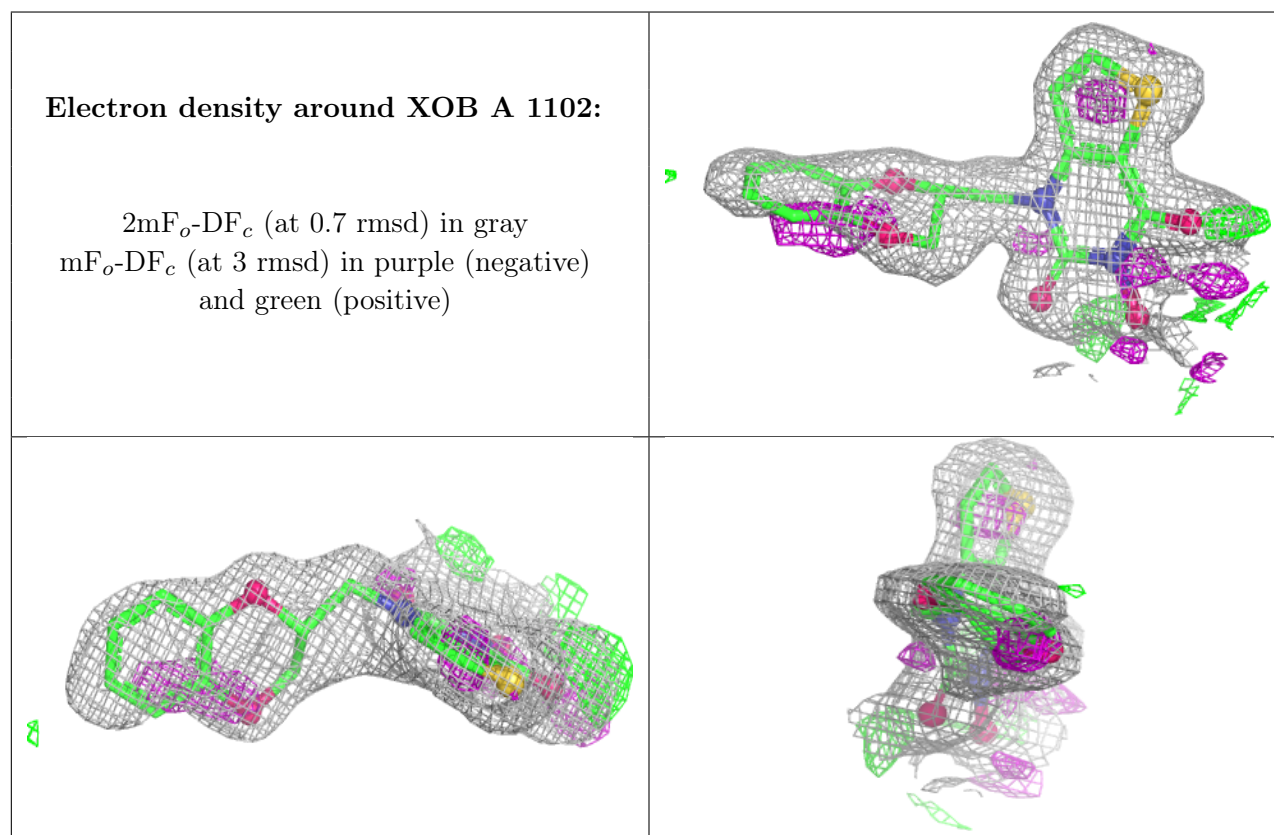
6.4 Ligands [i](#)

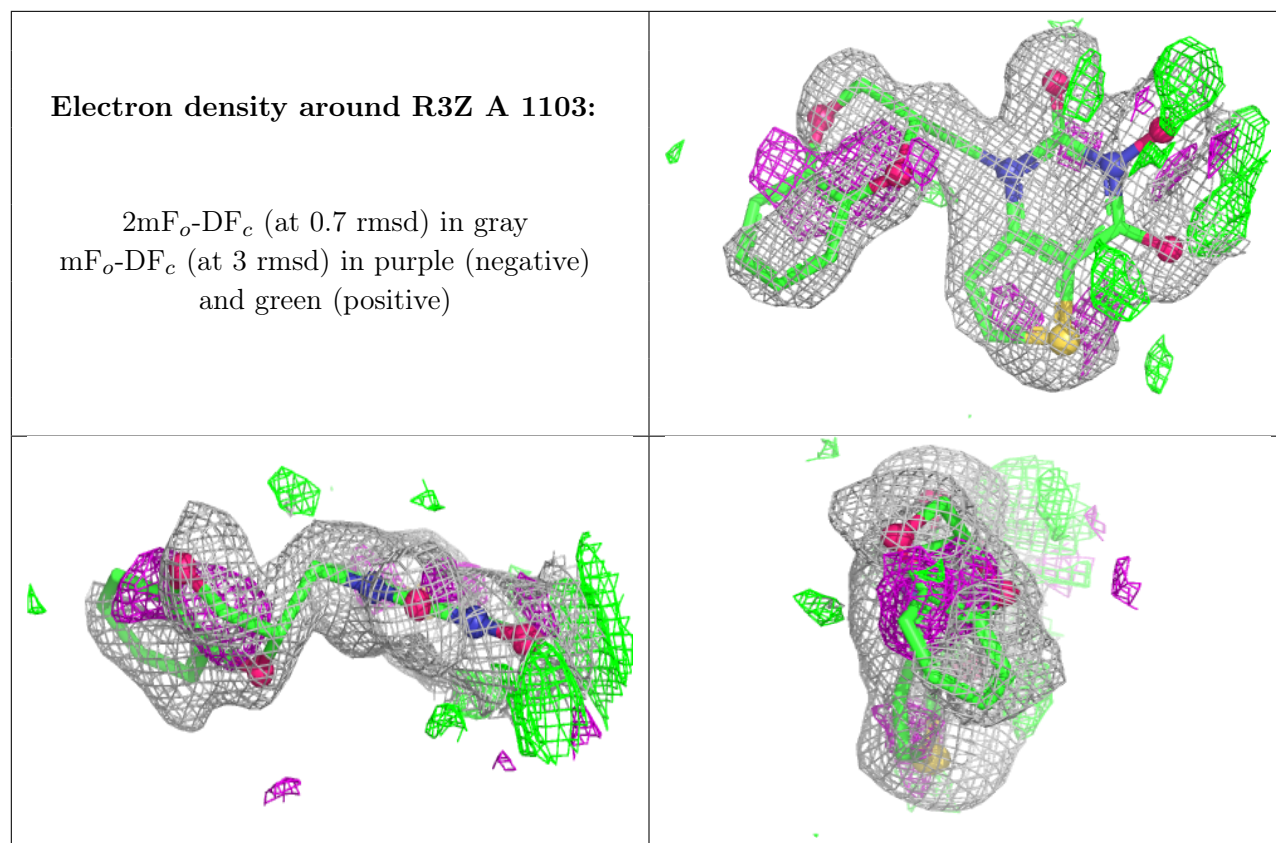
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,

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($\text{\AA}^2$)	Q<0.9
6	SO4	A	1105	5/5	0.85	0.12	40,44,52,52	0
3	XOB	A	1102	23/23	0.86	0.13	33,44,52,53	0
4	R3Z	A	1103	23/23	0.88	0.15	23,44,61,69	0
5	ZN	A	1104	1/1	0.96	0.06	28,28,28,28	0
2	NI	A	1101	1/1	0.98	0.04	21,21,21,21	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 ⓘ

There are no such residues in this entry.