WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Jan 15, 2024 – 02:35 pm GMT

PDB ID	:	6TP6
Title	:	Crystal structure of the Orexin-1 receptor in complex with filorexant
Authors	:	Rappas, M.; Ali, A.; Bennett, K.A.; Brown, J.D.; Bucknell, S.J.; Congreve, M.; Cooke, R.M.; Cseke, G.; de Graaf, C.; Dore, A.S.; Errey, J.C.; Jazayeri, A.; Marshall, F.H.; Mason, J.S.; Mould, R.; Patel, J.C.; Tehan, B.G.; Weir, M.; Christopher, J.A.
Deposited on	:	2019-12-12
Resolution	:	2.34 Å(reported)

The following versions of software and data (see references ①) 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 ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.34 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		0.262
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		0.2%
RSRZ outliers		16.1%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution
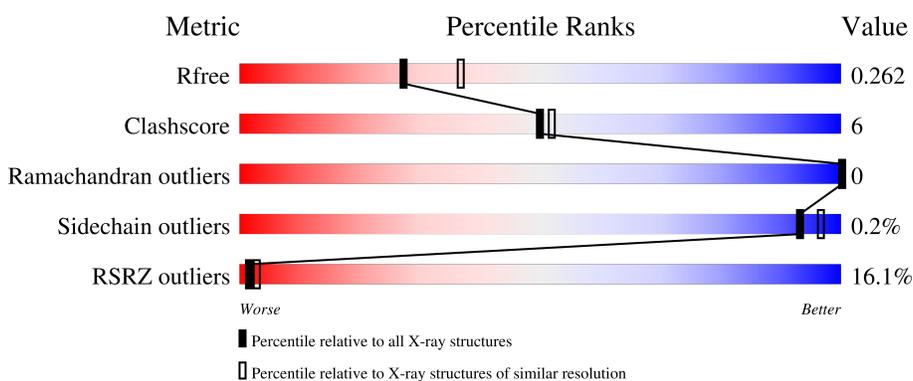
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	A	368	12% 73% 10% 17%
1	B	368	14% 70% 12% 18%

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:

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	PGW	A	411	X	-	-	-

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 5301 atoms, of which 50 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 Orexin receptor type 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	305	Total 2447	C 1627	N 402	O 401	S 17	0	0	0
1	B	303	Total 2404	C 1592	N 403	O 392	S 17	0	0	0

There are 52 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	25	ALA	-	expression tag	UNP O43613
A	26	ALA	-	expression tag	UNP O43613
A	27	SER	-	expression tag	UNP O43613
A	46	ALA	GLU	engineered mutation	UNP O43613
A	85	LEU	ILE	engineered mutation	UNP O43613
A	95	ALA	VAL	engineered mutation	UNP O43613
A	162	LEU	ARG	engineered mutation	UNP O43613
A	194	ALA	ASN	engineered mutation	UNP O43613
A	198	ALA	LEU	engineered mutation	UNP O43613
A	211	ALA	TYR	engineered mutation	UNP O43613
A	304	VAL	LEU	engineered mutation	UNP O43613
A	339	ALA	CYS	engineered mutation	UNP O43613
A	375	TRP	CYS	engineered mutation	UNP O43613
A	376	TRP	CYS	engineered mutation	UNP O43613
A	381	ALA	-	expression tag	UNP O43613
A	382	ALA	-	expression tag	UNP O43613
A	383	ALA	-	expression tag	UNP O43613
A	384	HIS	-	expression tag	UNP O43613
A	385	HIS	-	expression tag	UNP O43613
A	386	HIS	-	expression tag	UNP O43613
A	387	HIS	-	expression tag	UNP O43613
A	388	HIS	-	expression tag	UNP O43613
A	389	HIS	-	expression tag	UNP O43613
A	390	HIS	-	expression tag	UNP O43613
A	391	HIS	-	expression tag	UNP O43613

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	HIS	-	expression tag	UNP O43613
B	25	ALA	-	expression tag	UNP O43613
B	26	ALA	-	expression tag	UNP O43613
B	27	SER	-	expression tag	UNP O43613
B	46	ALA	GLU	engineered mutation	UNP O43613
B	85	LEU	ILE	engineered mutation	UNP O43613
B	95	ALA	VAL	engineered mutation	UNP O43613
B	162	LEU	ARG	engineered mutation	UNP O43613
B	194	ALA	ASN	engineered mutation	UNP O43613
B	198	ALA	LEU	engineered mutation	UNP O43613
B	211	ALA	TYR	engineered mutation	UNP O43613
B	304	VAL	LEU	engineered mutation	UNP O43613
B	339	ALA	CYS	engineered mutation	UNP O43613
B	375	TRP	CYS	engineered mutation	UNP O43613
B	376	TRP	CYS	engineered mutation	UNP O43613
B	381	ALA	-	expression tag	UNP O43613
B	382	ALA	-	expression tag	UNP O43613
B	383	ALA	-	expression tag	UNP O43613
B	384	HIS	-	expression tag	UNP O43613
B	385	HIS	-	expression tag	UNP O43613
B	386	HIS	-	expression tag	UNP O43613
B	387	HIS	-	expression tag	UNP O43613
B	388	HIS	-	expression tag	UNP O43613
B	389	HIS	-	expression tag	UNP O43613
B	390	HIS	-	expression tag	UNP O43613
B	391	HIS	-	expression tag	UNP O43613
B	392	HIS	-	expression tag	UNP O43613

- Molecule 2 is [(2 {R},5 {R})-5-[(5-fluoranylpyridin-2-yl)oxymethyl]-2-methyl-piperidin-1-yl]-(5-methyl-2-pyrimidin-2-yl-phenyl)methanone (three-letter code: NT5) (formula: $C_{24}H_{25}FN_4O_2$) (labeled as "Ligand of Interest" by depositor).
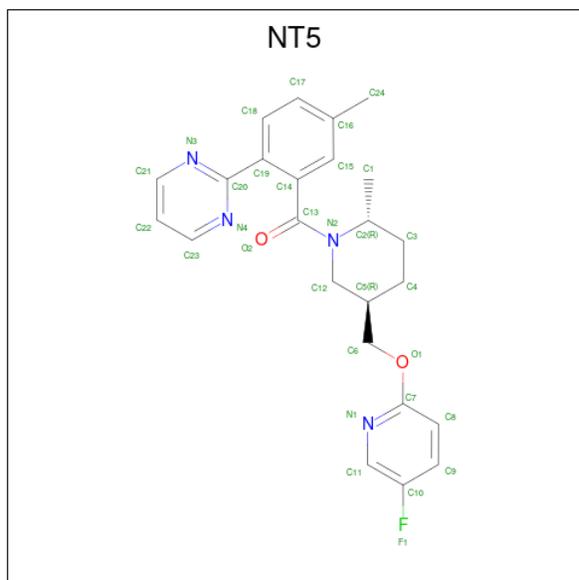
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	F	H	N	O	0	0
			56	24	1	25	4	2		
2	B	1	Total	C	F	H	N	O	0	0
			56	24	1	25	4	2		

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: $C_8H_{18}O_5$).
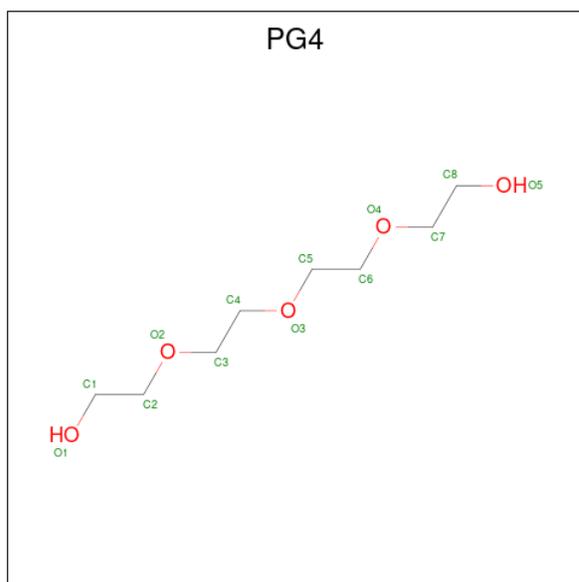
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	8	5		
3	B	1	Total	C	O	0	0
			9	6	3		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).
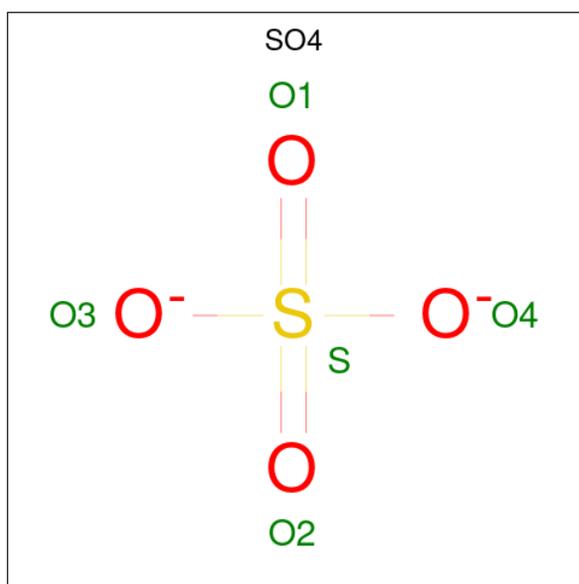
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 5	O 4	S 1	0	0
4	A	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0
4	B	1	Total 5	O 4	S 1	0	0

- Molecule 5 is octyl 1-thio-beta-D-glucopyranoside (three-letter code: SOG) (formula: $C_{14}H_{28}O_5S$).
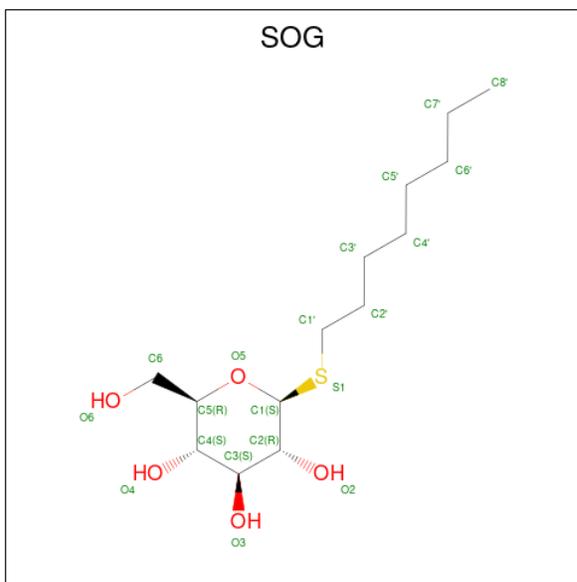
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 20	C 14	O 5	S 1	0	0
5	A	1	Total 20	C 14	O 5	S 1	0	0
5	A	1	Total 13	C 7	O 5	S 1	0	0
5	A	1	Total 20	C 14	O 5	S 1	0	0
5	A	1	Total 4	C 4			0	0
5	B	1	Total 20	C 14	O 5	S 1	0	0
5	B	1	Total 20	C 14	O 5	S 1	0	0
5	B	1	Total 20	C 14	O 5	S 1	0	0
5	B	1	Total 6	C 5	S 1		0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total 1	Cl 1	0	0

- Molecule 7 is (1R)-2-{[(S)-{[(2S)-2,3-dihydroxypropyl]oxy}(hydroxy)phosphoryl]oxy}-1-[(hexadecanoyloxy)methyl]ethyl (9Z)-octadec-9-enoate (three-letter code: PGW) (formula: $C_{40}H_{77}O_{10}P$).
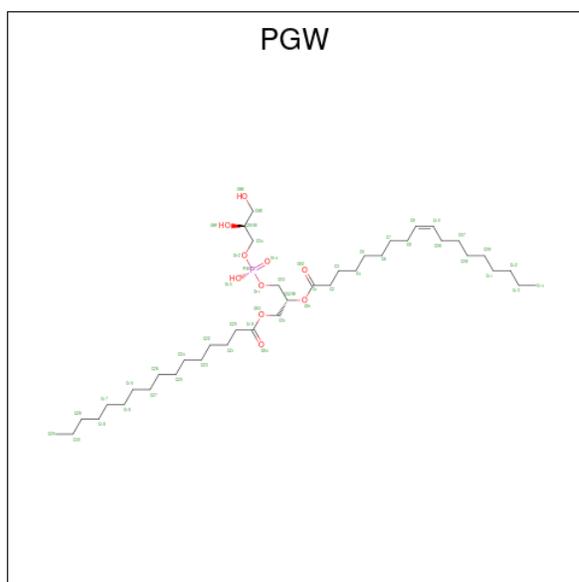
PGW

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total 51	C 40	O 10	P 1	0	0
7	B	1	Total 51	C 40	O 10	P 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	20	Total 20	O 20	0	0
8	B	30	Total 30	O 30	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: Orexin receptor type 1
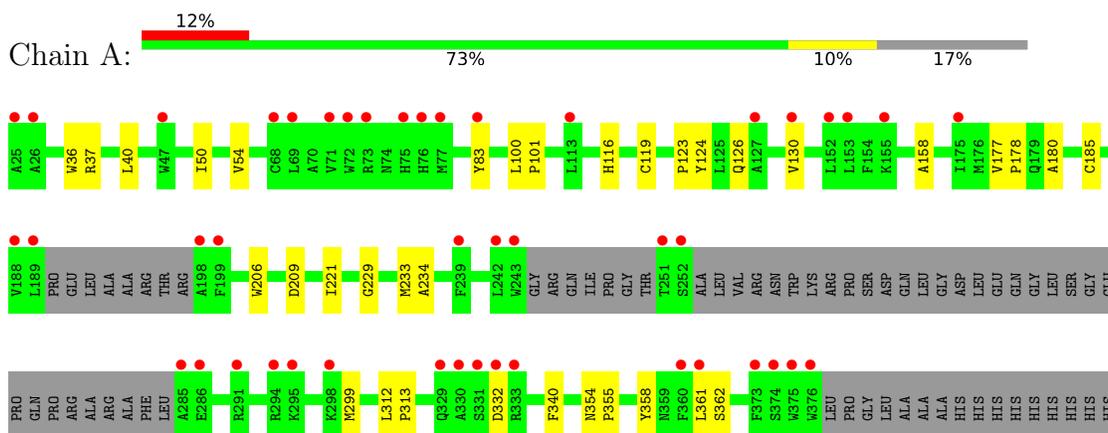
- Molecule 1: Orexin receptor type 1
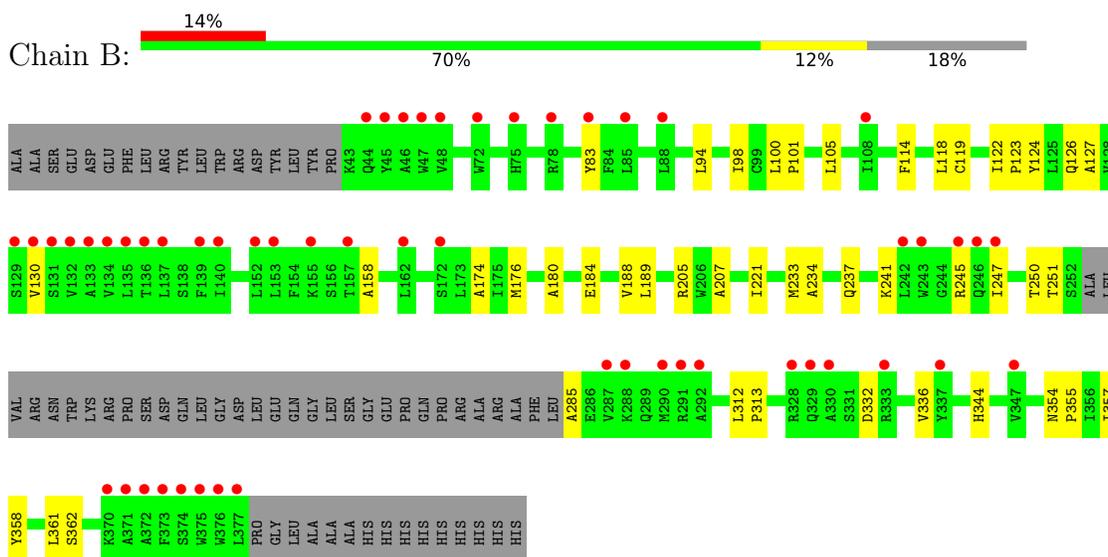
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.89Å 146.20Å 72.33Å 90.00° 111.12° 90.00°	Depositor
Resolution (Å)	34.99 – 2.34 34.99 – 2.33	Depositor EDS
% Data completeness (in resolution range)	60.0 (34.99-2.34) 81.3 (34.99-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.27 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.211 , 0.244 0.230 , 0.262	Depositor DCC
R_{free} test set	2472 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	52.0	Xtriage
Anisotropy	0.150	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 57.3	EDS
L-test for twinning [2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	5301	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.43% 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.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CL, SOG, NT5, PGW, PG4, SO4

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	0.26	0/2516	0.41	0/3429
1	B	0.25	0/2470	0.41	0/3368
All	All	0.25	0/4986	0.41	0/6797

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	2447	0	2489	25	0
1	B	2404	0	2474	36	0
2	A	31	25	0	0	0
2	B	31	25	0	1	0
3	A	13	0	18	2	0
3	B	9	0	11	1	0
4	A	10	0	0	0	0
4	B	10	0	0	0	0
5	A	77	0	102	2	0
5	B	66	0	92	1	0
6	A	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	51	0	76	5	0
7	B	51	0	76	8	0
8	A	20	0	0	0	0
8	B	30	0	0	0	0
All	All	5251	50	5338	61	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 6.

All (61) 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:B:234:ALA:HB2	7:B:409:PGW:H7	1.46	0.96
1:A:234:ALA:HB2	7:A:411:PGW:H7	1.52	0.90
1:B:184:GLU:HG2	1:B:205:ARG:HG2	1.55	0.86
1:B:237:GLN:HG2	7:B:409:PGW:H05	1.71	0.71
1:B:105:LEU:HD12	1:B:114:PHE:HZ	1.55	0.70
1:A:116:HIS:CD2	1:A:185:CYS:HB3	2.29	0.67
1:B:251:THR:HG22	1:B:285:ALA:H	1.60	0.66
1:B:233:MET:HB3	7:B:409:PGW:H2A	1.79	0.65
7:A:411:PGW:H02	7:A:411:PGW:H04	1.83	0.61
1:B:105:LEU:HD12	1:B:114:PHE:CZ	2.35	0.60
1:B:234:ALA:HB2	7:B:409:PGW:C7	2.25	0.59
1:B:184:GLU:CG	1:B:205:ARG:HG2	2.31	0.59
1:A:233:MET:HB3	7:A:411:PGW:H2A	1.85	0.59
1:A:177:VAL:HB	7:B:409:PGW:H4	1.85	0.58
1:A:83:TYR:CE2	1:A:158:ALA:HB1	2.41	0.56
1:B:94:LEU:HD12	1:B:98:ILE:HD12	1.88	0.56
1:B:83:TYR:CE2	1:B:158:ALA:HB1	2.41	0.55
1:A:126:GLN:O	1:A:130:VAL:HG23	2.06	0.55
1:B:124:TYR:HB2	1:B:180:ALA:HB2	1.89	0.54
1:A:340:PHE:HB2	3:A:402:PG4:H12	1.90	0.53
1:A:124:TYR:HB2	1:A:180:ALA:HB2	1.91	0.53
1:B:241:LYS:HD2	1:B:245:ARG:HH11	1.73	0.52
1:B:130:VAL:HG21	2:B:401:NT5:C23	2.40	0.51
1:A:229:GLY:HA3	7:A:411:PGW:H15	1.93	0.51
1:B:126:GLN:O	1:B:130:VAL:HG23	2.11	0.50
1:A:36:TRP:HA	1:A:40:LEU:HB2	1.92	0.50
1:B:188:VAL:HG13	1:B:189:LEU:N	2.25	0.50
1:A:206:TRP:HE3	5:A:407:SOG:HO3	1.58	0.50
1:B:344:HIS:CE1	3:B:402:PG4:H42	2.47	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:119:CYS:O	1:B:123:PRO:HG2	2.14	0.47
1:B:114:PHE:HB3	1:B:118:LEU:HD12	1.98	0.46
1:A:340:PHE:CB	3:A:402:PG4:H12	2.46	0.46
1:B:354:ASN:HB2	1:B:355:PRO:HD3	1.97	0.46
1:B:234:ALA:CB	7:B:409:PGW:H7	2.33	0.45
7:A:411:PGW:H6	1:B:174:ALA:HB1	1.98	0.45
1:A:354:ASN:HB2	1:A:355:PRO:HD3	1.98	0.45
1:B:127:ALA:HB1	1:B:176:MET:SD	2.56	0.45
1:A:119:CYS:O	1:A:123:PRO:HG2	2.17	0.45
1:A:312:LEU:HB3	1:A:313:PRO:HD3	1.99	0.44
1:B:241:LYS:HA	1:B:245:ARG:HD2	1.98	0.44
1:A:221:ILE:HD13	1:B:221:ILE:HD13	1.99	0.44
1:B:205:ARG:O	1:B:205:ARG:HG3	2.17	0.44
1:B:312:LEU:HB3	1:B:313:PRO:HD3	1.99	0.43
1:B:188:VAL:CG1	1:B:189:LEU:N	2.82	0.43
1:A:358:TYR:O	1:A:362:SER:HB3	2.18	0.43
1:A:299:MET:HA	1:A:361:LEU:HD13	2.01	0.43
1:A:332:ASP:OD1	1:A:332:ASP:N	2.52	0.43
1:B:207:ALA:HB3	5:B:407:SOG:H1	2.00	0.42
1:B:122:ILE:HB	1:B:123:PRO:HD3	2.01	0.42
1:B:247:ILE:O	1:B:250:THR:HG23	2.19	0.42
1:A:116:HIS:HD2	1:A:185:CYS:HB3	1.78	0.42
1:B:357:ILE:HG23	1:B:361:LEU:HD12	2.02	0.42
1:B:358:TYR:O	1:B:362:SER:HB3	2.20	0.42
7:B:409:PGW:O04	7:B:409:PGW:H02	2.20	0.41
1:A:178:PRO:HG3	7:B:409:PGW:H3	2.02	0.41
1:A:221:ILE:HD13	1:B:221:ILE:CD1	2.50	0.41
1:B:100:LEU:HB3	1:B:101:PRO:HD3	2.02	0.41
1:B:332:ASP:O	1:B:336:VAL:HG23	2.20	0.41
1:A:100:LEU:HB3	1:A:101:PRO:HD3	2.02	0.40
1:A:37:ARG:HG3	5:A:408:SOG:O6	2.22	0.40
1:A:50:ILE:O	1:A:54:VAL:HG23	2.21	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	A	299/368 (81%)	295 (99%)	4 (1%)	0	100	100
1	B	301/368 (82%)	297 (99%)	4 (1%)	0	100	100
All	All	600/736 (82%)	592 (99%)	8 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	256/305 (84%)	255 (100%)	1 (0%)	91	95
1	B	252/305 (83%)	252 (100%)	0	100	100
All	All	508/610 (83%)	507 (100%)	1 (0%)	93	96

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	209	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	116	HIS
1	B	344	HIS

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 20 ligands modelled in this entry, 1 is monoatomic - leaving 19 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	$\#|Z| > 2$
5	SOG	B	406	-	20,20,20	0.25	0	24,25,25	0.25	0
4	SO4	B	403	-	4,4,4	0.14	0	6,6,6	0.12	0
7	PGW	B	409	-	50,50,50	0.28	0	53,56,56	0.44	0
3	PG4	A	402	-	12,12,12	0.16	0	11,11,11	0.11	0
5	SOG	A	406	-	20,20,20	0.29	0	24,25,25	0.27	0
5	SOG	A	408	-	20,20,20	0.24	0	24,25,25	0.40	0
2	NT5	A	401	-	34,34,34	1.10	2 (5%)	45,47,47	1.64	7 (15%)
5	SOG	B	405	-	20,20,20	0.29	0	24,25,25	0.24	0
3	PG4	B	402	-	8,8,12	0.21	0	7,7,11	0.18	0
5	SOG	A	409	-	3,3,20	0.22	0	2,2,25	0.47	0
7	PGW	A	411	-	50,50,50	0.28	0	53,56,56	0.36	0
4	SO4	A	404	-	4,4,4	0.14	0	6,6,6	0.08	0
5	SOG	A	405	-	20,20,20	0.36	0	24,25,25	0.62	0
5	SOG	B	408	-	5,5,20	0.27	0	4,4,25	0.19	0
5	SOG	B	407	-	20,20,20	0.29	0	24,25,25	0.36	0
4	SO4	A	403	-	4,4,4	0.14	0	6,6,6	0.09	0
2	NT5	B	401	-	34,34,34	1.11	2 (5%)	45,47,47	1.56	7 (15%)
4	SO4	B	404	-	4,4,4	0.14	0	6,6,6	0.09	0
5	SOG	A	407	-	12,13,20	0.45	0	16,18,25	0.35	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
5	SOG	B	407	-	-	0/11/31/31	0/1/1/1
5	SOG	B	405	-	-	5/11/31/31	0/1/1/1
2	NT5	B	401	-	-	0/17/30/30	0/4/4/4
3	PG4	B	402	-	-	2/6/6/10	-
5	SOG	A	409	-	-	0/1/1/31	-
5	SOG	B	406	-	-	3/11/31/31	0/1/1/1
7	PGW	A	411	-	1/1/5/7	14/55/55/55	-
5	SOG	A	406	-	-	3/11/31/31	0/1/1/1
5	SOG	A	408	-	-	3/11/31/31	0/1/1/1
5	SOG	A	405	-	-	3/11/31/31	0/1/1/1
7	PGW	B	409	-	-	15/55/55/55	-
2	NT5	A	401	-	-	0/17/30/30	0/4/4/4
5	SOG	A	407	-	-	2/4/24/31	0/1/1/1
3	PG4	A	402	-	-	5/10/10/10	-
5	SOG	B	408	-	-	0/3/3/31	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	401	NT5	C7-N1	3.01	1.37	1.32
2	A	401	NT5	C7-N1	3.01	1.37	1.32
2	B	401	NT5	C11-C10	2.12	1.40	1.37
2	A	401	NT5	C11-C10	2.06	1.40	1.37

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	NT5	C9-C10-C11	-5.77	118.52	121.54
2	B	401	NT5	C9-C10-C11	-5.36	118.73	121.54
2	A	401	NT5	C4-C5-C12	4.11	113.33	108.68
2	B	401	NT5	C4-C5-C12	3.55	112.69	108.68
2	A	401	NT5	C21-N3-C20	3.24	119.47	116.08
2	B	401	NT5	C21-N3-C20	2.94	119.15	116.08
2	A	401	NT5	C10-C11-N1	2.74	123.11	121.46
2	B	401	NT5	C10-C11-N1	2.37	122.89	121.46
2	B	401	NT5	C19-C20-N3	2.34	119.61	116.99
2	A	401	NT5	C14-C13-N2	2.32	121.13	117.92
2	A	401	NT5	O1-C6-C5	2.19	113.94	108.21

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	NT5	O1-C6-C5	2.17	113.87	108.21
2	B	401	NT5	C5-C12-N2	-2.14	105.88	110.94
2	A	401	NT5	C8-C7-N1	-2.03	121.84	124.87

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	A	411	PGW	C05

All (55) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	408	SOG	O5-C1-S1-C1'
7	A	411	PGW	C03-O11-P-O14
5	A	407	SOG	O5-C5-C6-O6
5	A	405	SOG	O5-C5-C6-O6
5	A	406	SOG	S1-C1'-C2'-C3'
3	A	402	PG4	O2-C3-C4-O3
5	A	407	SOG	C4-C5-C6-O6
7	A	411	PGW	O12-C04-C05-CAD
3	B	402	PG4	O3-C5-C6-O4
5	B	406	SOG	S1-C1'-C2'-C3'
5	A	405	SOG	C4-C5-C6-O6
5	B	406	SOG	C1'-C2'-C3'-C4'
5	B	405	SOG	S1-C1'-C2'-C3'
5	B	405	SOG	C3'-C4'-C5'-C6'
7	A	411	PGW	O12-C04-C05-OAF
7	A	411	PGW	C08-C09-C11-C12
7	A	411	PGW	C07-C08-C09-C11
7	B	409	PGW	C20-C19-O03-C01
7	B	409	PGW	C25-C26-C27-C15
5	B	406	SOG	O5-C5-C6-O6
5	B	405	SOG	O5-C5-C6-O6
7	A	411	PGW	C25-C26-C27-C15
7	A	411	PGW	C05-C04-O12-P
7	B	409	PGW	C1-C2-C3-C4
7	B	409	PGW	C19-C20-C21-C22
5	A	406	SOG	C2'-C3'-C4'-C5'
5	A	408	SOG	O5-C5-C6-O6
5	A	406	SOG	O5-C5-C6-O6
5	B	405	SOG	C2'-C3'-C4'-C5'
3	A	402	PG4	O4-C7-C8-O5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	409	PGW	O12-C04-C05-CAD
7	A	411	PGW	C2-C1-O01-C02
7	A	411	PGW	C20-C19-O03-C01
7	B	409	PGW	C01-C02-C03-O11
7	A	411	PGW	O02-C1-O01-C02
7	B	409	PGW	O04-C19-O03-C01
5	B	405	SOG	C5'-C6'-C7'-C8'
7	B	409	PGW	C3-C4-C5-C6
5	A	405	SOG	C1'-C2'-C3'-C4'
7	A	411	PGW	C17-C18-C28-C30
7	B	409	PGW	C2-C3-C4-C5
3	A	402	PG4	C8-C7-O4-C6
7	A	411	PGW	O04-C19-O03-C01
7	B	409	PGW	C04-O12-P-O11
5	A	408	SOG	C2-C1-S1-C1'
7	B	409	PGW	O01-C02-C03-O11
7	B	409	PGW	C06-C10-C9-C8
3	B	402	PG4	C5-C6-O4-C7
7	A	411	PGW	C04-C05-CAD-OAE
7	A	411	PGW	C07-C06-C10-C9
7	B	409	PGW	O12-C04-C05-OAF
3	A	402	PG4	C5-C6-O4-C7
7	B	409	PGW	C02-C01-O03-C19
3	A	402	PG4	C3-C4-O3-C5
7	B	409	PGW	O01-C1-C2-C3

There are no ring outliers.

8 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	409	PGW	8	0
3	A	402	PG4	2	0
5	A	408	SOG	1	0
3	B	402	PG4	1	0
7	A	411	PGW	5	0
5	B	407	SOG	1	0
2	B	401	NT5	1	0
5	A	407	SOG	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 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 SOG B 406

Bond lengths

Bond angles

Torsions

Rings

Ligand PGW B 409

Bond lengths

Bond angles

Torsions

Rings

Ligand SOG A 406

Bond lengths

Bond angles

Torsions

Rings

Ligand SOG A 408

Bond lengths

Bond angles

Torsions

Rings

Ligand NT5 A 401

Bond lengths

Bond angles

Torsions

Rings

Ligand SOG B 405

Bond lengths

Bond angles

Torsions

Rings

Ligand PGW A 411

Bond lengths

Bond angles

Torsions

Rings

Ligand SOG A 405	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand NT5 B 401	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand SOG A 407

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 ⓘ

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(Å²)	Q<0.9
1	A	305/368 (82%)	0.76	45 (14%)	2	3	38, 65, 135, 179	0
1	B	303/368 (82%)	0.94	53 (17%)	1	2	41, 69, 128, 164	0
All	All	608/736 (82%)	0.85	98 (16%)	1	3	38, 67, 133, 179	0

All (98) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	245	ARG	11.2
1	A	189	LEU	9.0
1	B	45	TYR	8.9
1	A	69	LEU	8.8
1	A	330	ALA	7.7
1	B	47	TRP	7.1
1	B	287	VAL	6.5
1	A	252	SER	6.4
1	A	188	VAL	6.0
1	B	288	LYS	5.9
1	B	247	ILE	5.8
1	B	72	TRP	5.8
1	A	332	ASP	5.5
1	B	242	LEU	5.4
1	A	26	ALA	5.3
1	A	286	GLU	5.3
1	B	377	LEU	5.2
1	B	246	GLN	4.9
1	A	285	ALA	4.8
1	A	72	TRP	4.7
1	B	129	SER	4.7
1	B	152	LEU	4.7
1	B	375	TRP	4.7
1	A	243	TRP	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	375	TRP	4.6
1	B	292	ALA	4.6
1	A	47	TRP	4.5
1	B	155	LYS	4.4
1	B	290	MET	4.4
1	B	330	ALA	4.2
1	B	44	GLN	4.2
1	A	251	THR	4.2
1	B	373	PHE	4.2
1	A	152	LEU	4.0
1	A	76	HIS	4.0
1	A	373	PHE	3.9
1	B	337	TYR	3.9
1	A	361	LEU	3.9
1	A	329	GLN	3.9
1	A	376	TRP	3.9
1	A	242	LEU	3.9
1	B	157	THR	3.9
1	B	132	VAL	3.9
1	B	374	SER	3.9
1	A	298	LYS	3.5
1	A	239	PHE	3.5
1	A	199	PHE	3.4
1	A	291	ARG	3.4
1	B	48	VAL	3.3
1	A	374	SER	3.3
1	A	83	TYR	3.3
1	B	88	LEU	3.2
1	B	130	VAL	3.2
1	A	333	ARG	3.2
1	A	71	VAL	3.2
1	A	75	HIS	3.1
1	B	162	LEU	3.1
1	A	73	ARG	3.1
1	B	371	ALA	3.1
1	B	370	LYS	3.1
1	B	83	TYR	3.1
1	A	198	ALA	3.0
1	B	133	ALA	3.0
1	B	46	ALA	2.9
1	B	376	TRP	2.9
1	B	291	ARG	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	329	GLN	2.7
1	A	68	CYS	2.7
1	A	294	ARG	2.7
1	A	295	LYS	2.7
1	B	140	ILE	2.7
1	B	134	VAL	2.6
1	B	243	TRP	2.6
1	A	360	PHE	2.6
1	B	137	LEU	2.6
1	A	25	ALA	2.5
1	B	372	ALA	2.5
1	B	172	SER	2.5
1	B	131	SER	2.5
1	B	136	THR	2.5
1	A	113	LEU	2.5
1	B	75	HIS	2.5
1	B	108	ILE	2.4
1	B	153	LEU	2.4
1	A	77	MET	2.4
1	B	139	PHE	2.4
1	B	78	ARG	2.3
1	A	153	LEU	2.3
1	A	175	ILE	2.3
1	B	328	ARG	2.2
1	B	333	ARG	2.2
1	A	331	SER	2.2
1	B	347	VAL	2.2
1	B	135	LEU	2.1
1	A	130	VAL	2.1
1	A	127	ALA	2.1
1	B	85	LEU	2.1
1	A	155	LYS	2.1

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
5	SOG	B	407	20/20	0.67	0.34	123,129,140,145	0
5	SOG	A	408	20/20	0.70	0.21	124,138,144,148	0
3	PG4	A	402	13/13	0.72	0.35	62,97,110,110	0
5	SOG	A	407	13/20	0.80	0.19	134,142,147,151	0
7	PGW	B	409	51/51	0.80	0.26	40,72,132,137	0
5	SOG	B	408	6/20	0.83	0.52	86,89,95,95	0
4	SO4	A	404	5/5	0.83	0.20	122,129,133,133	0
5	SOG	A	406	20/20	0.85	0.27	49,108,119,126	0
3	PG4	B	402	9/13	0.86	0.29	74,81,83,83	0
7	PGW	A	411	51/51	0.87	0.18	47,80,127,131	0
5	SOG	B	405	20/20	0.87	0.31	80,98,111,112	0
4	SO4	B	404	5/5	0.89	0.10	119,122,127,131	0
5	SOG	B	406	20/20	0.89	0.29	41,98,110,111	0
6	CL	A	410	1/1	0.93	0.17	68,68,68,68	0
5	SOG	A	405	20/20	0.93	0.14	42,64,76,88	0
4	SO4	A	403	5/5	0.93	0.13	103,106,112,114	0
2	NT5	B	401	31/31	0.94	0.17	50,66,83,89	0
2	NT5	A	401	31/31	0.95	0.17	40,53,65,72	0
5	SOG	A	409	4/20	0.96	0.23	52,62,66,67	0
4	SO4	B	403	5/5	0.97	0.09	76,81,91,93	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 SOG A 408:

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 SOG A 407:

$2mF_o-DF_c$ (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
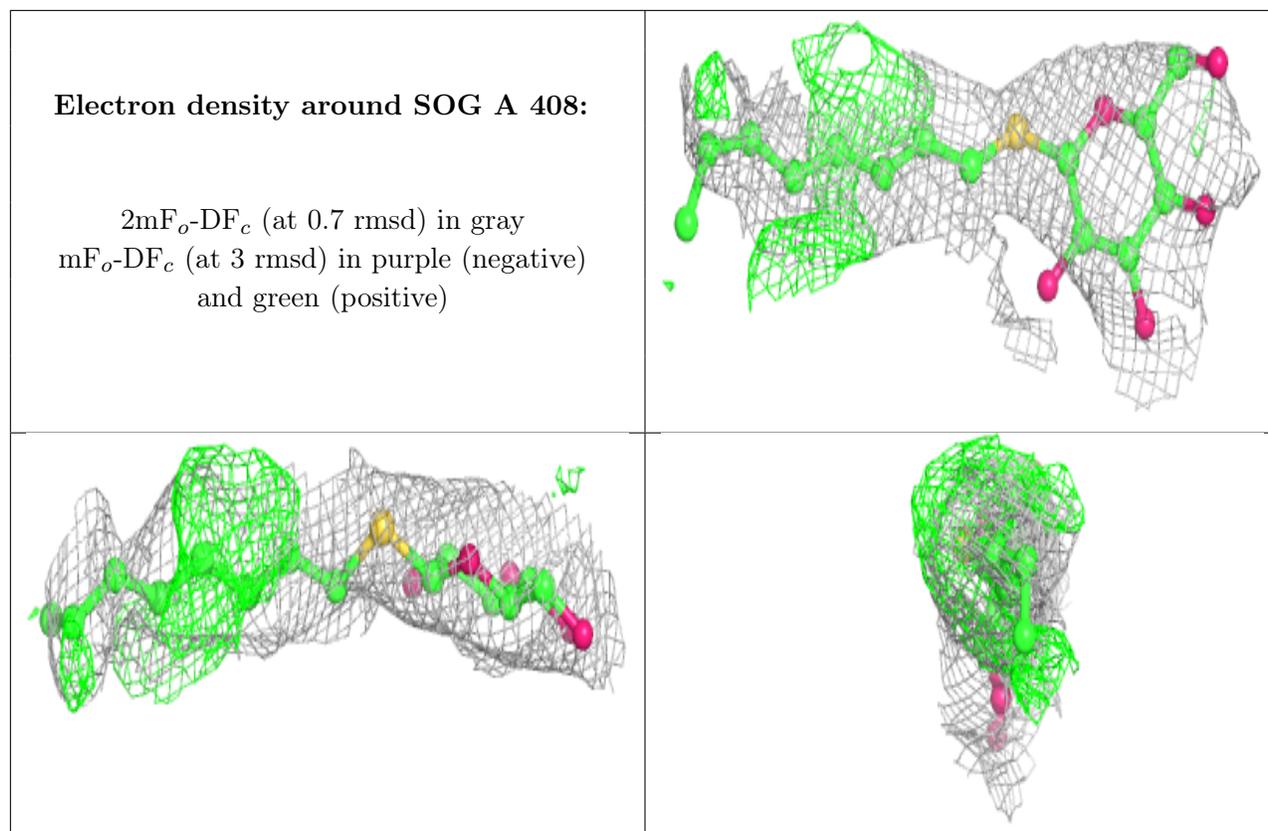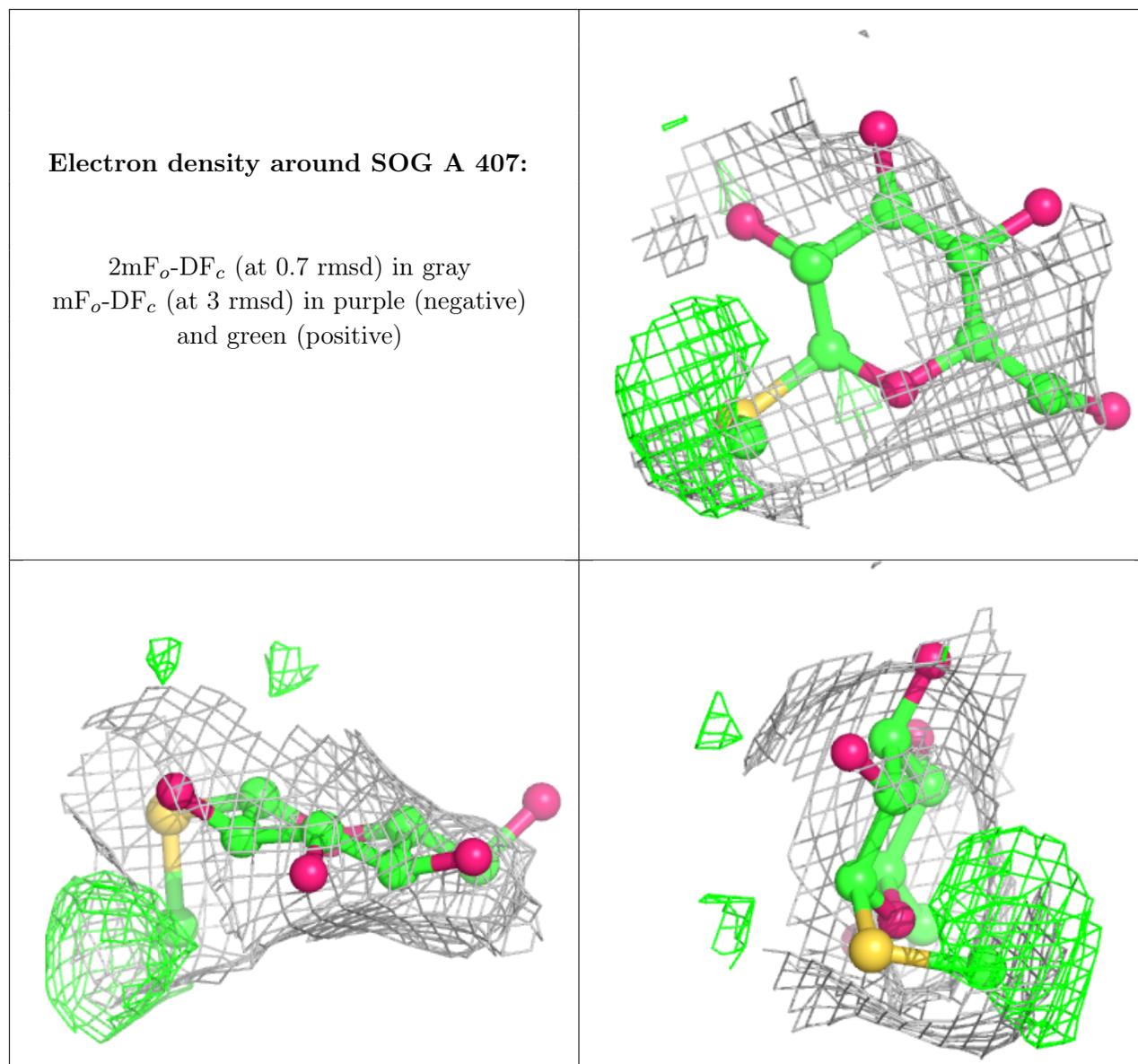
**Electron density around PGW B 409:**

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
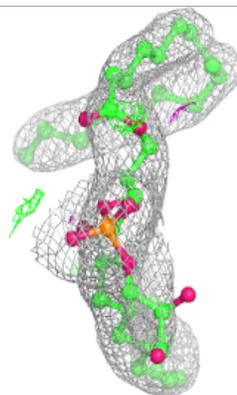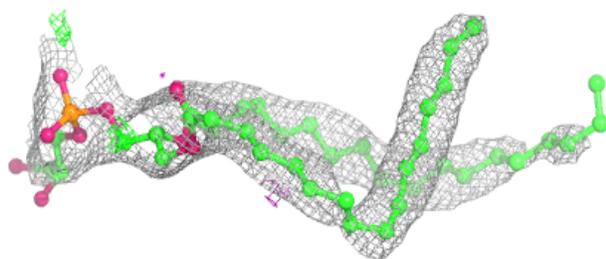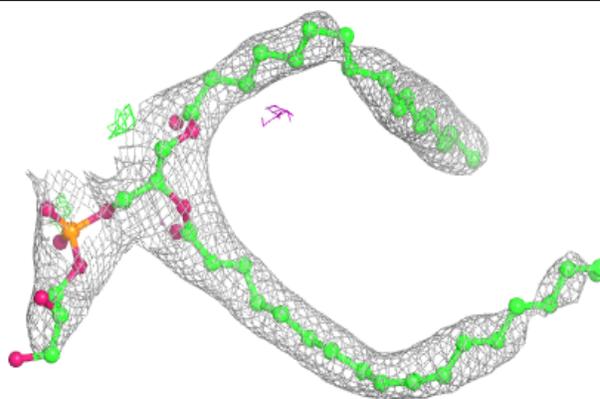
Electron density around SOG A 406:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
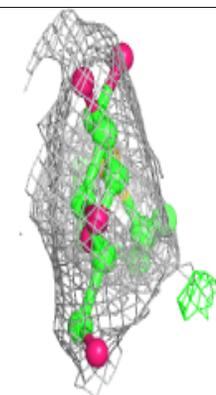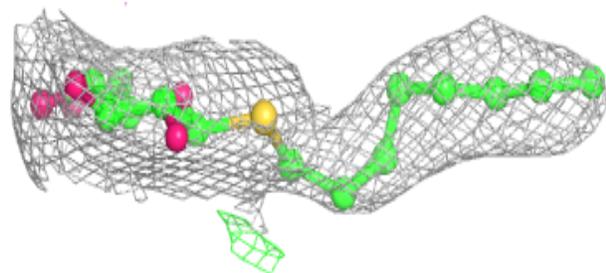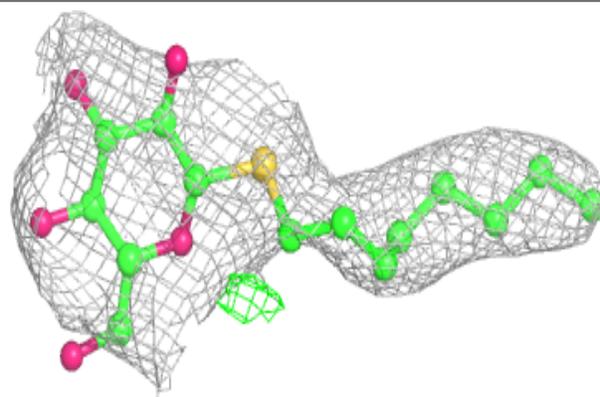
Electron density around PGW A 411:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
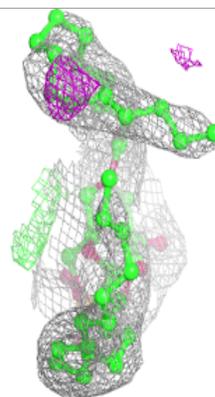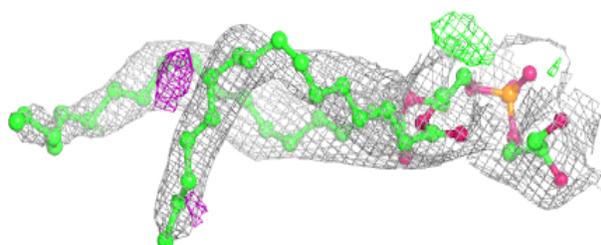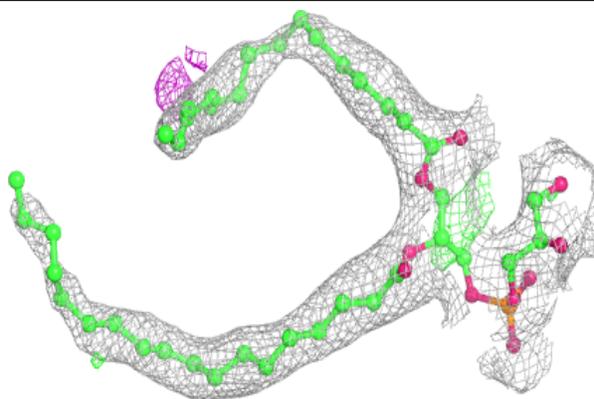
Electron density around SOG B 405:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
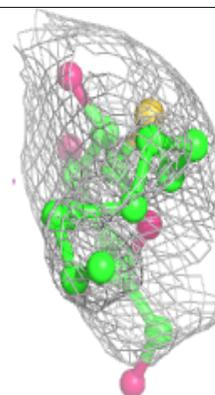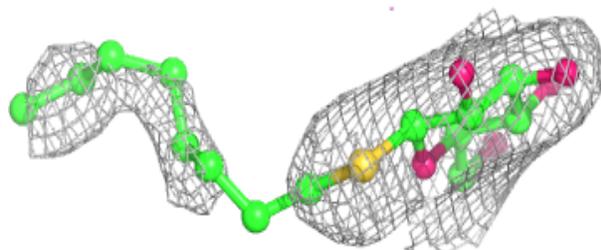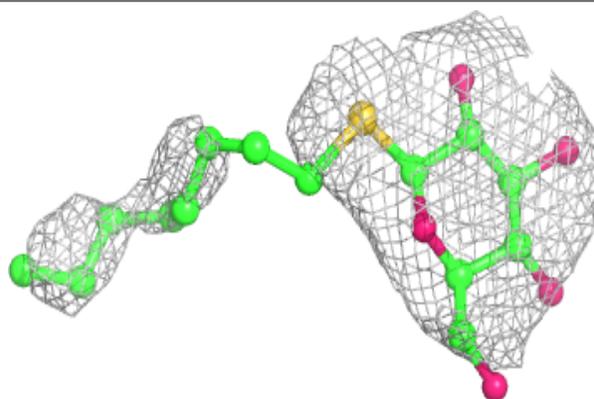
Electron density around SOG B 406:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
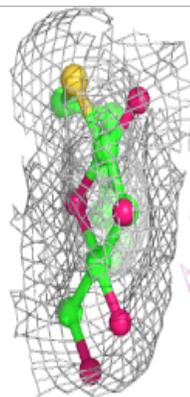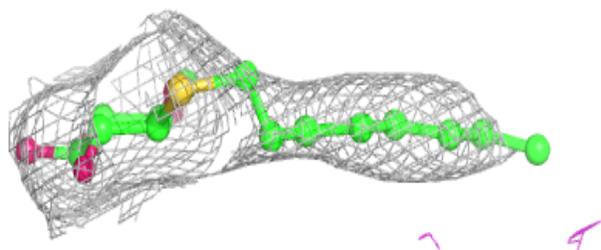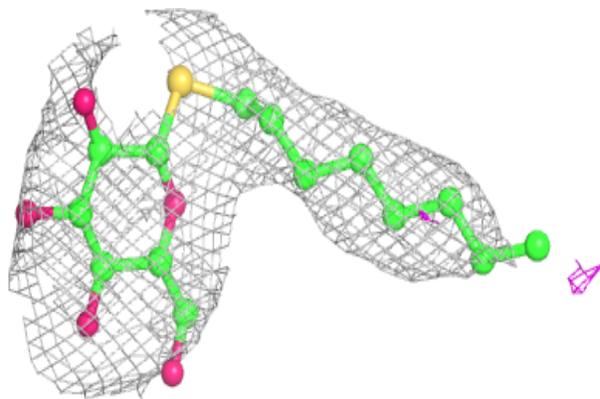
Electron density around SOG A 405:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
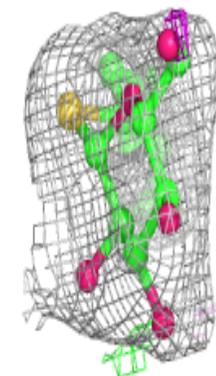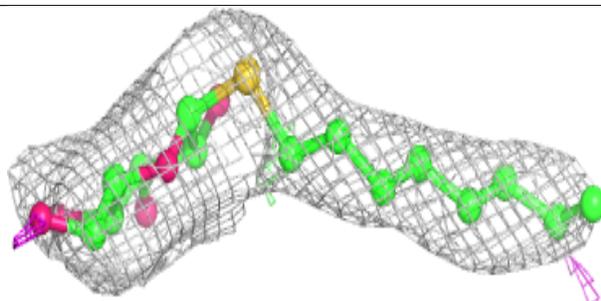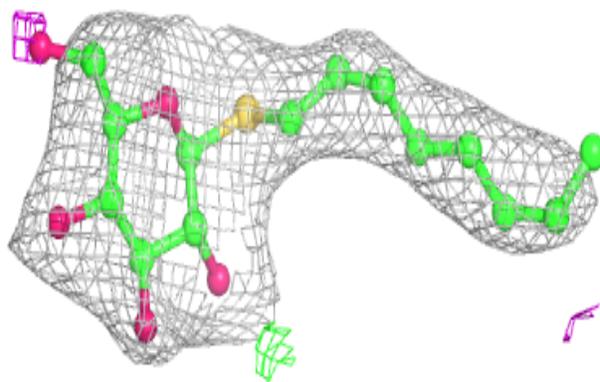
Electron density around NT5 B 401:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)
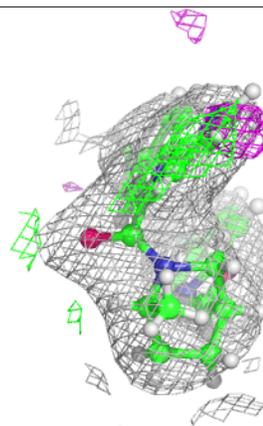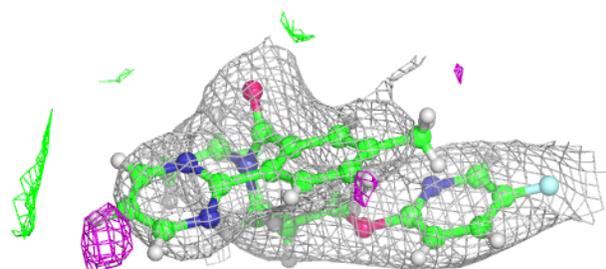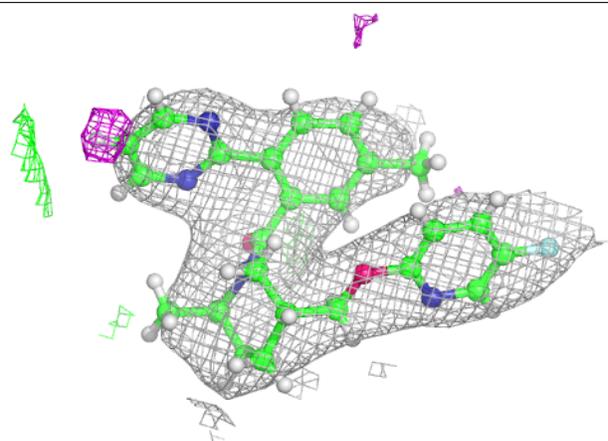
Electron density around NT5 A 401:

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 ⓘ

There are no such residues in this entry.