W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Aug 14, 2021 – 04:06 pm BST

PDB ID : 6ZIS
Title : Crystal structure of a CGRP receptor ectodomain heterodimer with bound high affinity inhibitor
Authors : Southall, S.M.
Deposited on : 2020-06-26
Resolution : 1.73 Å(reported)

1 Overall quality at a glance (i)

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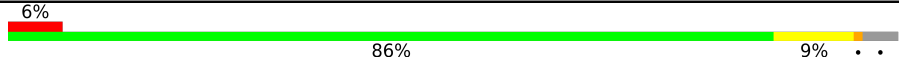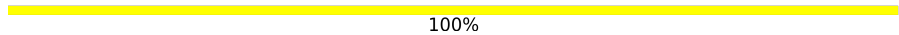*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.73 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.
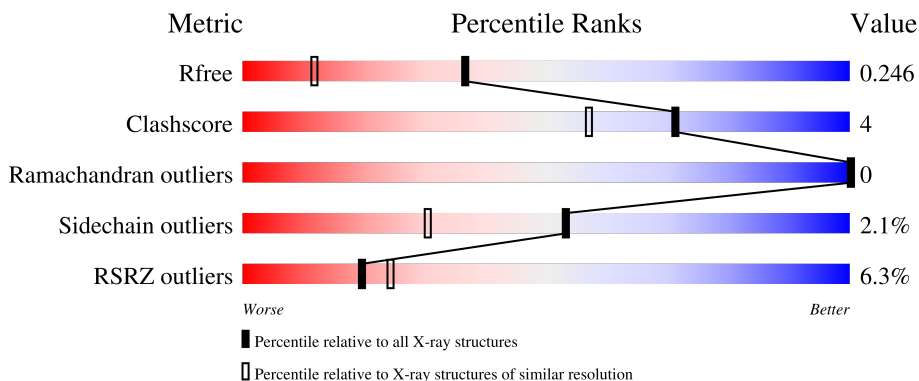
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3764 (1.76-1.72)
Clashscore	141614	3923 (1.76-1.72)
Ramachandran outliers	138981	3878 (1.76-1.72)
Sidechain outliers	138945	3878 (1.76-1.72)
RSRZ outliers	127900	3705 (1.76-1.72)

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	594	6% ... 86% ... 9% ..
2	B	2	100%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4911 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 Maltose/maltodextrin-binding periplasmic protein,Receptor activity-modifying protein 1,Calcitonin gene-related peptide type 1 receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	569	4513	2890	744	856	23	0	4	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	ALA	-	expression tag	UNP P0AEX9
A	1	SER	-	expression tag	UNP P0AEX9
A	369	ASN	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P0AEX9
A	371	ALA	-	linker	UNP P0AEX9
A	372	ALA	-	linker	UNP P0AEX9
A	1022	GLU	-	linker	UNP P0AEX9
A	1023	PHE	-	linker	UNP P0AEX9
A	2020	GLY	-	linker	UNP O60894
A	2021	SER	-	linker	UNP O60894
A	2022	ALA	-	linker	UNP O60894
A	2023	GLY	-	linker	UNP O60894
A	2024	SER	-	linker	UNP O60894
A	2025	ALA	-	linker	UNP O60894
A	2026	GLY	-	linker	UNP O60894
A	2027	SER	-	linker	UNP O60894
A	2028	ALA	-	linker	UNP O60894
A	2066	GLN	ASN	conflict	UNP Q16602
A	2118	GLN	ASN	conflict	UNP Q16602
A	2123	ASP	ASN	conflict	UNP Q16602
A	2145	HIS	-	expression tag	UNP Q16602
A	2146	HIS	-	expression tag	UNP Q16602
A	2147	HIS	-	expression tag	UNP Q16602
A	2148	HIS	-	expression tag	UNP Q16602
A	2149	HIS	-	expression tag	UNP Q16602
A	2150	HIS	-	expression tag	UNP Q16602

- Molecule 2 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
2	B	2	Total 23	C 12	O 11	0	0	0

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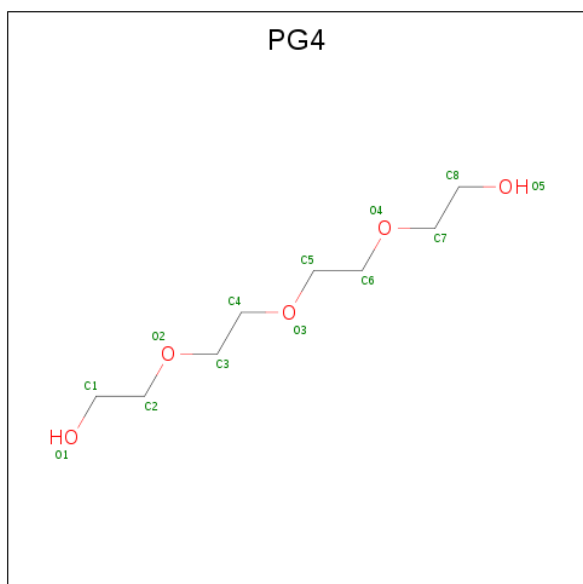
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 13	C 8	O 5	0	0
3	A	1	Total 13	C 8	O 5	0	0
3	A	1	Total 13	C 8	O 5	0	0

- Molecule 4 is N-{(1S)-5-amino-1-[(4-pyridin-4-ylpiperazin-1-yl)carbonyl]pentyl}-3,5-dibrom o-Nalpha-{[4-(2-oxo-1,4-dihydroquinazolin-3 (2H)-yl)piperidin-1-yl]carbonyl}-D-tyrosinam ide (three-letter code: 3N6) (formula: $C_{38}H_{47}Br_2N_9O_5$) (labeled as "Ligand of Interest" by depositor).
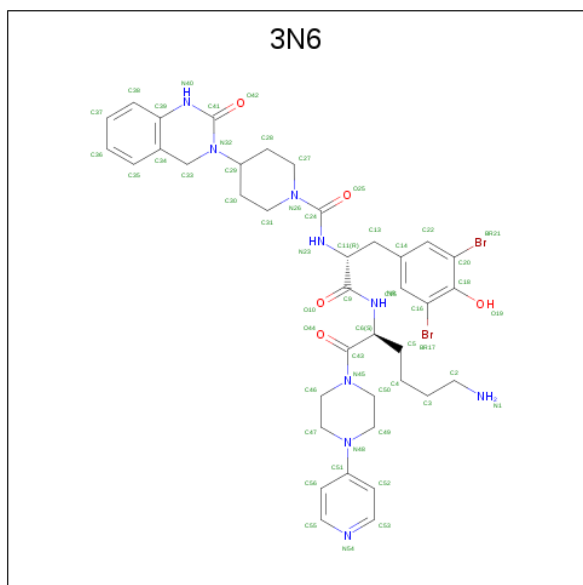
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	Br	C	N	O		
4	A	1	54	2	38	9	5	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	282	282	282	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Maltose/maltodextrin-binding periplasmic protein,Receptor activity-modifying protein 1,Calcitonin gene-related peptide type 1 receptor

Chain A:
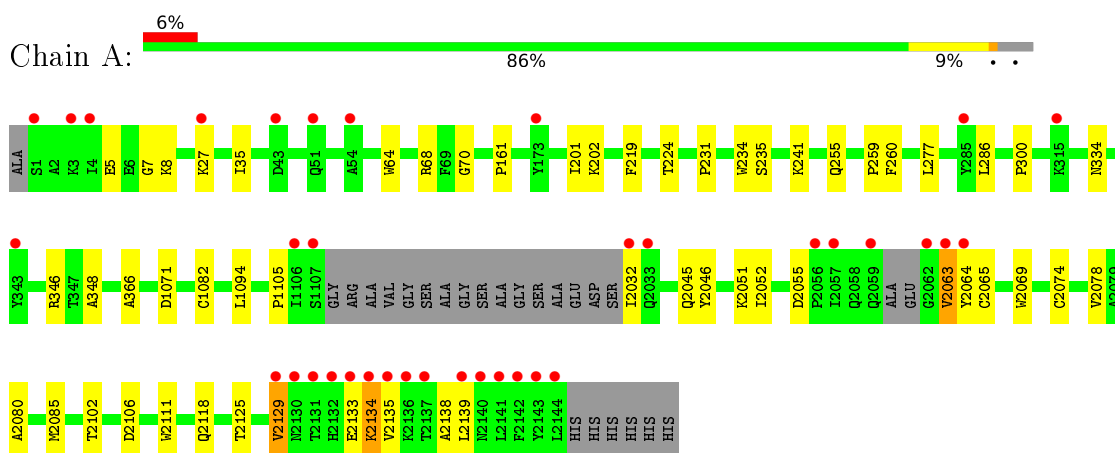
• Molecule 2: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain B:

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.14Å 76.98Å 98.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.62 – 1.73 60.60 – 1.73	Depositor EDS
% Data completeness (in resolution range)	99.9 (57.62-1.73) 99.9 (60.60-1.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 1.73Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.199 , 0.244 0.199 , 0.246	Depositor DCC
R_{free} test set	2721 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	27.0	Xtriage
Anisotropy	0.505	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 44.4	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	0.96	EDS
Total number of atoms	4911	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: 3N6, GLC, PG4

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	0.43	2/4639 (0.0%)	0.55	0/6302

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1082	CYS	CB-SG	-5.51	1.72	1.81
1	A	2074	CYS	CB-SG	-5.03	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4513	0	4370	32	0
2	B	23	0	21	0	0
3	A	39	0	54	5	0
4	A	54	0	46	0	0
5	A	282	0	0	3	0
All	All	4911	0	4491	32	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 4.

The worst 5 of 32 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:2118:GLN:NE2	5:A:2301:HOH:O	2.04	0.63
1:A:1105:PRO:HD2	3:A:2203:PG4:H11	1.81	0.61
1:A:35:ILE:HG13	1:A:277:LEU:HD22	1.81	0.61
1:A:2106:ASP:HA	3:A:2202:PG4:H32	1.84	0.59
1:A:346:ARG:NH1	5:A:2307:HOH:O	2.35	0.58

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	567/594 (96%)	561 (99%)	6 (1%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	472/484 (98%)	462 (98%)	10 (2%)	53	30

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2133	GLU
1	A	2134	LYS
1	A	2135	VAL
1	A	1071	ASP
1	A	2032	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2107	GLN
1	A	2132	HIS
1	A	2140	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 ⓘ

2 monosaccharides 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	GLC	B	1	2	12,12,12	0.32	0	17,17,17	1.18	3 (17%)
2	GLC	B	2	2	11,11,12	0.56	0	15,15,17	1.27	1 (6%)

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
2	GLC	B	1	2	-	0/2/22/22	0/1/1/1
2	GLC	B	2	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2	GLC	C1-O5-C5	3.60	117.08	112.19
2	B	1	GLC	O5-C1-C2	2.38	114.54	110.28
2	B	1	GLC	O2-C2-C3	-2.19	105.29	110.35
2	B	1	GLC	O3-C3-C2	-2.04	105.64	110.35

There are no chirality outliers.

There are no torsion outliers.

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 oligosaccharide.

Oligosaccharide Chain B

Bond lengths

Bond angles

Torsions

Rings
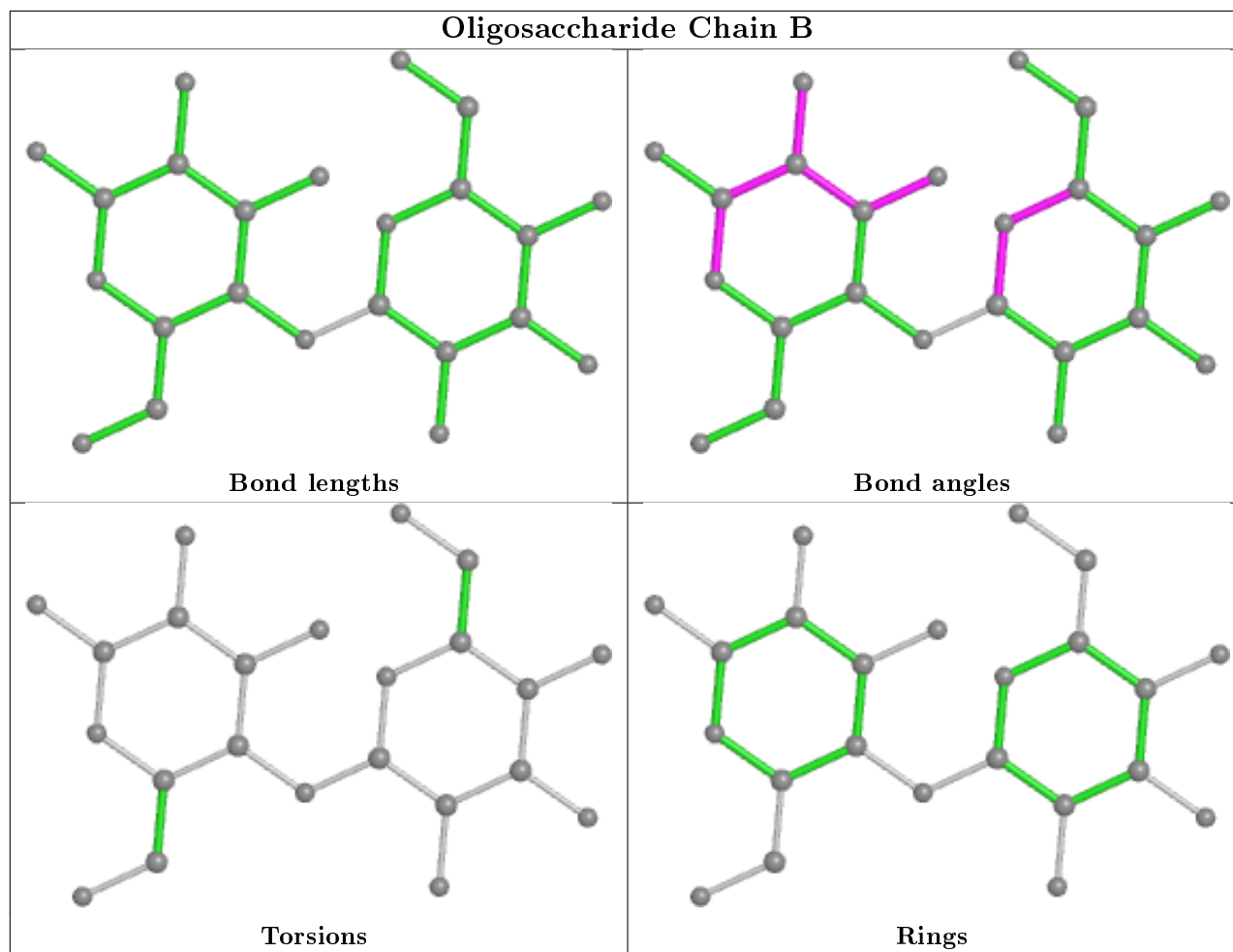
5.6 Ligand geometry (i)

4 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	#$|Z| > 2$
4	3N6	A	2204	-	59,59,59	0.39	0	78,82,82	0.71	3 (3%)
3	PG4	A	2203	-	12,12,12	0.56	0	11,11,11	0.19	0
3	PG4	A	2201	-	12,12,12	0.47	0	11,11,11	0.40	0
3	PG4	A	2202	-	12,12,12	0.50	0	11,11,11	0.28	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
4	3N6	A	2204	-	-	1/41/73/73	0/6/6/6
3	PG4	A	2203	-	-	3/10/10/10	-
3	PG4	A	2201	-	-	3/10/10/10	-
3	PG4	A	2202	-	-	3/10/10/10	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	2204	3N6	C47-N48-C51	2.31	124.33	118.09
4	A	2204	3N6	O42-C41-N32	-2.21	120.33	123.11
4	A	2204	3N6	C49-N48-C51	2.14	123.89	118.09

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	2202	PG4	O4-C7-C8-O5
3	A	2203	PG4	O1-C1-C2-O2
4	A	2204	3N6	C6-C43-N45-C50
3	A	2201	PG4	O3-C5-C6-O4
3	A	2202	PG4	C5-C6-O4-C7

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	2203	PG4	2	0
3	A	2201	PG4	2	0
3	A	2202	PG4	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 3N6 A 2204

Bond lengths — Bond angles — Torsions — Rings
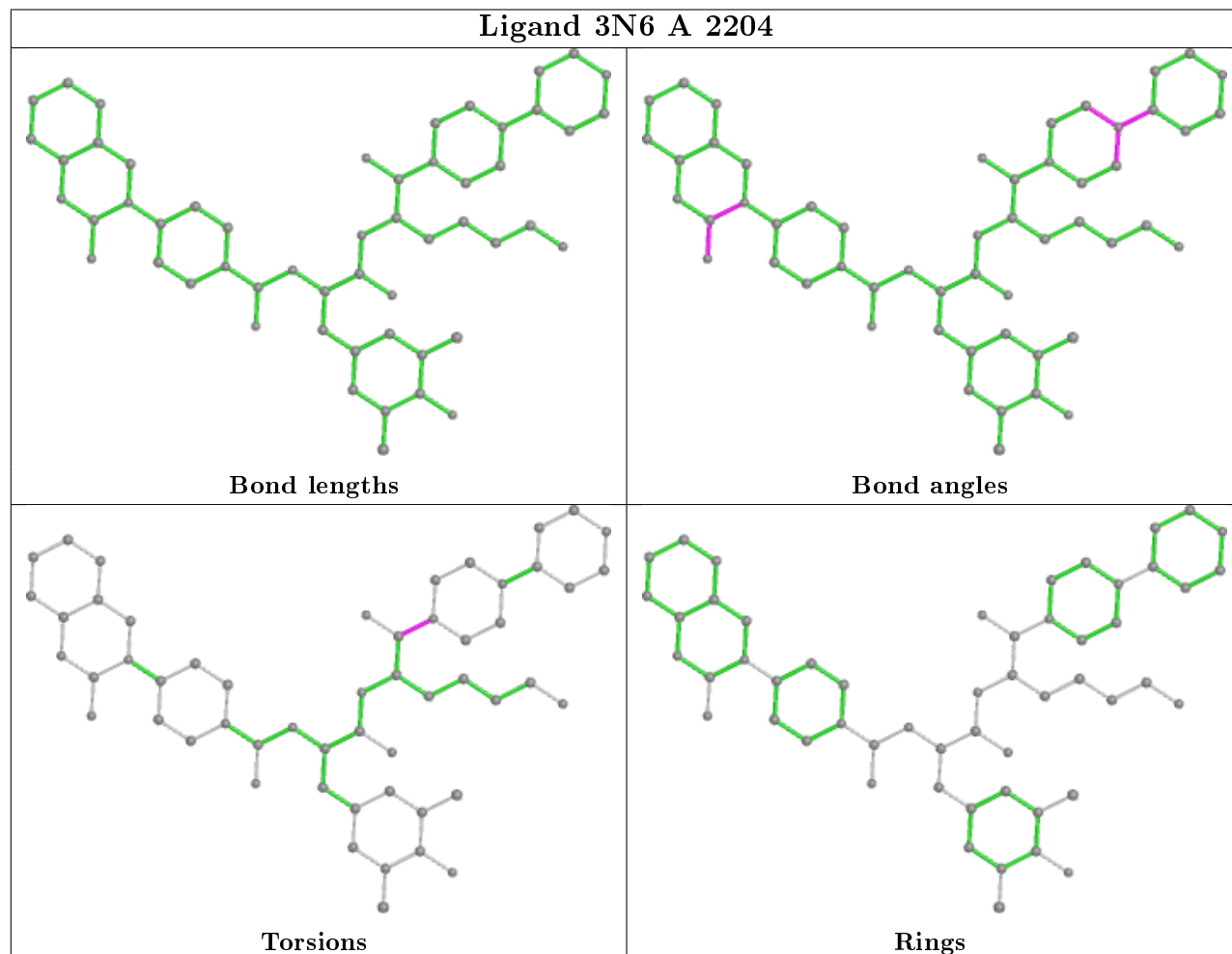
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	569/594 (95%)	0.37	36 (6%)	20	24	20, 31, 60, 111	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2137	THR	8.0
1	A	2141	LEU	5.2
1	A	2135	VAL	4.6
1	A	2139	LEU	4.2
1	A	1107	SER	4.2

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)

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	GLC	B	1	12/12	0.96	0.10	20,24,33,35	0
2	GLC	B	2	11/12	0.96	0.09	20,23,25,26	0

6.4 Ligands (i)

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

median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	PG4	A	2202	13/13	0.68	0.24	64,66,75,76	0
3	PG4	A	2203	13/13	0.68	0.20	42,51,77,81	0
3	PG4	A	2201	13/13	0.74	0.26	56,69,75,81	0
4	3N6	A	2204	54/54	0.97	0.10	21,29,42,45	2

6.5 Other polymers (i)

There are no such residues in this entry.