WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 3, 2023 – 08:40 am GMT

PDB ID	:	2JFA
Title	:	ESTROGEN RECEPTOR ALPHA LBD IN COMPLEX WITH AN AFFINITY-SELECTED COREPRESSOR PEPTIDE
Authors	:	Heldring, N.; Pawson, T.; McDonnell, D.; Treuter, E.; Gustafsson, J.A.; Pike, A.C.W.
Deposited on	:	2007-01-29
Resolution	:	2.55 Å(reported)

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

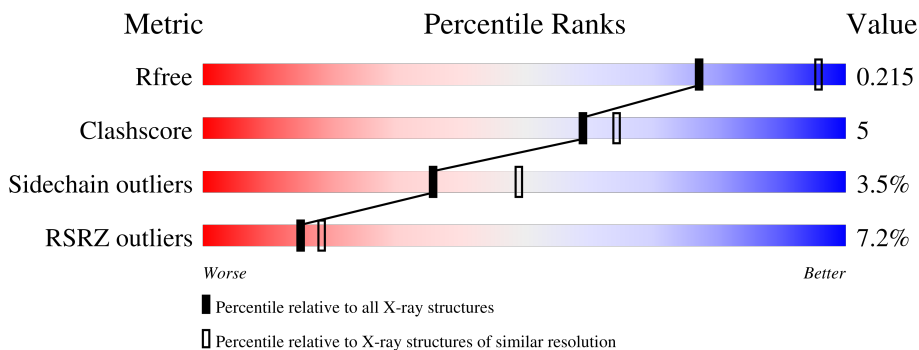|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.55 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		0.215
Clashscore		5
Sidechain outliers		3.5%
RSRZ outliers		7.2%

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	1284 (2.56-2.52)
Clashscore	141614	1332 (2.56-2.52)
Sidechain outliers	138945	1315 (2.56-2.52)
RSRZ outliers	127900	1272 (2.56-2.52)

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 3896 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 ESTROGEN RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	3	0
			1742	1113	300	312	17			

- Molecule 2 is a protein called ESTROGEN RECEPTOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	223	Total	C	N	O	S	0	3	0
			1769	1132	307	313	17			

- Molecule 3 is a protein called COREPRESSOR PEPTIDE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	P	15	Total	C	N	O	0	0	0
			112	73	20	19			
3	Q	15	Total	C	N	O	0	0	0
			99	64	19	16			

- Molecule 4 is RALOXIFENE (three-letter code: RAL) (formula: $C_{28}H_{27}NO_4S$).
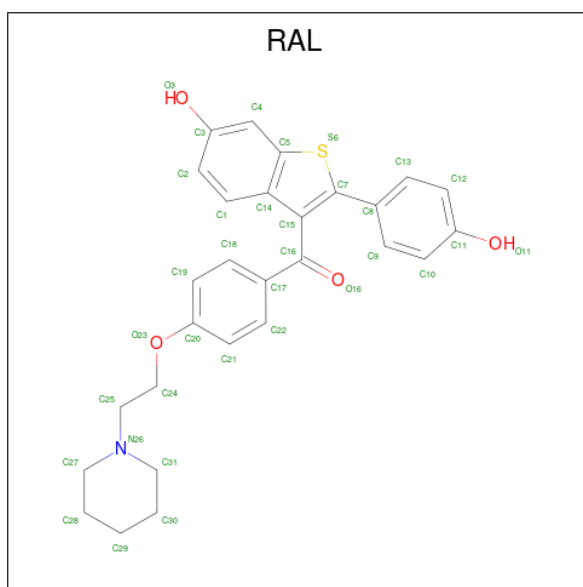
RAL

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	N	O	S	0	0
			34	28	1	4	1		
4	B	1	Total	C	N	O	S	0	0
			34	28	1	4	1		

• Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 5	O 4	S 1	0	0
5	B	1	Total 5	O 4	S 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	46	Total 46	O 46	0	0
6	B	40	Total 40	O 40	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	126.57Å 126.57Å 113.43Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.55 29.37 – 2.55	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-2.55) 99.9 (29.37-2.55)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.75 (at 2.54Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.193 , 0.213 0.194 , 0.215	Depositor DCC
R_{free} test set	1735 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	50.4	Xtriage
Anisotropy	0.058	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.2	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3896	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.85% 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.

4   Model quality ⓘ

4.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, RAL, CCS

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	0.73	2/1789 (0.1%)	0.74	2/2416 (0.1%)
2	B	0.73	3/1808 (0.2%)	0.75	1/2445 (0.0%)
3	P	0.57	0/112	0.47	0/150
3	Q	0.49	0/98	0.66	0/132
All	All	0.72	5/3807 (0.1%)	0.74	3/5143 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	397	GLU	CG-CD	6.17	1.61	1.51
2	B	397	GLU	CB-CG	5.99	1.63	1.52
1	A	471	GLU	CD-OE2	-5.30	1.19	1.25
1	A	385	GLU	CD-OE2	-5.10	1.20	1.25
2	B	471	GLU	CD-OE2	-5.01	1.20	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	471	GLU	OE1-CD-OE2	-6.52	115.48	123.30
1	A	385	GLU	OE1-CD-OE2	-6.15	115.92	123.30
2	B	471	GLU	OE1-CD-OE2	-5.97	116.13	123.30

There are no chirality outliers.

There are no planarity outliers.

4.2   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	1742	0	1755	22	0
2	B	1769	0	1790	21	0
3	P	112	0	113	1	0
3	Q	99	0	98	0	0
4	A	34	0	26	2	0
4	B	34	0	25	2	0
5	A	10	0	0	0	0
5	B	10	0	0	0	0
6	A	46	0	0	0	0
6	B	40	0	0	0	0
All	All	3896	0	3807	36	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 36 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:501[B]:HIS:CD2	2:B:501[B]:HIS:CD2	2.31	1.19
1:A:501[B]:HIS:CD2	2:B:501[B]:HIS:HD2	1.86	0.94
1:A:501[B]:HIS:HD2	2:B:501[B]:HIS:CD2	1.80	0.91
1:A:501[B]:HIS:NE2	2:B:501[B]:HIS:NE2	2.28	0.82
1:A:501[B]:HIS:CD2	2:B:501[B]:HIS:NE2	2.54	0.74

There are no symmetry-related clashes.

4.3 Torsion angles ⓘ

4.3.1 Protein backbone ⓘ

There are no protein backbone outliers to report in this entry.

4.3.2 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	192/226 (85%)	185 (96%)	7 (4%)	35	47
2	B	194/225 (86%)	190 (98%)	4 (2%)	53	68
3	P	10/14 (71%)	7 (70%)	3 (30%)	0	0
3	Q	7/14 (50%)	7 (100%)	0	100	100
All	All	403/479 (84%)	389 (96%)	14 (4%)	36	49

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	413	ASN
2	B	465	THR
3	P	15	ASP
3	P	3	PHE
3	P	10	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

4.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	CCS	B	417	2	8,9,10	2.71	1 (12%)	6,10,12	1.69	1 (16%)

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	CCS	B	417	2	-	4/6/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	417	CCS	CD-SG	-7.39	1.62	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	417	CCS	CB-SG-CD	3.52	137.13	104.44

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	417	CCS	CA-CB-SG-CD
2	B	417	CCS	CE-CD-SG-CB
2	B	417	CCS	SG-CD-CE-OZ1
2	B	417	CCS	SG-CD-CE-OZ2

There are no ring outliers.

No monomer is involved in short contacts.

4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

4.6 Ligand geometry ⓘ

6 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	$\#|Z| > 2$
5	SO4	A	620	-	4,4,4	0.16	0	6,6,6	1.00	0
5	SO4	B	620	-	4,4,4	0.20	0	6,6,6	0.99	0
5	SO4	B	610	-	4,4,4	0.22	0	6,6,6	1.48	2 (33%)
4	RAL	B	600	-	32,38,38	1.63	3 (9%)	42,53,53	1.20	6 (14%)
4	RAL	A	600	-	32,38,38	1.27	2 (6%)	42,53,53	1.04	3 (7%)
5	SO4	A	610	-	4,4,4	0.34	0	6,6,6	0.94	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	RAL	B	600	-	-	1/11/26/26	0/5/5/5
4	RAL	A	600	-	-	0/11/26/26	0/5/5/5

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	600	RAL	C5-S6	-7.44	1.68	1.74
4	A	600	RAL	C5-S6	-5.57	1.69	1.74
4	B	600	RAL	C1-C14	-2.50	1.37	1.42
4	B	600	RAL	C13-C12	2.49	1.41	1.36
4	A	600	RAL	C1-C14	-2.22	1.37	1.42

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	600	RAL	O23-C24-C25	2.83	114.65	107.68
4	A	600	RAL	C24-O23-C20	2.71	125.02	117.93
5	B	610	SO4	O3-S-O2	2.50	122.37	109.31
4	A	600	RAL	C31-N26-C27	2.49	114.44	108.83
4	B	600	RAL	C28-C27-N26	2.30	115.01	111.33

There are no chirality outliers.
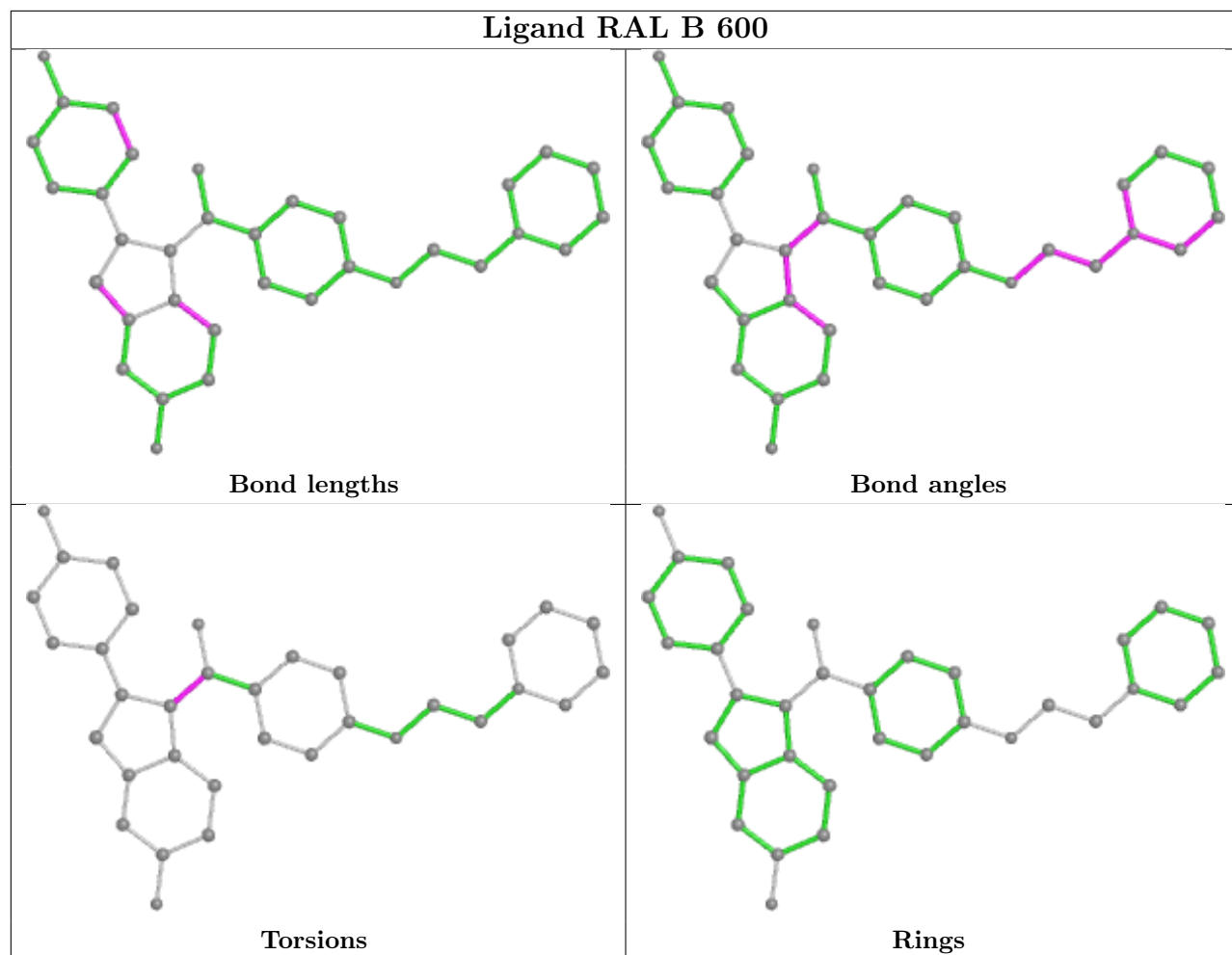
All (1) torsion outliers are listed below:
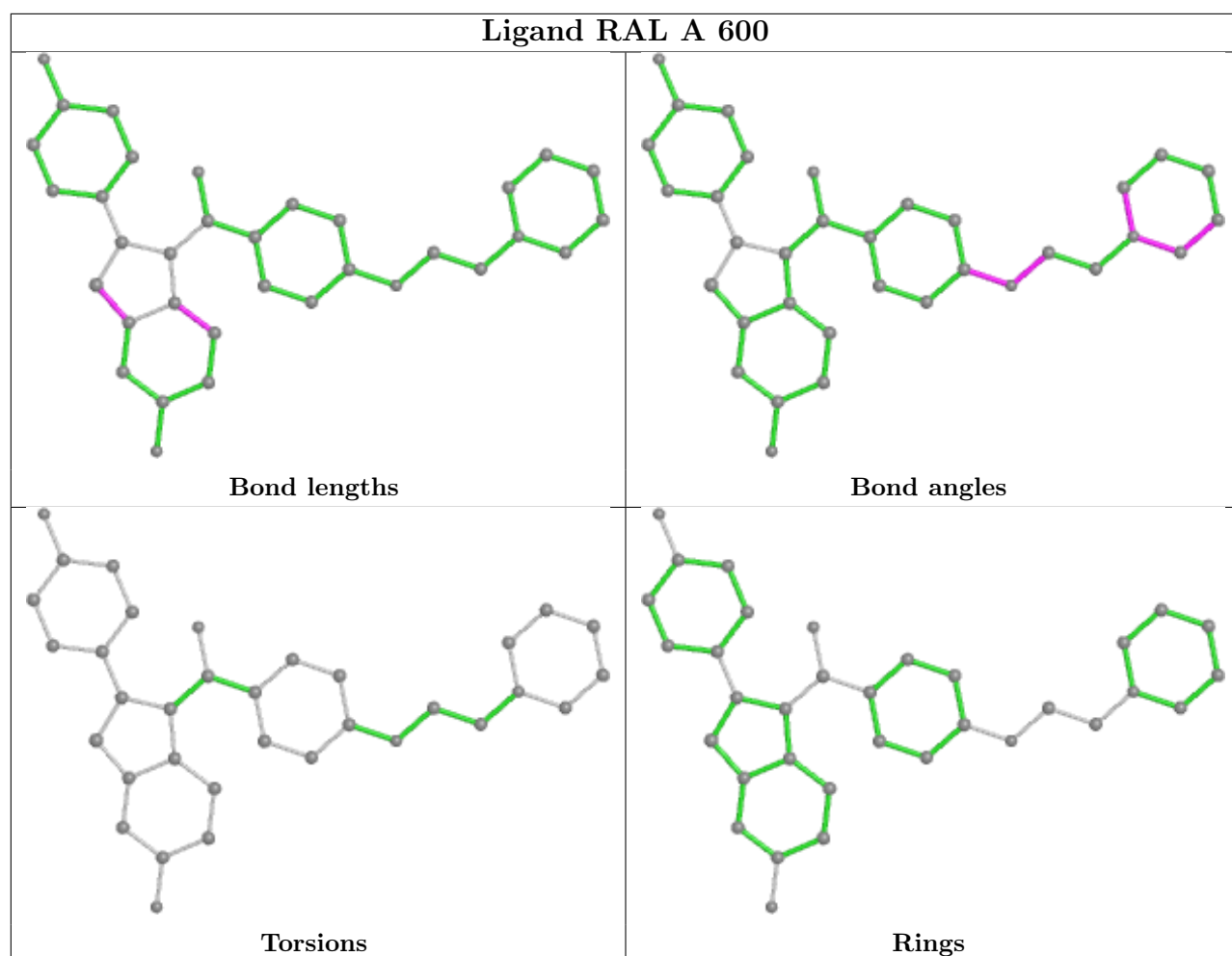
Mol	Chain	Res	Type	Atoms
4	B	600	RAL	C14-C15-C16-O16

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	600	RAL	2	0
4	A	600	RAL	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 RAL B 600

Bond lengths

Bond angles

Torsions

Rings

Ligand RAL A 600

Bond lengths

Bond angles

Torsions

Rings

4.7 Other polymers ⓘ

There are no such residues in this entry.

4.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 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	221/252 (87%)	0.34	19 (8%)	10	12	42, 51, 67, 78	0
2	B	222/252 (88%)	0.15	10 (4%)	33	40	40, 50, 63, 72	0
3	P	15/16 (93%)	1.37	3 (20%)	1	1	61, 67, 81, 82	0
3	Q	15/16 (93%)	0.86	2 (13%)	3	4	59, 66, 81, 82	0
All	All	473/536 (88%)	0.30	34 (7%)	15	18	40, 51, 70, 82	0

The worst 5 of 34 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	464	SER	5.2
3	P	3	PHE	4.7
2	B	526	TYR	4.6
1	A	337	PHE	4.6
1	A	526	TYR	4.4

## 5.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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	CCS	B	417	10/11	0.95	0.15	56,57,75,76	0

5.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.4 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
4	RAL	B	600	34/34	0.91	0.17	45,50,60,61	0
4	RAL	A	600	34/34	0.94	0.16	48,50,55,56	0
5	SO4	A	610	5/5	0.97	0.18	53,53,56,56	0
5	SO4	B	620	5/5	0.97	0.18	57,59,60,62	0
5	SO4	A	620	5/5	0.98	0.19	59,63,65,67	0
5	SO4	B	610	5/5	0.99	0.10	42,44,50,52	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 RAL B 600:

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

Electron density around RAL A 600:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)
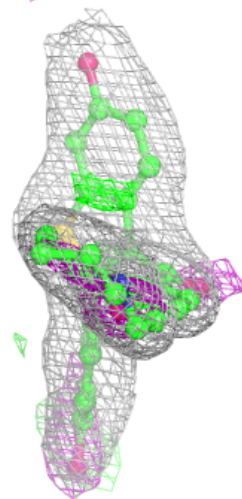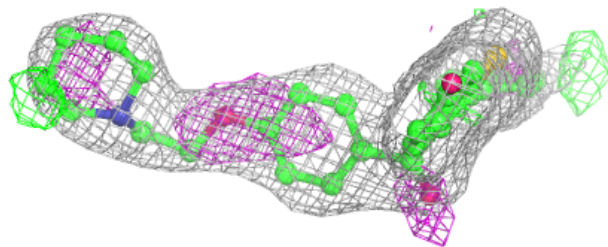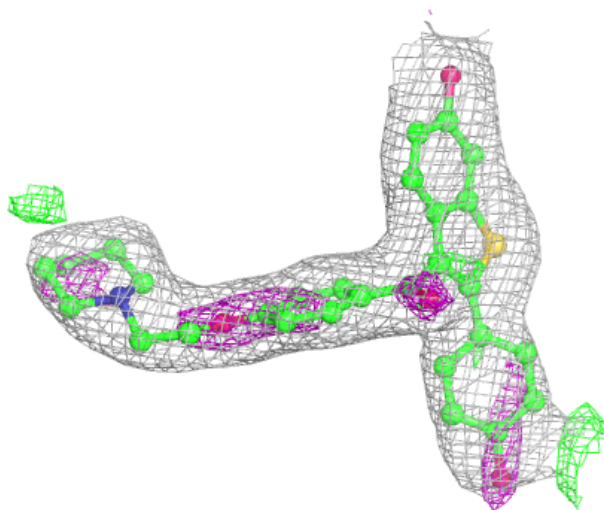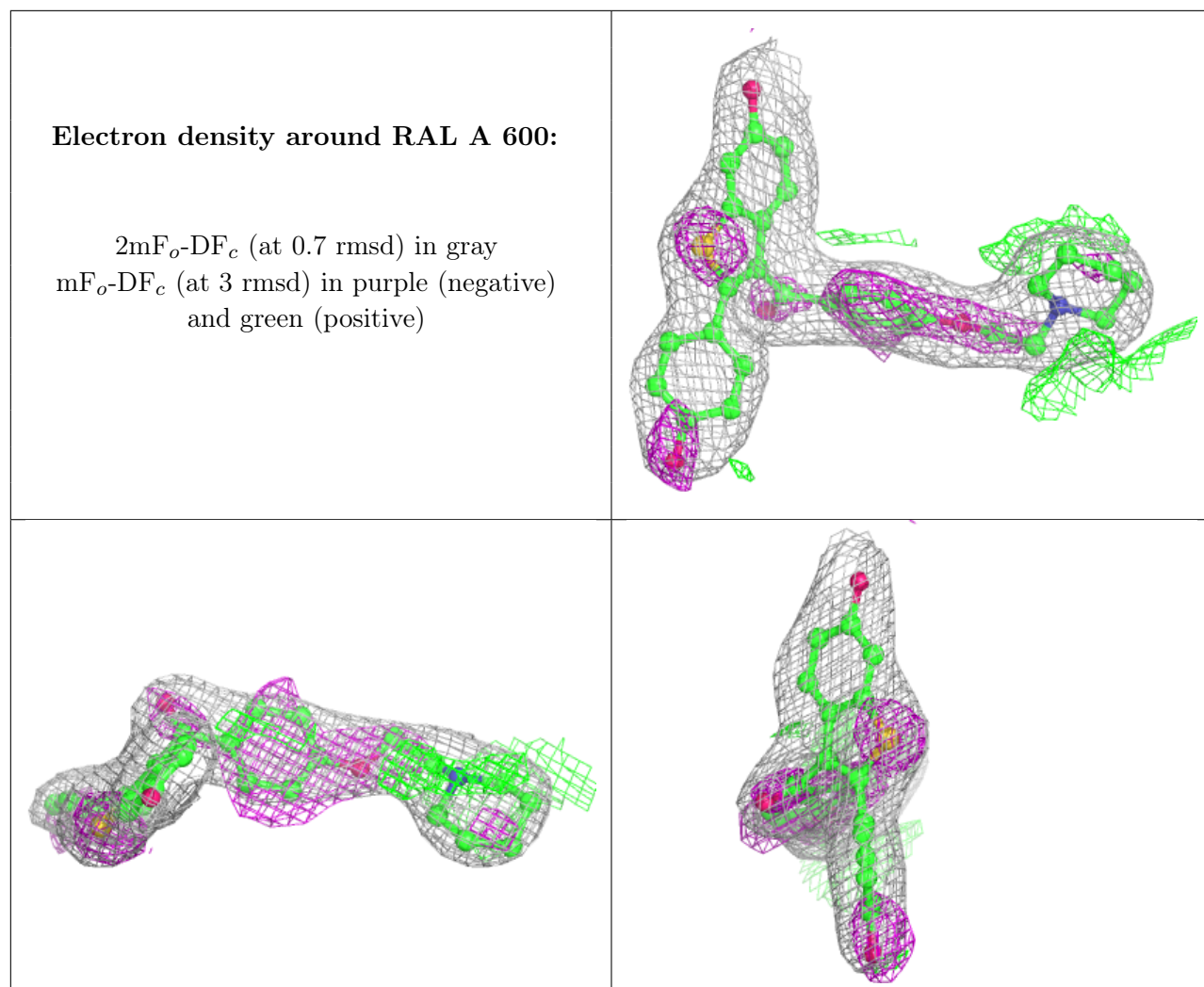
5.5 Other polymers (i)

There are no such residues in this entry.