WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 7, 2023 – 11:36 PM EDT

PDB ID	:	4QEB
Title	:	Dcps in complex with covalent inhibitor targeting Tyrosine
Authors	:	Liu, S.
Deposited on	:	2014-05-15
Resolution	:	3.21 Å(reported)

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

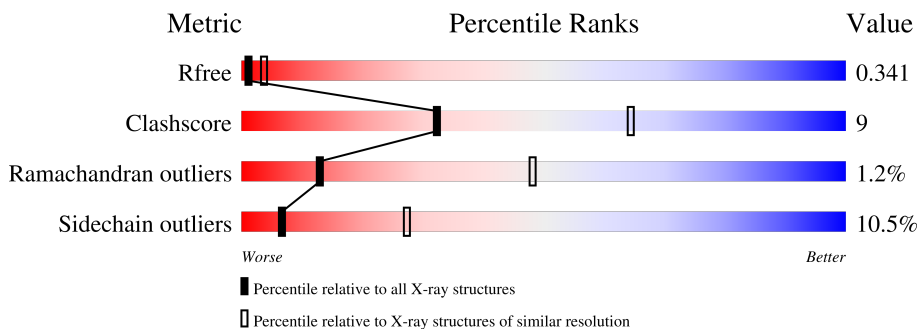|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 3.21 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		0.341
Clashscore		9
Ramachandran outliers		1.2%
Sidechain outliers		10.5%

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	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)

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%

Mol	Chain	Length	Quality of chain
1	A	337	58%　　25%　•　15%
1	B	337	55%　　28%　•　15%
1	C	337	56%　　27%　•　14%
1	D	337	61%　　25%　•　13%

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	PO4	B	402	-	-	X	-

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9617 atoms, of which 13 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 m7GpppX diphosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	287	Total	C	N	O	S	0	0	0
			2377	1523	425	426	3			
1	B	288	Total	C	N	O	S	0	0	0
			2384	1526	427	428	3			
1	C	289	Total	C	N	O	S	0	0	0
			2389	1529	427	430	3			
1	D	292	Total	C	N	O	S	0	0	0
			2411	1542	430	436	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP Q96C86
B	1	GLY	-	expression tag	UNP Q96C86
C	1	GLY	-	expression tag	UNP Q96C86
D	1	GLY	-	expression tag	UNP Q96C86

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
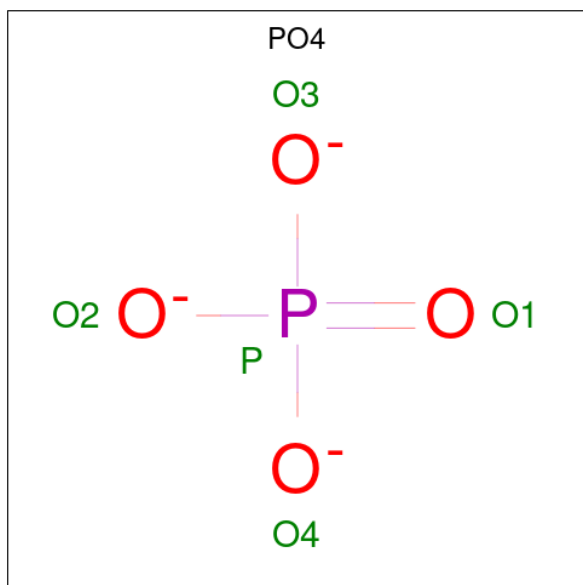
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	P 1	0	0
2	A	1	Total 5	O 4	P 1	0	0
2	B	1	Total 5	O 4	P 1	0	0
2	C	1	Total 5	O 4	P 1	0	0

- Molecule 3 is 3-{[(2,4-diaminoquinazolin-5-yl)oxy]methyl}benzenesulfonic acid (three-letter code: 31G) (formula: $C_{15}H_{14}N_4O_4S$).
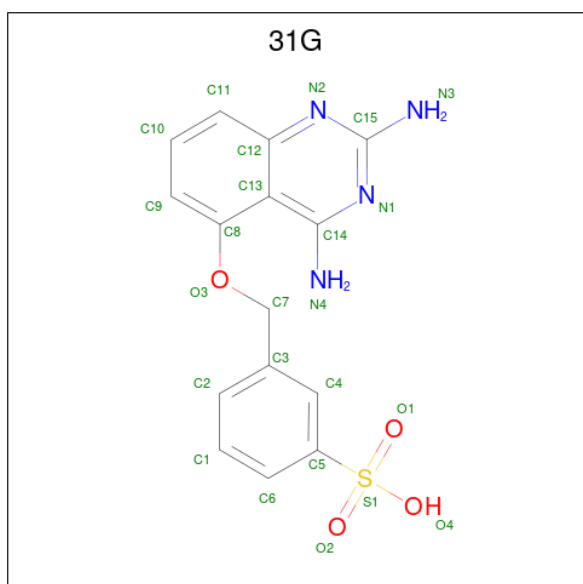
| Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	H	N	O	S		
3	B	1	36	15	13	4	3	1	13	0

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. 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. 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.
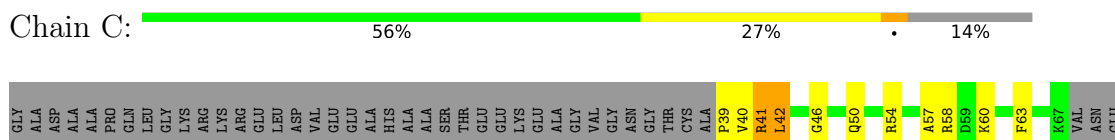
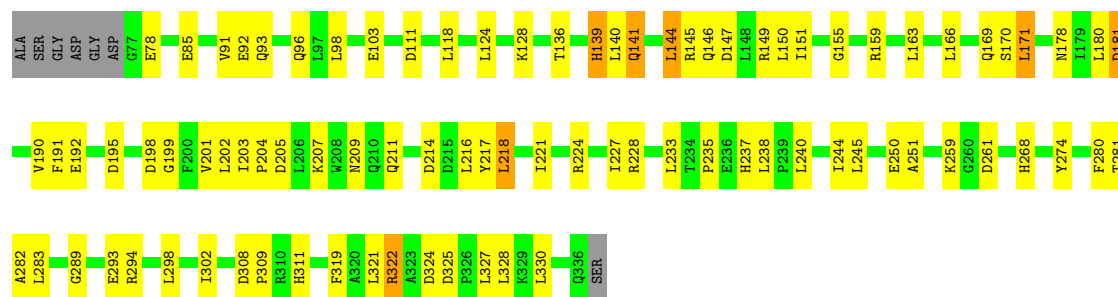- Molecule 1: m7GpppX diphosphatase
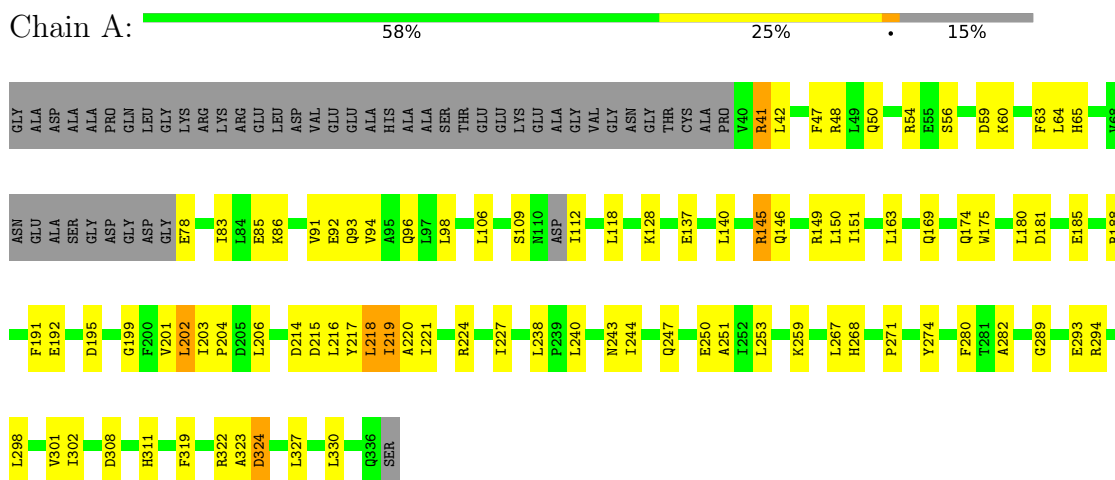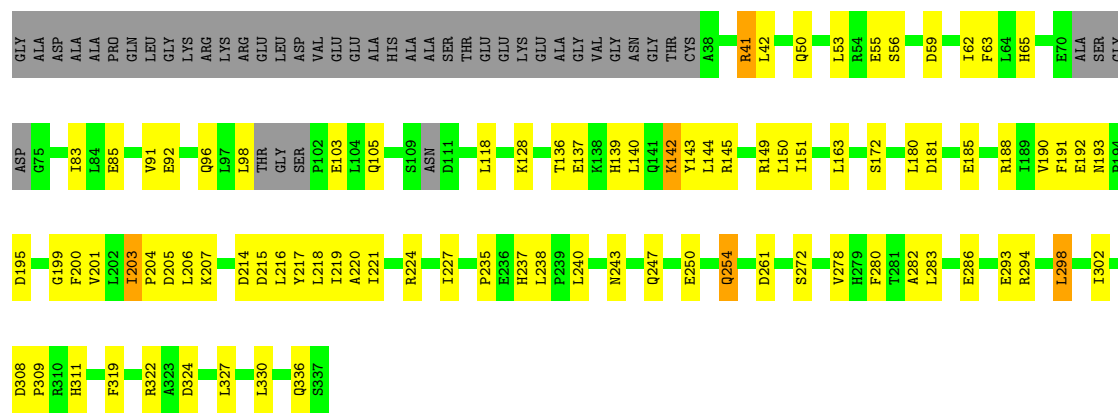
Chain A:

- Molecule 1: m7GpppX diphosphatase
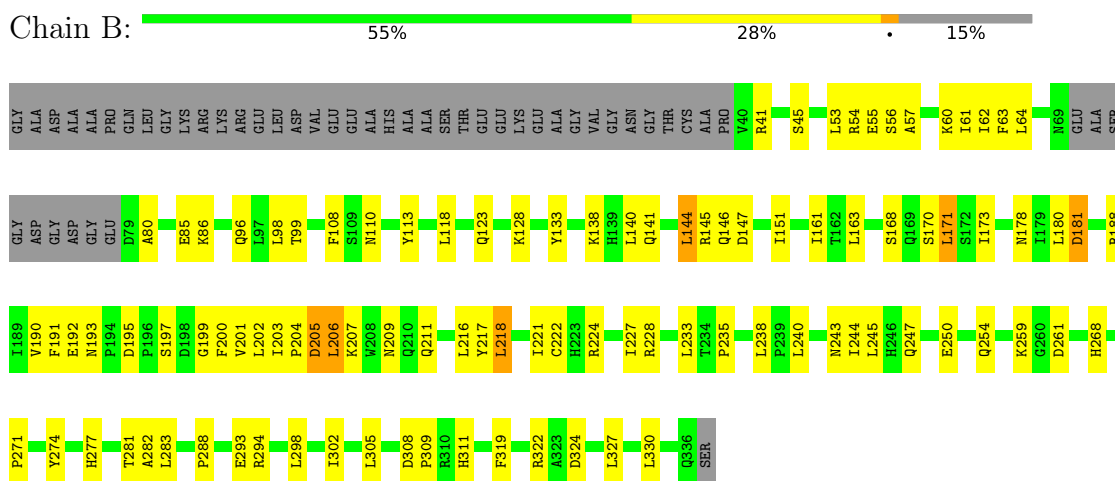
Chain B:

- Molecule 1: m7GpppX diphosphatase

Chain C:

• Molecule 1: m7GpppX diphosphatase

Chain D:

4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	101.25Å 105.23Å 139.96Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.89 – 3.21 82.03 – 3.21	Depositor EDS
% Data completeness (in resolution range)	83.2 (33.89-3.21) 83.4 (82.03-3.21)	Depositor EDS
R_{merge}	0.32	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 3.19Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, R_{free}	0.211 , 0.276 0.276 , 0.341	Depositor DCC
R_{free} test set	1069 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	64.1	Xtriage
Anisotropy	0.752	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 84.6	EDS
L-test for twinning [2]	$< |L| > = 0.35, < L^2 > = 0.19$	Xtriage
Estimated twinning fraction	0.036 for k,h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9617	wwPDB-VP
Average B, all atoms (Å^2)	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.68 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1064e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 31G, PO4

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	0.56	0/2432	0.78	0/3298
1	B	0.54	0/2440	0.75	1/3311 (0.0%)
1	C	0.52	0/2446	0.76	0/3318
1	D	0.53	0/2466	0.77	1/3342 (0.0%)
All	All	0.54	0/9784	0.76	2/13269 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	193	ASN	N-CA-C	-5.07	97.32	111.00
1	B	193	ASN	N-CA-C	-5.03	97.42	111.00

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	2377	0	2384	44	0
1	B	2384	0	2388	55	0
1	C	2389	0	2391	47	0
1	D	2411	0	2406	40	0
2	A	10	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	5	0	0	4	0
2	C	5	0	0	1	0
3	B	23	13	13	2	0
All	All	9604	13	9582	168	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 9.

The worst 5 of 168 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:146:GLN:HG3	1:A:324:ASP:HB3	1.63	0.80
1:A:243:ASN:O	1:A:247:GLN:HB2	1.85	0.74
1:C:57:ALA:HB1	1:D:215:ASP:HB3	1.68	0.74
1:B:277:HIS:HE1	2:B:402:PO4:O4	1.70	0.73
1:B:62:ILE:HD12	1:B:64:LEU:HD11	1.75	0.69

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	A	281/337 (83%)	251 (89%)	27 (10%)	3 (1%)	14	50
1	B	284/337 (84%)	257 (90%)	24 (8%)	3 (1%)	14	50
1	C	285/337 (85%)	260 (91%)	21 (7%)	4 (1%)	11	45
1	D	284/337 (84%)	254 (89%)	26 (9%)	4 (1%)	11	45
All	All	1134/1348 (84%)	1022 (90%)	98 (9%)	14 (1%)	13	48

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	169	GLN
1	A	91	VAL
1	A	174	GLN
1	C	46	GLY
1	C	91	VAL

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	A	261/294 (89%)	234 (90%)	27 (10%)	7	28
1	B	262/294 (89%)	234 (89%)	28 (11%)	6	26
1	C	262/294 (89%)	233 (89%)	29 (11%)	6	25
1	D	264/294 (90%)	238 (90%)	26 (10%)	8	30
All	All	1049/1176 (89%)	939 (90%)	110 (10%)	7	27

5 of 110 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	50	GLN
1	C	170	SER
1	D	324	ASP
1	D	214	ASP
1	C	78	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	146	GLN
1	D	254	GLN
1	D	249	GLN
1	C	110	ASN
1	D	69	ASN

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 ⓘ

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	PO4	B	402	-	4,4,4	2.68	2 (50%)	6,6,6	1.00	0
2	PO4	C	401	-	4,4,4	1.89	1 (25%)	6,6,6	1.10	1 (16%)
2	PO4	A	402	-	4,4,4	2.49	1 (25%)	6,6,6	0.19	0
2	PO4	A	401	-	4,4,4	2.48	1 (25%)	6,6,6	1.32	1 (16%)
3	31G	B	401	1	22,25,26	0.38	0	33,35,38	1.73	6 (18%)

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	31G	B	401	1	-	2/9/9/11	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	402	PO4	P-O1	4.25	1.60	1.50
2	A	402	PO4	P-O1	4.13	1.60	1.50
2	A	401	PO4	P-O1	4.02	1.60	1.50
2	B	402	PO4	P-O3	2.06	1.60	1.54
2	C	401	PO4	P-O3	2.05	1.60	1.54

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	401	31G	C7-O3-C8	5.20	124.86	117.56
3	B	401	31G	O3-C7-C3	4.20	121.75	109.16
3	B	401	31G	O2-S1-C5	3.48	112.50	104.58
3	B	401	31G	C8-C13-C14	3.43	129.10	127.36
3	B	401	31G	C14-C13-C12	-3.23	112.65	114.86

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	401	31G	C13-C8-O3-C7
3	B	401	31G	C9-C8-O3-C7

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	402	PO4	4	0
2	C	401	PO4	1	0
3	B	401	31G	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 31G B 401

Bond lengths

Bond angles

Torsions

Rings
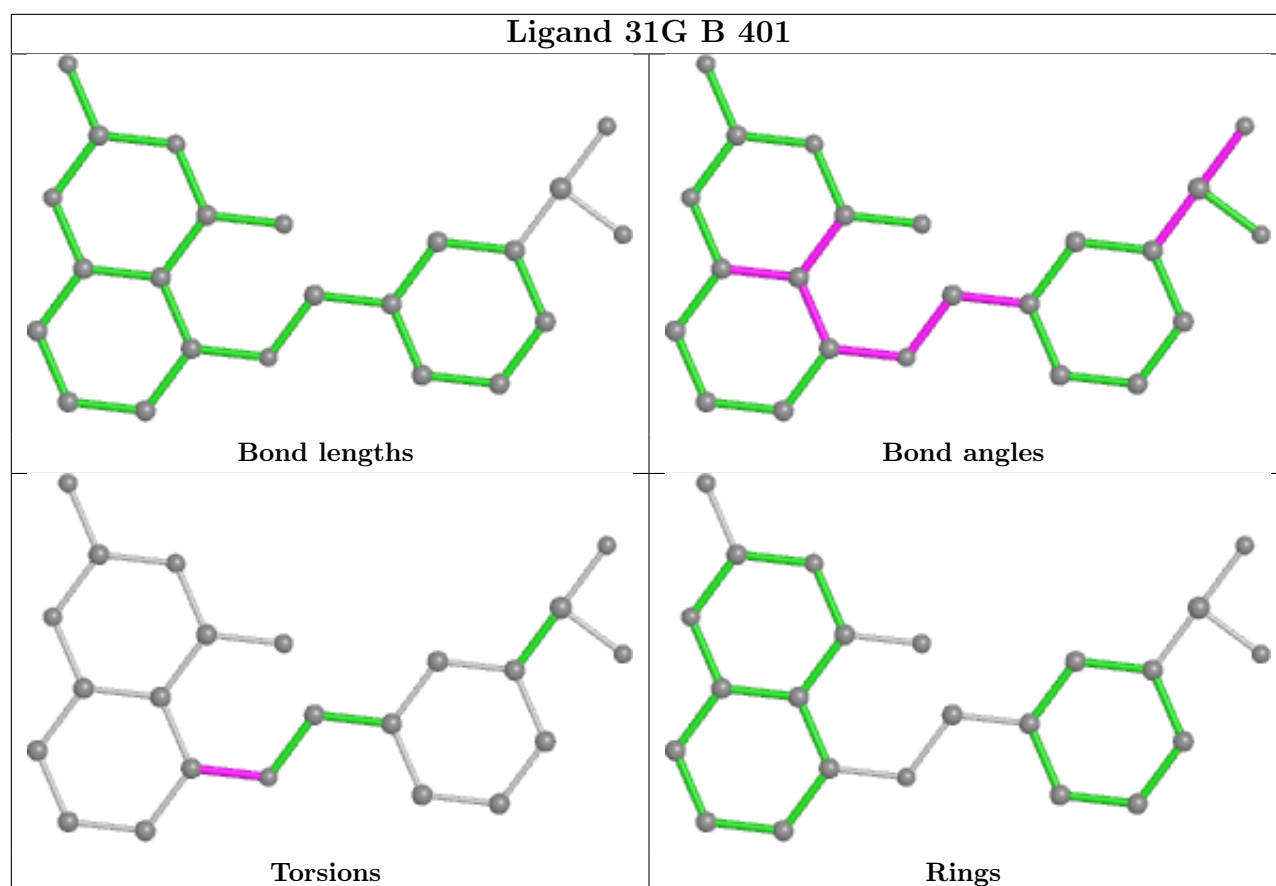
5.7 Other polymers ⓘ

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

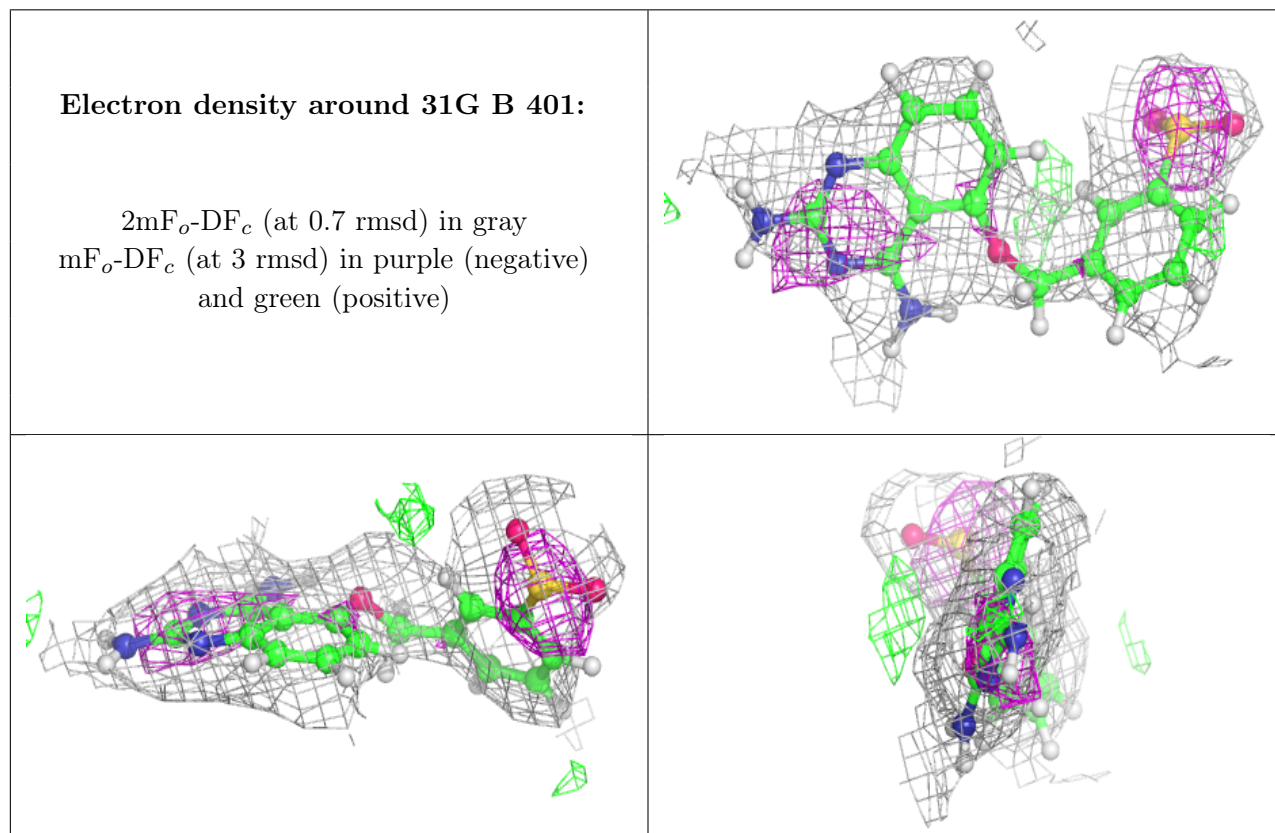Unable to reproduce the depositors R factor - this section is therefore empty.

6.3   Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4   Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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 31G B 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)

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.