WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

Jun 24, 2024 – 07:22 PM EDT

PDB ID	:	6TN0
Title	:	Rapid optimisation of fragments and hits to lead compounds from screening of crude reaction mixtures
Authors	:	Baker, L.M.; Aimon, A.; Murray, J.B.; Surgenor, A.E.; Matassova, N.; Roughley, S.D.; von Delft, F.; Hubbard, R.E.
Deposited on	:	2019-12-05
Resolution	:	1.91 Å(reported)

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

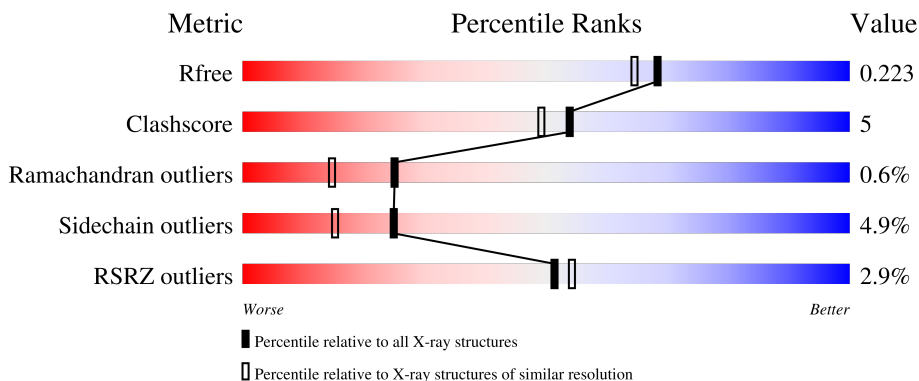|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.91 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	Percentile Ranks	Value
Rfree		0.223
Clashscore		5
Ramachandran outliers		0.6%
Sidechain outliers		4.9%
RSRZ outliers		2.9%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	AAA	394	3% / 74% 11% · 13%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 5658 atoms, of which 2745 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 [Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 2, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	341	Total 5433	C 1754	H 2707	N 444	O 512	S 16	142	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	6	GLY	ALA	conflict	UNP Q15119
AAA	7	SER	GLY	conflict	UNP Q15119

- Molecule 2 is N-(2-AMINOETHYL)-2-{3-CHLORO-4-[(4-ISOPROPYLBENZYL)OXY]PHENYL} ACETAMIDE (three-letter code: TF3) (formula: $C_{20}H_{25}ClN_2O_2$).
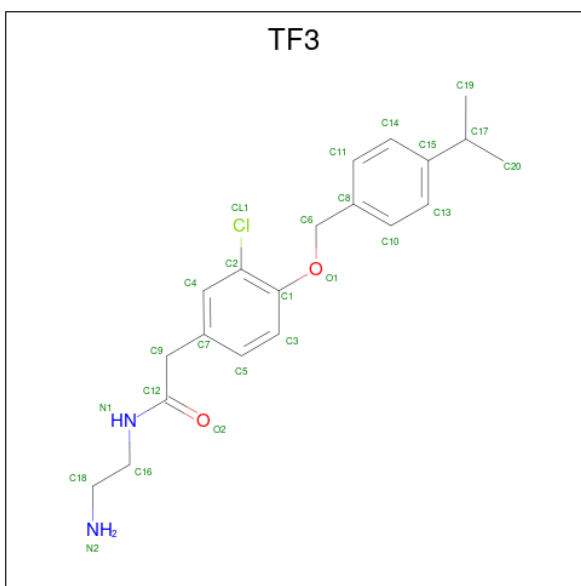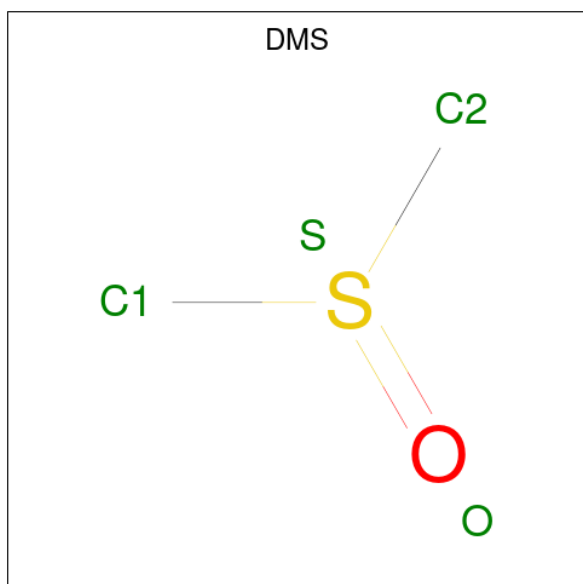

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	AAA	1	Total 50	C 20	Cl 1	H 25	N 2	O 2	0	0

- Molecule 3 is 2,4-bis(oxidanyl)benzamide (three-letter code: NLZ) (formula: $C_7H_7NO_3$) (labeled as "Ligand of Interest" by depositor).
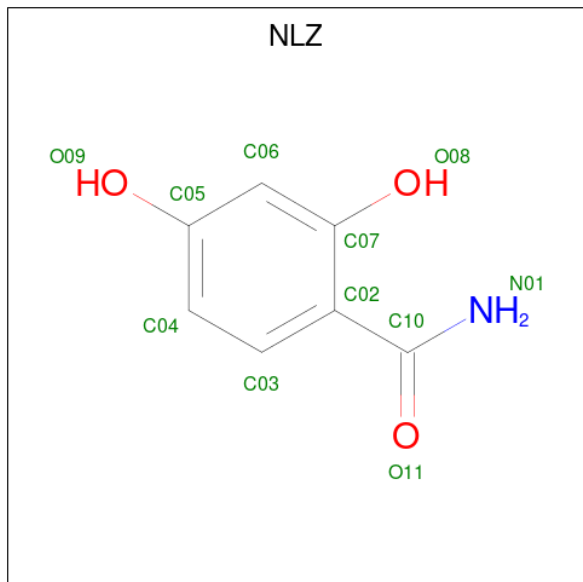
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	AAA	1	Total	C	H	N	O	2	0
			18	7	7	1	3		

- Molecule 4 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	AAA	1	Total	C	H	O	S	0	0
			10	2	6	1	1		

• Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
5	AAA	1	1	1	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	AAA	146	146	146	0	0

3 Residue-property plots ⓘ

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: [Pyruvate dehydrogenase (acetyl-transferring)] kinase isozyme 2, mitochondrial
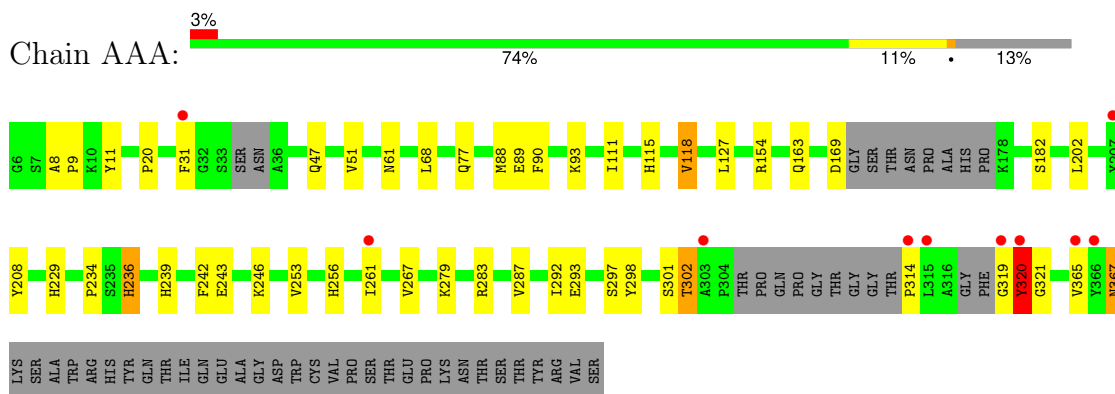
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α, β, γ	109.05Å 109.05Å 83.94Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	94.44 – 1.91 94.44 – 1.91	Depositor EDS
% Data completeness (in resolution range)	62.6 (94.44-1.91) 62.6 (94.44-1.91)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 1.91Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.176 , 0.221 0.182 , 0.223	Depositor DCC
R_{free} test set	1332 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	43.1	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 42.5	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.040 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5658	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.65% 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.

Full wwPDB X-ray Structure Validation Report

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: DMS, CL, NLZ, TF3

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	AAA	0.71	1/2791 (0.0%)	0.83	0/3780

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	243	GLU	CD-OE1	7.12	1.33	1.25

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	AAA	2726	2707	2682	27	0
2	AAA	25	25	25	0	0
3	AAA	11	7	0	2	0
4	AAA	4	6	6	1	0
5	AAA	1	0	0	0	0
6	AAA	146	0	0	6	0
All	All	2913	2745	2713	29	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 5.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

All (29) 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:AAA:163:GLN:HE22	1:AAA:182:SER:H	1.15	0.93
1:AAA:163:GLN:NE2	1:AAA:182:SER:H	1.87	0.73
1:AAA:298:TYR:O	1:AAA:302:THR:OG1	2.08	0.70
1:AAA:229:HIS:HB2	6:AAA:622:HOH:O	1.97	0.64
1:AAA:163:GLN:HE22	1:AAA:182:SER:N	1.93	0.60
1:AAA:20:PRO:HD2	1:AAA:365:VAL:HG12	1.86	0.57
1:AAA:118:VAL:HG11	1:AAA:154:ARG:CD	2.36	0.56
1:AAA:319:GLY:O	1:AAA:320:TYR:HB2	2.08	0.53
1:AAA:367:ASN:C	1:AAA:367:ASN:HD22	2.15	0.50
3:AAA:402:NLZ:N01	6:AAA:501:HOH:O	2.35	0.50
1:AAA:11:TYR:CG	1:AAA:88:MET:CE	2.94	0.50
1:AAA:242:PHE:CZ	1:AAA:246:LYS:HD2	2.49	0.47
1:AAA:297:SER:O	1:AAA:301:SER:OG	2.23	0.47
1:AAA:267:VAL:HA	1:AAA:279:LYS:O	2.15	0.46
1:AAA:111:ILE:O	1:AAA:115:HIS:HD2	1.99	0.45
1:AAA:182:SER:O	1:AAA:234:PRO:HD2	2.16	0.45
1:AAA:292:ILE:HG21	4:AAA:403:DMS:C2	2.47	0.45
1:AAA:61:ASN:OD1	1:AAA:77:GLN:NE2	2.50	0.44
1:AAA:321:GLY:HA2	6:AAA:627:HOH:O	2.18	0.44
1:AAA:11:TYR:CG	1:AAA:88:MET:HE3	2.53	0.44
1:AAA:88:MET:O	1:AAA:90:PHE:N	2.51	0.44
1:AAA:11:TYR:CD1	1:AAA:88:MET:HE3	2.53	0.43
1:AAA:256:HIS:HE1	1:AAA:283:ARG:O	2.01	0.43
1:AAA:208:TYR:CG	1:AAA:253:VAL:HG11	2.54	0.43
1:AAA:8:ALA:HB3	1:AAA:9:PRO:HD3	2.01	0.42
1:AAA:236:HIS:HD2	6:AAA:519:HOH:O	2.03	0.42
1:AAA:239:HIS:NE2	6:AAA:503:HOH:O	2.37	0.41
3:AAA:402:NLZ:C10	6:AAA:501:HOH:O	2.68	0.41
1:AAA:47:GLN:O	1:AAA:51:VAL:HG23	2.22	0.40

There are no symmetry-related clashes.

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	AAA	332/394 (84%)	320 (96%)	10 (3%)	2 (1%)	25	15

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	320	TYR
1	AAA	89	GLU

5.3.2 Protein sidechains ⓘ

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	AAA	305/353 (86%)	290 (95%)	15 (5%)	25	15

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	31	PHE
1	AAA	68	LEU
1	AAA	93	LYS
1	AAA	118	VAL
1	AAA	127	LEU
1	AAA	169	ASP
1	AAA	202	LEU
1	AAA	236	HIS
1	AAA	261	ILE
1	AAA	287	VAL
1	AAA	293	GLU
1	AAA	302	THR
1	AAA	314	PRO
1	AAA	320	TYR
1	AAA	367	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 1 is 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	DMS	AAA	403	-	3,3,3	0.29	0	3,3,3	0.32	0
3	NLZ	AAA	402	-	11,11,11	2.41	1 (9%)	15,15,15	1.15	2 (13%)
2	TF3	AAA	401	-	26,26,26	0.32	0	33,34,34	0.69	1 (3%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NLZ	AAA	402	-	-	2/4/4/4	0/1/1/1
2	TF3	AAA	401	-	-	6/17/17/17	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	AAA	402	NLZ	C02-C10	-7.87	1.39	1.50

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	402	NLZ	O08-C07-C02	-2.77	116.57	121.71
3	AAA	402	NLZ	C02-C10-N01	-2.11	115.02	118.29
2	AAA	401	TF3	O1-C6-C8	2.02	115.09	109.16

There are no chirality outliers.

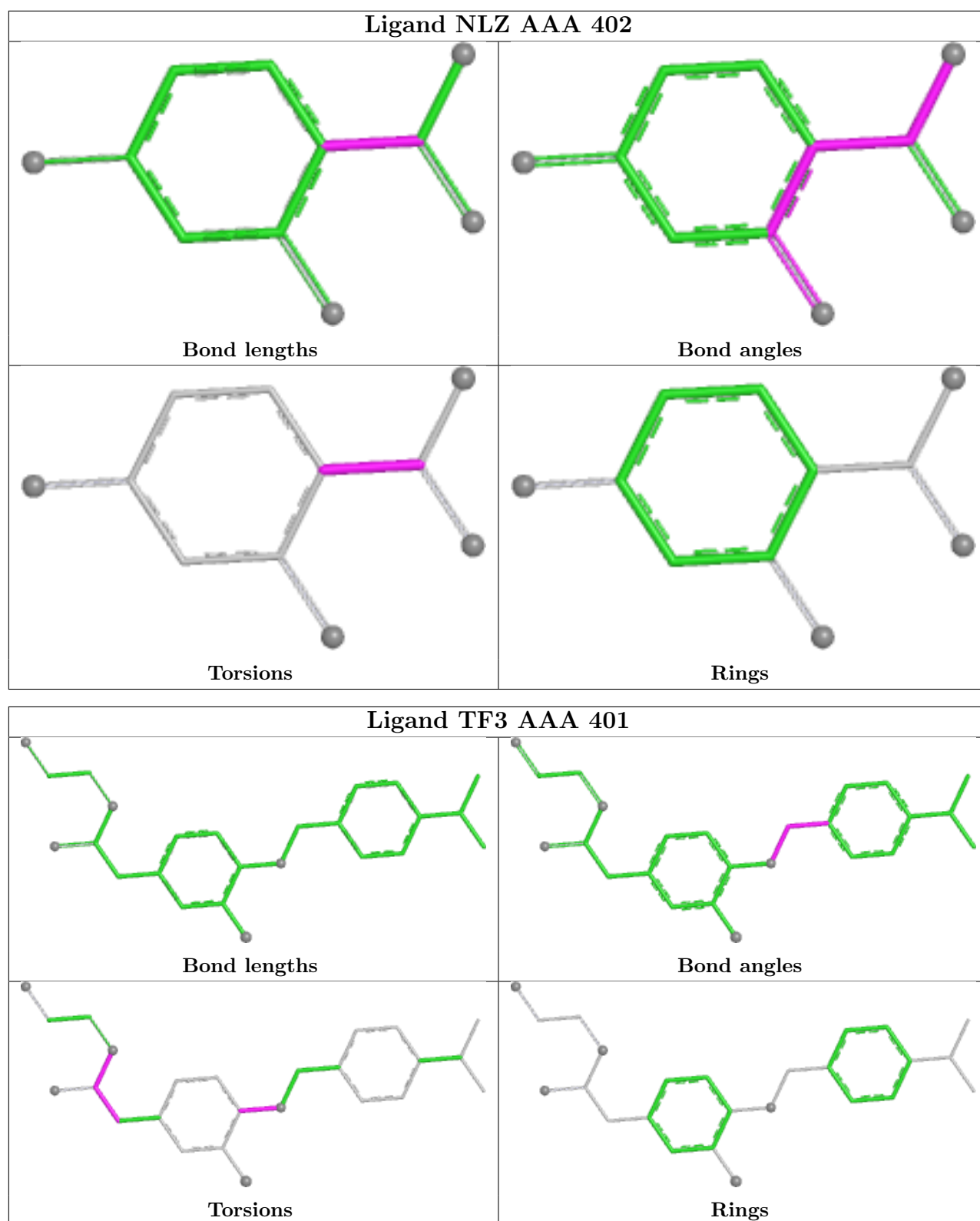All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	AAA	401	TF3	C2-C1-O1-C6
2	AAA	401	TF3	O2-C12-N1-C16
2	AAA	401	TF3	O2-C12-C9-C7
2	AAA	401	TF3	N1-C12-C9-C7
2	AAA	401	TF3	C9-C12-N1-C16
3	AAA	402	NLZ	C07-C02-C10-O11
3	AAA	402	NLZ	C07-C02-C10-N01
2	AAA	401	TF3	C3-C1-O1-C6

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	AAA	403	DMS	1	0
3	AAA	402	NLZ	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 NLZ AAA 402

Bond lengths

Bond angles

Torsions

Rings

Ligand TF3 AAA 401

Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

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	AAA	341/394 (86%)	0.32	10 (2%)	51	54	30, 47, 78, 117	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	314	PRO	4.1
1	AAA	365	VAL	3.8
1	AAA	261	ILE	3.8
1	AAA	366	TYR	3.4
1	AAA	320	TYR	3.3
1	AAA	207	TYR	3.0
1	AAA	303	ALA	2.9
1	AAA	31	PHE	2.7
1	AAA	319	GLY	2.3
1	AAA	315	LEU	2.0

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

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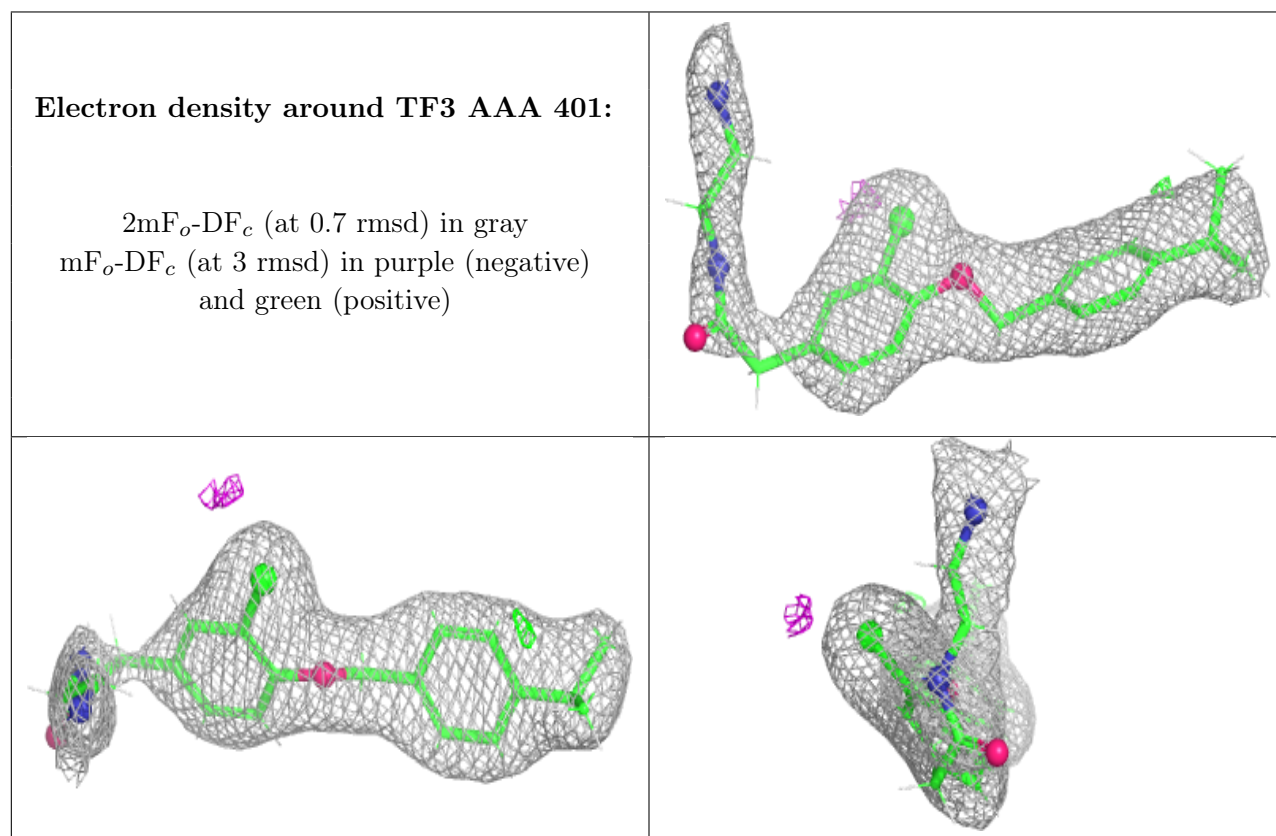
6.3   Carbohydrates   ⓘ

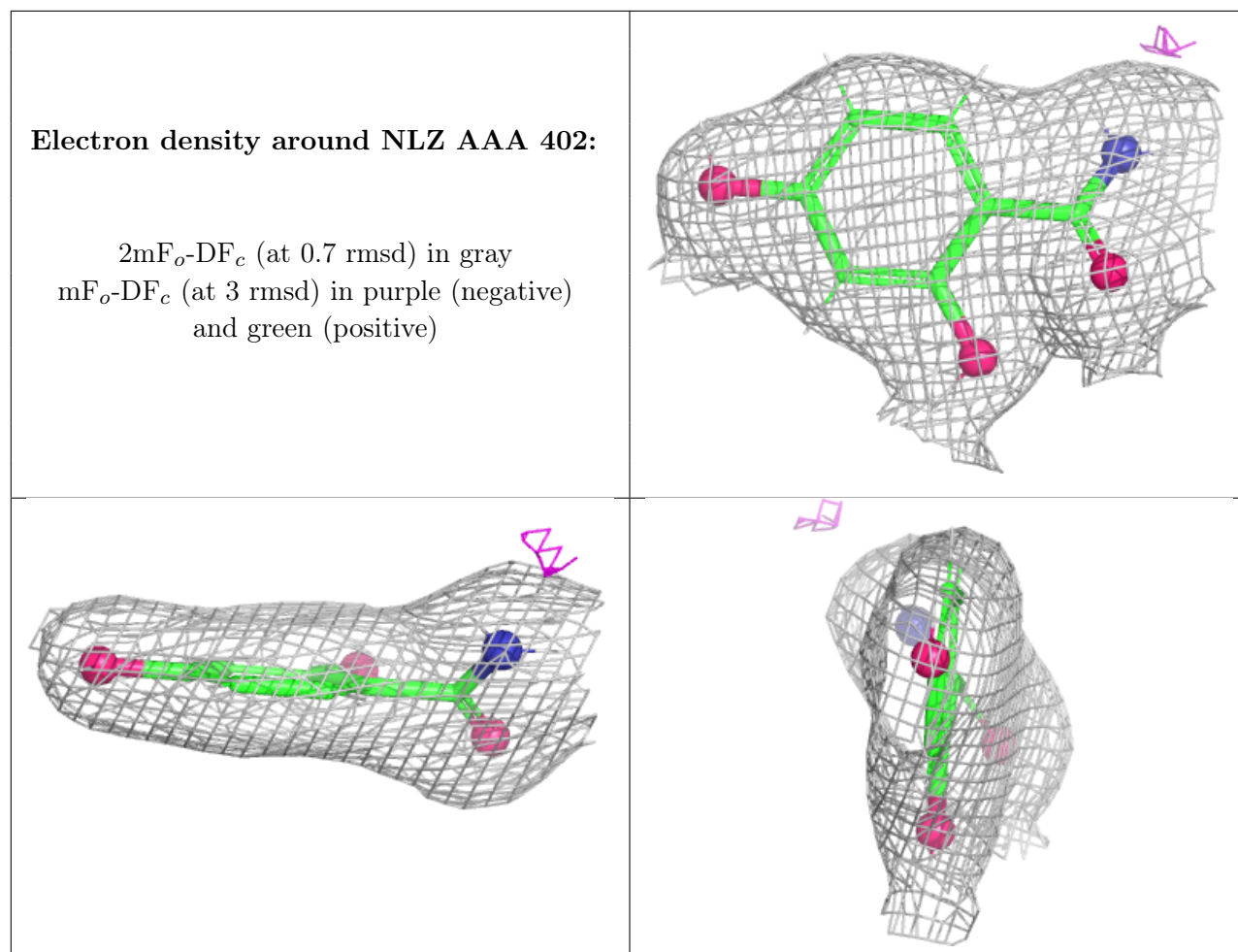There are no monosaccharides in this entry.

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, 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
2	TF3	AAA	401	25/25	0.94	0.18	39,63,97,104	0
3	NLZ	AAA	402	11/11	0.96	0.13	37,44,46,50	2
5	CL	AAA	404	1/1	0.97	0.24	77,77,77,77	0
4	DMS	AAA	403	4/4	0.98	0.19	44,49,63,63	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 TF3 AAA 401:

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

Electron density around NLZ AAA 402:

$2mF_o$-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.