WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 02:59 AM EDT

PDB ID	:	6AWQ
Title	:	Anomalous chloride signal reveals the position of sertraline complexed with the serotonin transporter at the central site
Authors	:	Coleman, J.A.; Gouaux, E.
Deposited on	:	2017-09-06
Resolution	:	4.05 Å(reported)

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

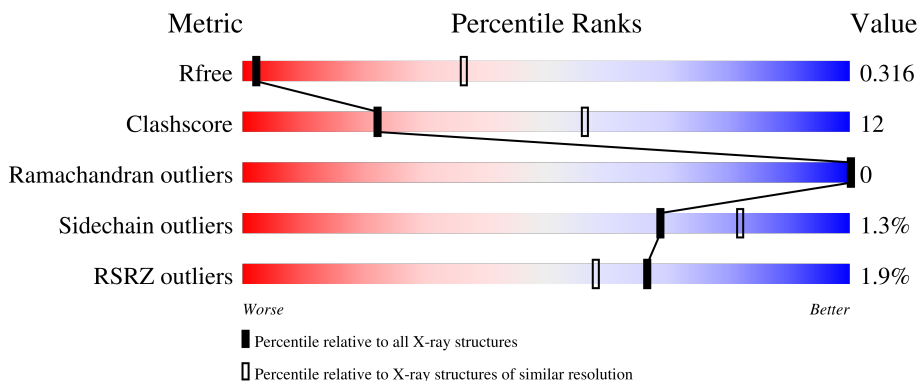|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 4.05 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		0.316
Clashscore		12
Ramachandran outliers		0
Sidechain outliers		1.3%
RSRZ outliers		1.9%

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	1105 (4.38-3.70)
Clashscore	141614	1005 (4.36-3.72)
Ramachandran outliers	138981	1125 (4.38-3.70)
Sidechain outliers	138945	1115 (4.38-3.70)
RSRZ outliers	127900	1003 (4.40-3.68)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	549	69% 29% ••
2	B	229	6% 79% 17% 5%
3	C	214	2% 75% 25%

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
4	SRE	A	701	-	-	X	-
5	NAG	A	704	-	-	-	X

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 7677 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 Sodium-dependent serotonin transporter.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	544	Total 4293	C 2870	N 663	O 736	S 24	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	74	GLY	-	expression tag	UNP P31645
A	75	SER	-	expression tag	UNP P31645
A	110	ALA	TYR	engineered mutation	UNP P31645
A	291	ALA	ILE	engineered mutation	UNP P31645
A	439	SER	THR	engineered mutation	UNP P31645
A	554	ALA	CYS	engineered mutation	UNP P31645
A	580	ALA	CYS	engineered mutation	UNP P31645
A	619	LEU	-	expression tag	UNP P31645
A	620	VAL	-	expression tag	UNP P31645
A	621	PRO	-	expression tag	UNP P31645
A	622	ARG	-	expression tag	UNP P31645

- Molecule 2 is a protein called 8B6 antibody FAB heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	218	Total 1643	C 1038	N 266	O 331	S 8	0	0	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	238	ARG	-	expression tag	UNP A0A0F7R1P3
B	239	GLY	-	expression tag	UNP A0A0F7R1P3
B	240	SER	-	expression tag	UNP A0A0F7R1P3
B	241	HIS	-	expression tag	UNP A0A0F7R1P3
B	242	HIS	-	expression tag	UNP A0A0F7R1P3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	243	HIS	-	expression tag	UNP A0A0F7R1P3
B	244	HIS	-	expression tag	UNP A0A0F7R1P3
B	245	HIS	-	expression tag	UNP A0A0F7R1P3
B	246	HIS	-	expression tag	UNP A0A0F7R1P3
B	247	HIS	-	expression tag	UNP A0A0F7R1P3
B	248	HIS	-	expression tag	UNP A0A0F7R1P3

- Molecule 3 is a protein called 8B6 antibody FAB light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	214	Total 1662	C 1037	N 280	O 337	S 8	0	0	0

- Molecule 4 is (1S,4S)-4-(3,4-dichlorophenyl)-N-methyl-1,2,3,4-tetrahydronaphthalen-1-amine (three-letter code: SRE) (formula: $C_{17}H_{17}Cl_2N$).
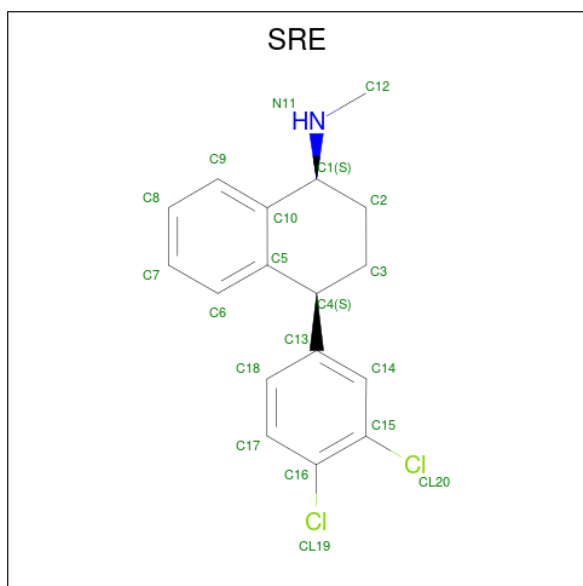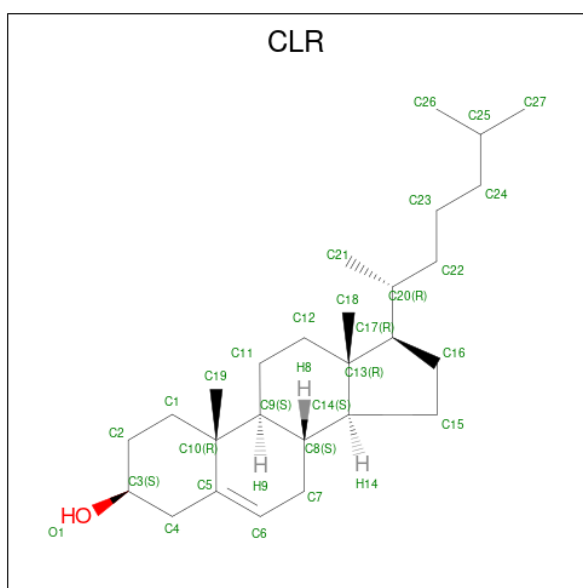

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 20	C 17	Cl 2	N 1	0	0

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).

NAG

O6
OH

O1
HO
O5
O

C6

C1(R)   C5(R)

C2(R)   C4(S)
C3(R)

N2
HN

OH
O4

C7

OH
O3

C8
O
O7

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is CHOLESTEROL (three-letter code: CLR) (formula: $C_{27}H_{46}O$).

CLR

C26   C25   C27

C23   C24

C21   C22

C18   C20(R)

C12   C17(R)

C11   H8   C13(R)   C16

C19   C9(S)   C14(S)

C1   C10(R)   C8(S)   C15

C2   C5   H9   C7   H14

C3(S)

HO
O1   C4   C6

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			28	27	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

• Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	2	Total 2	Na 2	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.
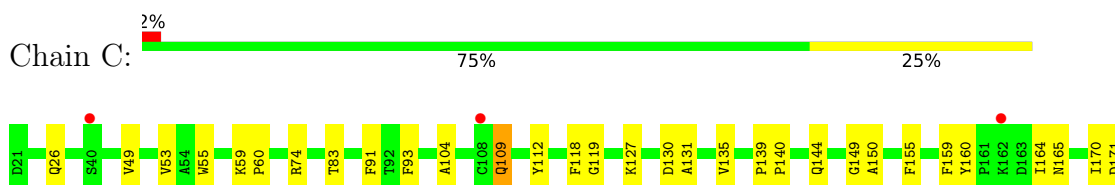
- Molecule 1: Sodium-dependent serotonin transporter
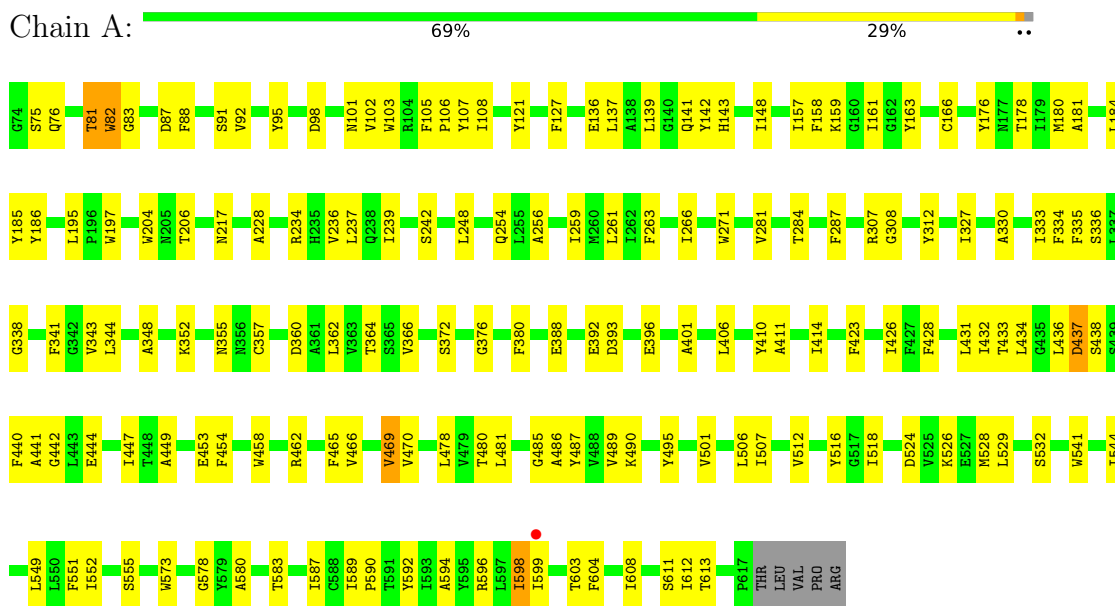
- Molecule 2: 8B6 antibody FAB heavy chain
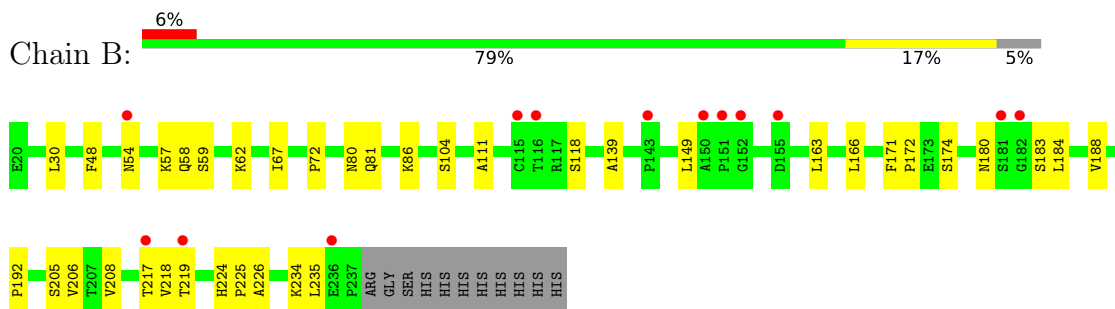
- Molecule 3: 8B6 antibody FAB light chain

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	130.55Å 165.90Å 142.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 4.05 19.99 – 4.05	Depositor EDS
% Data completeness (in resolution range)	98.9 (19.99-4.05) 98.7 (19.99-4.05)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 4.07Å)	Xtriage
Refinement program	PHENIX (dev_2597)	Depositor
R, R_{free}	0.269 , 0.316 0.269 , 0.316	Depositor DCC
R_{free} test set	634 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	173.2	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.20 , 73.5	EDS
L-test for twinning[2]	$< |L| > = 0.29, < L^2 > = 0.13$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7677	wwPDB-VP
Average B, all atoms (Å^2)	209.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.51% 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: NAG, NA, CLR, CL, SRE

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	0.35	1/4430 (0.0%)	0.52	2/6052 (0.0%)
2	B	0.25	0/1688	0.46	0/2309
3	C	0.25	0/1700	0.44	0/2307
All	All	0.31	1/7818 (0.0%)	0.49	2/10668 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	87	ASP	C-N	10.39	1.57	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	81	THR	O-C-N	-11.22	104.74	122.70
1	A	81	THR	CA-C-N	6.66	131.85	117.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	81	THR	Mainchain

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	4293	0	4257	122	0
2	B	1643	0	1589	21	0
3	C	1662	0	1585	33	0
4	A	20	0	17	7	0
5	A	28	0	26	0	0
6	A	28	0	46	5	0
7	A	1	0	0	0	0
8	A	2	0	0	0	0
All	All	7677	0	7520	176	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 12.

The worst 5 of 176 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:573:TRP:CE3	6:A:703:CLR:H71	1.85	1.11
1:A:573:TRP:CZ3	6:A:703:CLR:H71	2.09	0.87
3:C:211:SER:HA	3:C:229:PHE:O	1.77	0.85
1:A:335:PHE:HD1	4:A:701:SRE:H8	1.50	0.77
1:A:335:PHE:CD1	4:A:701:SRE:H8	2.23	0.74

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	542/549 (99%)	520 (96%)	22 (4%)	0	100	100
2	B	216/229 (94%)	204 (94%)	12 (6%)	0	100	100
3	C	212/214 (99%)	199 (94%)	13 (6%)	0	100	100
All	All	970/992 (98%)	923 (95%)	47 (5%)	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	450/462 (97%)	440 (98%)	10 (2%)	52	71
2	B	190/201 (94%)	190 (100%)	0	100	100
3	C	189/190 (100%)	188 (100%)	1 (0%)	88	93
All	All	829/853 (97%)	818 (99%)	11 (1%)	69	82

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	598	ILE
1	A	603	THR
3	C	109	GLN
1	A	604	PHE
1	A	437	ASP

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	111	GLN
1	A	141	GLN
3	C	26	GLN

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 7 ligands modelled in this entry, 3 are monoatomic - leaving 4 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	NAG	A	702	1	14,14,15	0.24	0	17,19,21	0.53	0
5	NAG	A	704	1	14,14,15	1.99	2 (14%)	17,19,21	2.00	2 (11%)
4	SRE	A	701	-	22,22,22	2.37	5 (22%)	31,31,31	1.83	4 (12%)
6	CLR	A	703	-	31,31,31	0.83	0	48,48,48	1.31	5 (10%)

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	NAG	A	702	1	-	2/6/23/26	0/1/1/1
5	NAG	A	704	1	-	3/6/23/26	0/1/1/1
4	SRE	A	701	-	-	0/5/19/19	0/3/3/3
6	CLR	A	703	-	-	6/10/68/68	0/4/4/4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	701	SRE	C10-C1	8.11	1.59	1.52
5	A	704	NAG	O5-C1	6.84	1.54	1.43
4	A	701	SRE	C5-C4	5.89	1.61	1.51
5	A	704	NAG	C1-C2	2.72	1.56	1.52
4	A	701	SRE	C3-C4	-2.53	1.50	1.54

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	704	NAG	C1-O5-C5	7.19	121.94	112.19
4	A	701	SRE	C2-C3-C4	5.74	115.56	110.04
4	A	701	SRE	C2-C1-C10	5.58	118.85	111.34
4	A	701	SRE	C3-C4-C5	3.56	117.82	109.10
6	A	703	CLR	C8-C7-C6	-3.14	108.22	112.73

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	704	NAG	O5-C5-C6-O6
5	A	704	NAG	C4-C5-C6-O6
5	A	702	NAG	O5-C5-C6-O6
6	A	703	CLR	C17-C20-C22-C23
6	A	703	CLR	C21-C20-C22-C23

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	701	SRE	7	0
6	A	703	CLR	5	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 SRE A 701

Bond lengths   Bond angles

Torsions   Rings

Ligand CLR A 703

Bond lengths   Bond angles

Torsions   Rings
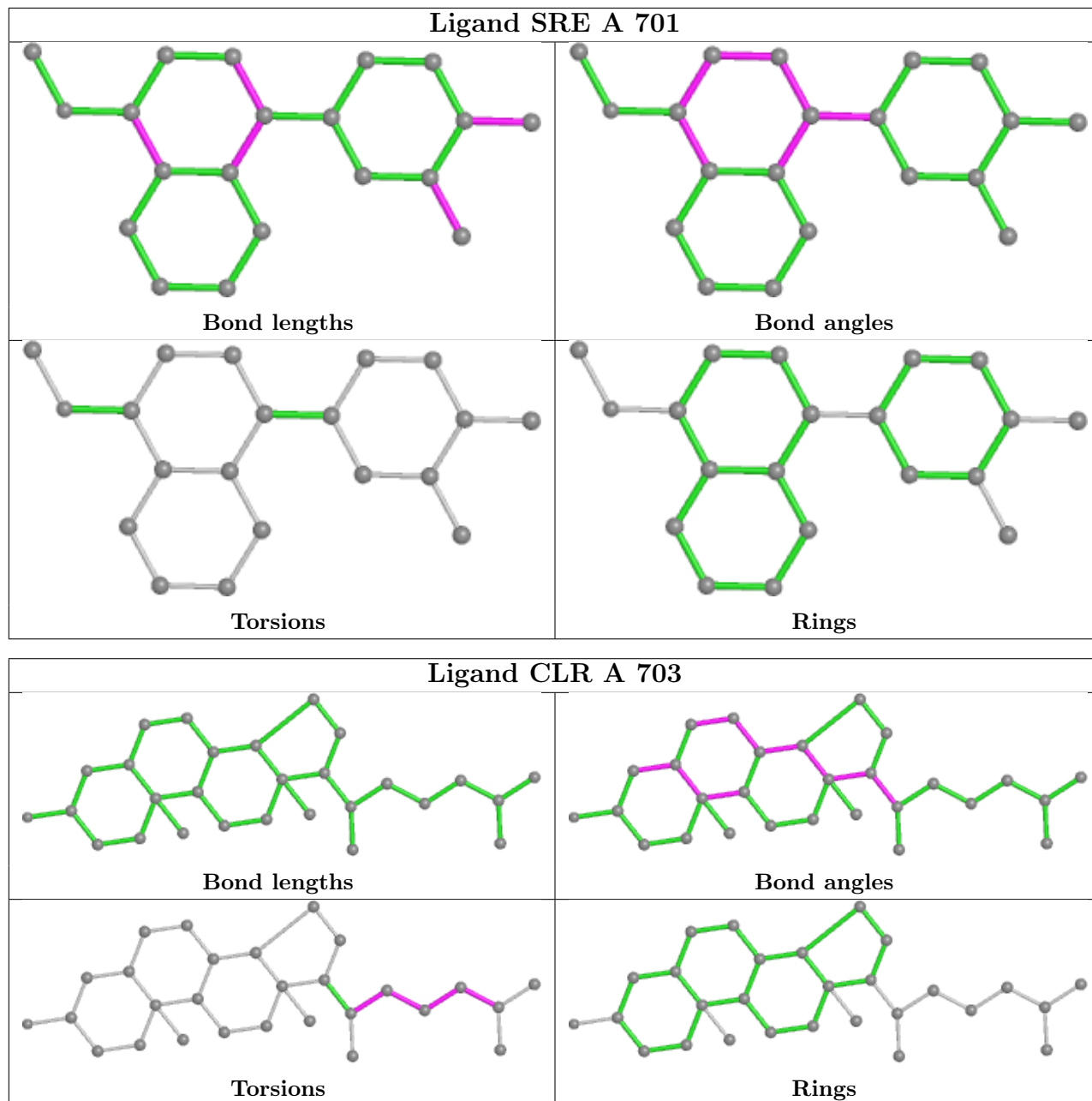
5.7 Other polymers ⓘ

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(Å^2)	Q<0.9
1	A	544/549 (99%)	-0.40	1 (0%)	95	93	124, 195, 255, 299	0
2	B	218/229 (95%)	-0.08	13 (5%)	21	17	122, 206, 318, 404	0
3	C	214/214 (100%)	-0.02	5 (2%)	60	51	136, 220, 300, 341	0
All	All	976/992 (98%)	-0.24	19 (1%)	66	58	122, 202, 296, 404	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	151	PRO	6.7
3	C	40	SER	4.1
2	B	54	ASN	3.9
2	B	115	CYS	3.6
2	B	236	GLU	3.5

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

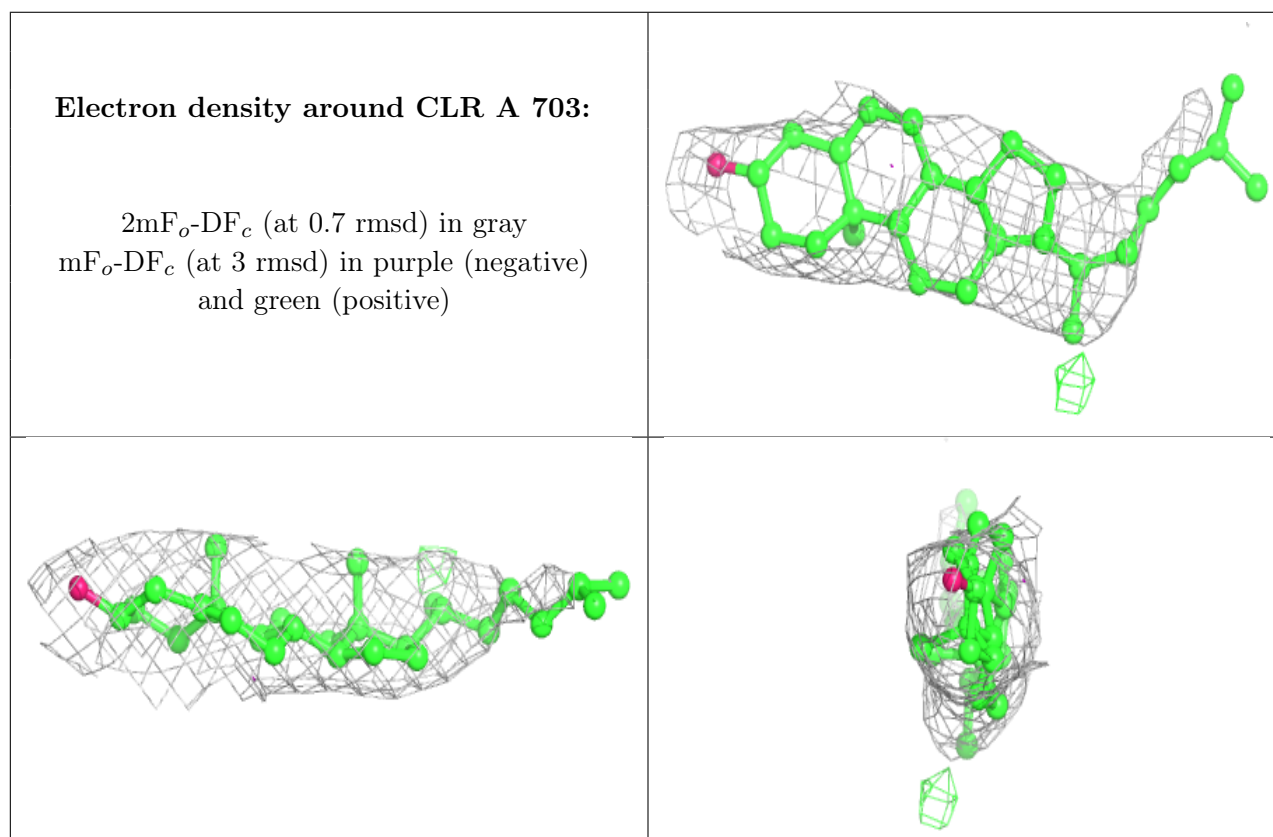There are no non-standard protein/DNA/RNA residues in this entry.

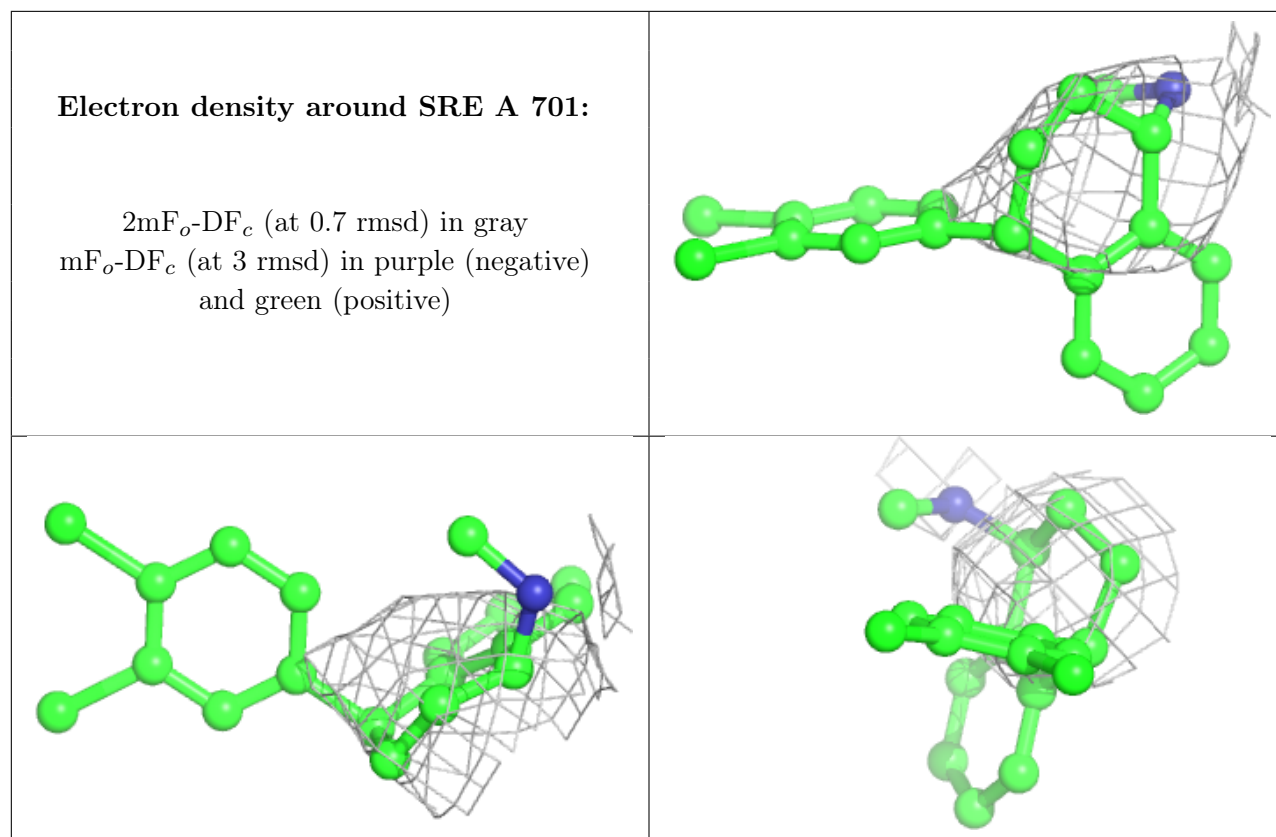## 6.3 Carbohydrates ⓘ

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
8	NA	A	707	1/1	0.75	0.31	196,196,196,196	0
5	NAG	A	704	14/15	0.76	0.54	169,222,241,244	0
6	CLR	A	703	28/28	0.85	0.43	161,184,213,215	0
4	SRE	A	701	20/20	0.85	0.58	194,210,234,239	0
7	CL	A	705	1/1	0.86	0.39	214,214,214,214	0
8	NA	A	706	1/1	0.88	0.11	185,185,185,185	0
5	NAG	A	702	14/15	0.97	0.20	126,149,184,184	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 CLR A 703:

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

Electron density around SRE A 701:

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

6.5   Other polymers ⓘ

There are no such residues in this entry.