wwPDB X-ray Structure Validation Summary Report (i)

Nov 5, 2023 – 10:08 AM EST

PDB ID	:	6PIT
Title	:	Estrogen Receptor Alpha Ligand Binding Domain Y537S Mutant in Complex with SRC2 Stapled Peptide 41A and Estradiol
Authors	:	Fanning, S.W.; Montgomery, J.E.; Greene, G.L.; Moellering, R.E.
Deposited on	:	2019-06-27
Resolution	:	2.25 Å(reported)

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

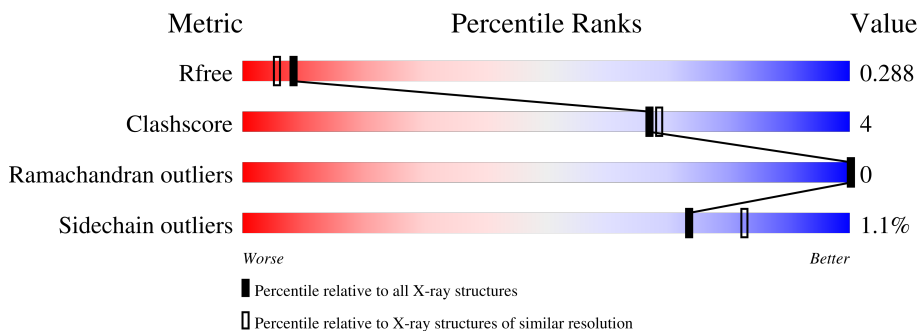|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.25 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		0.288
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		1.1%

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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)

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

Mol	Chain	Length	Quality of chain
1	A	262	79% 5% 16%
2	B	262	76% 7% 17%
3	C	12	25% 75%
3	D	12	8% 25% 33% 33%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3717 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 1723	C 1106	N 294	O 302	S 21	0	3	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	293	GLU	-	expression tag	UNP P03372
A	294	MET	-	expression tag	UNP P03372
A	295	ASP	-	expression tag	UNP P03372
A	296	PRO	-	expression tag	UNP P03372
A	537	SER	TYR	engineered mutation	UNP P03372

- Molecule 2 is a protein called Estrogen receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	218	Total 1711	C 1095	N 297	O 300	S 19	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	293	GLU	-	expression tag	UNP P03372
B	294	MET	-	expression tag	UNP P03372
B	295	ASP	-	expression tag	UNP P03372
B	296	PRO	-	expression tag	UNP P03372
B	537	SER	TYR	engineered mutation	UNP P03372

- Molecule 3 is a protein called Stapled Peptide 41A.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	8	Total 74	C 48	N 13	O 13	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	3	Total C N O 17 11 3 3				0	0	0

- Molecule 4 is ESTRADIOL (three-letter code: EST) (formula: $C_{18}H_{24}O_2$).
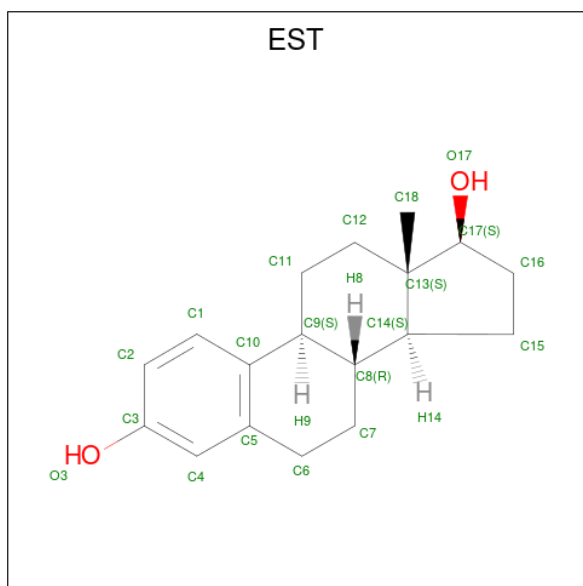
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total C O 20 18 2			0	0
4	B	1	Total C O 20 18 2			0	0

- Molecule 5 is water.

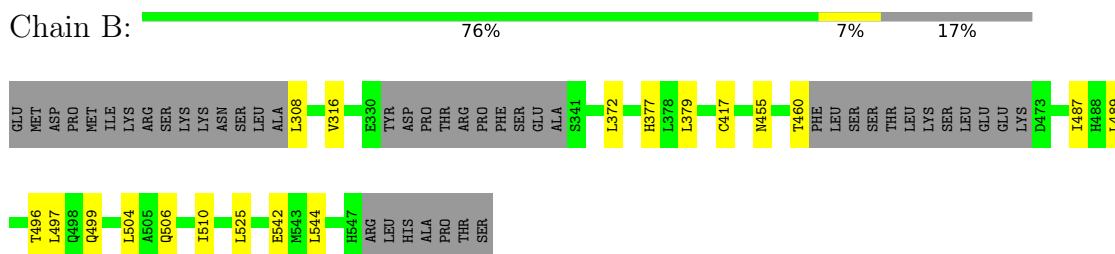| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	84	Total O 84 84		0	0
5	B	68	Total O 68 68		0	0

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. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
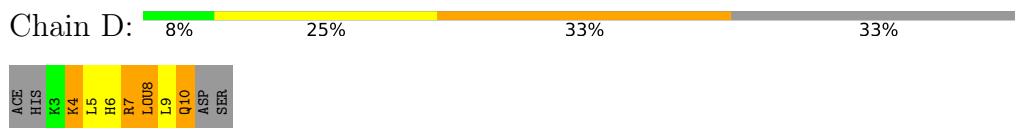
- Molecule 1: Estrogen receptor

Chain A:
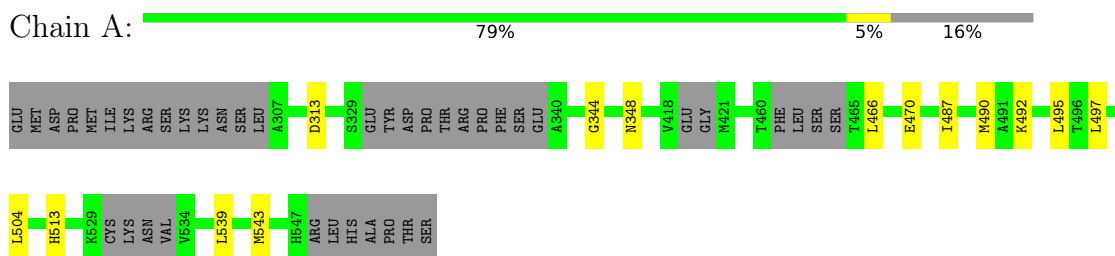
- Molecule 2: Estrogen receptor

Chain B:

- Molecule 3: Stapled Peptide 41A

Chain D:

- Molecule 3: Stapled Peptide 41A

Chain C:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	55.53Å 71.03Å 58.08Å 90.00° 108.68° 90.00°	Depositor
Resolution (Å)	43.50 – 2.25 43.50 – 2.19	Depositor EDS
% Data completeness (in resolution range)	99.5 (43.50-2.25) 84.4 (43.50-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 2.20Å)	Xtriage
Refinement program	PHENIX v1.9	Depositor
R, R_{free}	0.214 , 0.247 0.264 , 0.288	Depositor DCC
R_{free} test set	945 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å²)	30.0	Xtriage
Anisotropy	0.160	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.39 , 47.1	EDS
L-test for twinning [2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.035 for l,-k,h	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	3717	wwPDB-VP
Average B, all atoms (Å²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.64% 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: LOU, SCH, EST

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	0.20	0/1737	0.36	0/2347
2	B	0.24	0/1730	0.37	0/2336
3	C	0.13	0/11	0.37	0/12
3	D	0.90	0/53	1.08	0/68
All	All	0.24	0/3531	0.39	0/4763

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
3	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	7	ARG	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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1723	0	1725	8	0
2	B	1711	0	1756	16	0
3	C	17	0	11	0	0
3	D	74	0	50	12	0
4	A	20	0	23	0	0
4	B	20	0	23	0	0
5	A	84	0	0	0	0
5	B	68	0	0	0	0
All	All	3717	0	3588	28	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 28 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)
3:D:8:LOU:C22	3:D:8:LOU:N23	1.69	1.52
3:D:8:LOU:C22	3:D:8:LOU:C24	2.03	1.35
3:D:8:LOU:C22	3:D:8:LOU:O25	2.12	0.95
2:B:542:GLU:OE1	3:D:4:LYS:HB2	1.70	0.92
2:B:542:GLU:CD	3:D:4:LYS:HB2	2.01	0.79

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	212/262 (81%)	210 (99%)	2 (1%)	0	100	100
2	B	211/262 (80%)	208 (99%)	3 (1%)	0	100	100
3	D	5/12 (42%)	5 (100%)	0	0	100	100
All	All	428/536 (80%)	423 (99%)	5 (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	182/236 (77%)	182 (100%)	0	100	100
2	B	189/236 (80%)	188 (100%)	1 (0%)	88	92
3	C	1/10 (10%)	1 (100%)	0	100	100
3	D	5/10 (50%)	2 (40%)	3 (60%)	0	0
All	All	377/492 (77%)	373 (99%)	4 (1%)	73	82

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	308	LEU
3	D	4	LYS
3	D	5	LEU
3	D	10	GLN

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	476	HIS
2	B	499	GLN
3	D	6	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

5 non-standard protein/DNA/RNA residues 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	#$|Z| > 2$
3	LOU	D	8	3	21,22,24	4.71	16 (76%)	21,34,38	11.44	14 (66%)
1	SCH	A	381[A]	-	6,7,8	1.08	0	3,7,9	1.15	0
1	SCH	A	381[B]	-	6,7,8	1.10	0	3,7,9	1.06	0
2	SCH	B	417	2	6,7,8	1.99	1 (16%)	3,7,9	1.14	0
3	LOU	C	8	3	3,3,24	0.69	0	0,2,38	-	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	LOU	D	8	3	-	7/8/51/54	0/4/3/3
1	SCH	A	381[A]	-	-	0/2/6/8	-
1	SCH	A	381[B]	-	-	0/2/6/8	-
2	SCH	B	417	2	-	1/2/6/8	-
3	LOU	C	8	3	-	0/0/1/54	-

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	8	LOU	C22-N23	10.92	1.69	1.47
3	D	8	LOU	C54-CA	-9.42	1.42	1.54
3	D	8	LOU	O56-C55	-7.25	1.35	1.45
3	D	8	LOU	C28-C57	-6.94	1.43	1.54
3	D	8	LOU	C55-C27	-5.73	1.44	1.55

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	8	LOU	C22-N23-C24	-44.44	83.19	123.43
3	D	8	LOU	C22-N23-C26	14.47	136.53	123.43
3	D	8	LOU	C28-C24-N23	10.13	117.12	108.45
3	D	8	LOU	C26-N23-C24	-9.85	104.11	113.18

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	8	LOU	O56-C57-C58	-9.39	89.85	104.01

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	8	LOU	C21-C22-N23-C24
3	D	8	LOU	C55-C54-CA-N
3	D	8	LOU	C55-C54-CA-C
3	D	8	LOU	CA-C54-C55-C27
3	D	8	LOU	CA-C54-C55-O56

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	8	LOU	4	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 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	EST	A	601	-	23,23,23	1.44	2 (8%)	36,36,36	1.02	2 (5%)
4	EST	B	601	-	23,23,23	1.44	2 (8%)	36,36,36	1.06	2 (5%)

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	EST	A	601	-	-	-	0/4/4/4
4	EST	B	601	-	-	-	0/4/4/4

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	601	EST	C5-C10	5.66	1.49	1.40
4	A	601	EST	C5-C10	5.61	1.49	1.40
4	B	601	EST	C10-C9	-2.53	1.48	1.52
4	A	601	EST	C10-C9	-2.51	1.48	1.52

All (4) bond angle outliers are listed below:

| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	601	EST	C13-C14-C8	-2.78	110.27	114.38
4	A	601	EST	C13-C14-C8	-2.61	110.52	114.38
4	A	601	EST	C14-C13-C17	2.34	101.74	99.27
4	B	601	EST	C14-C13-C17	2.32	101.72	99.27

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 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 EST A 601

Ligand EST B 601
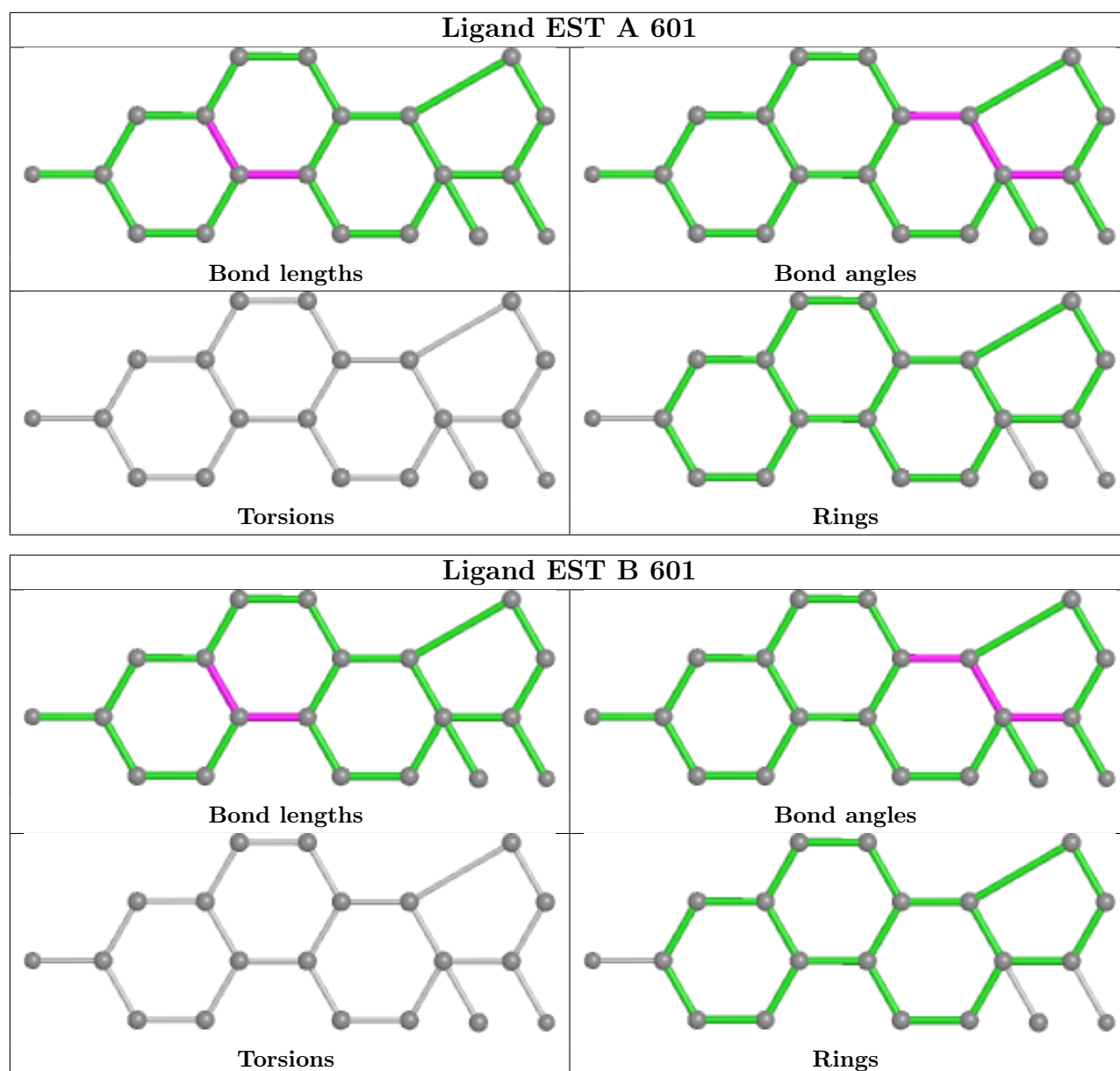
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

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

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

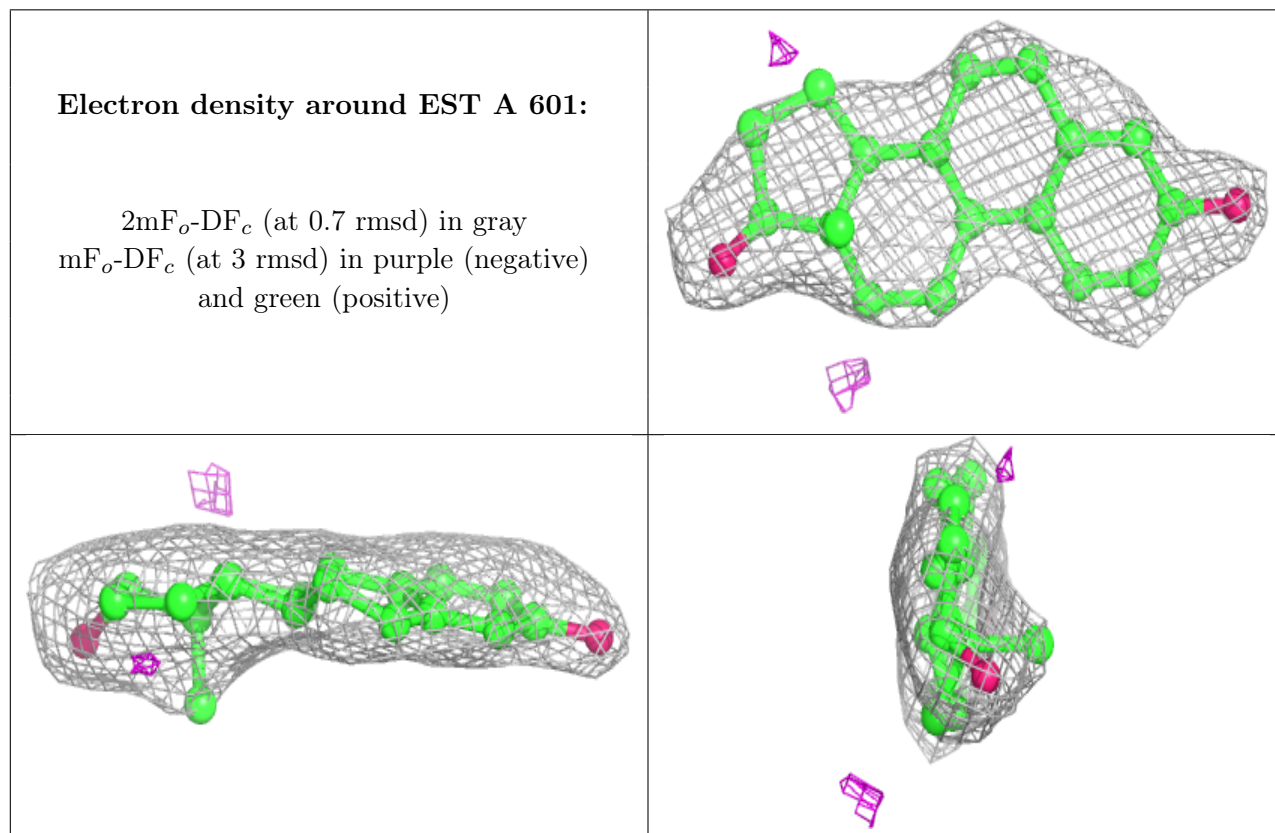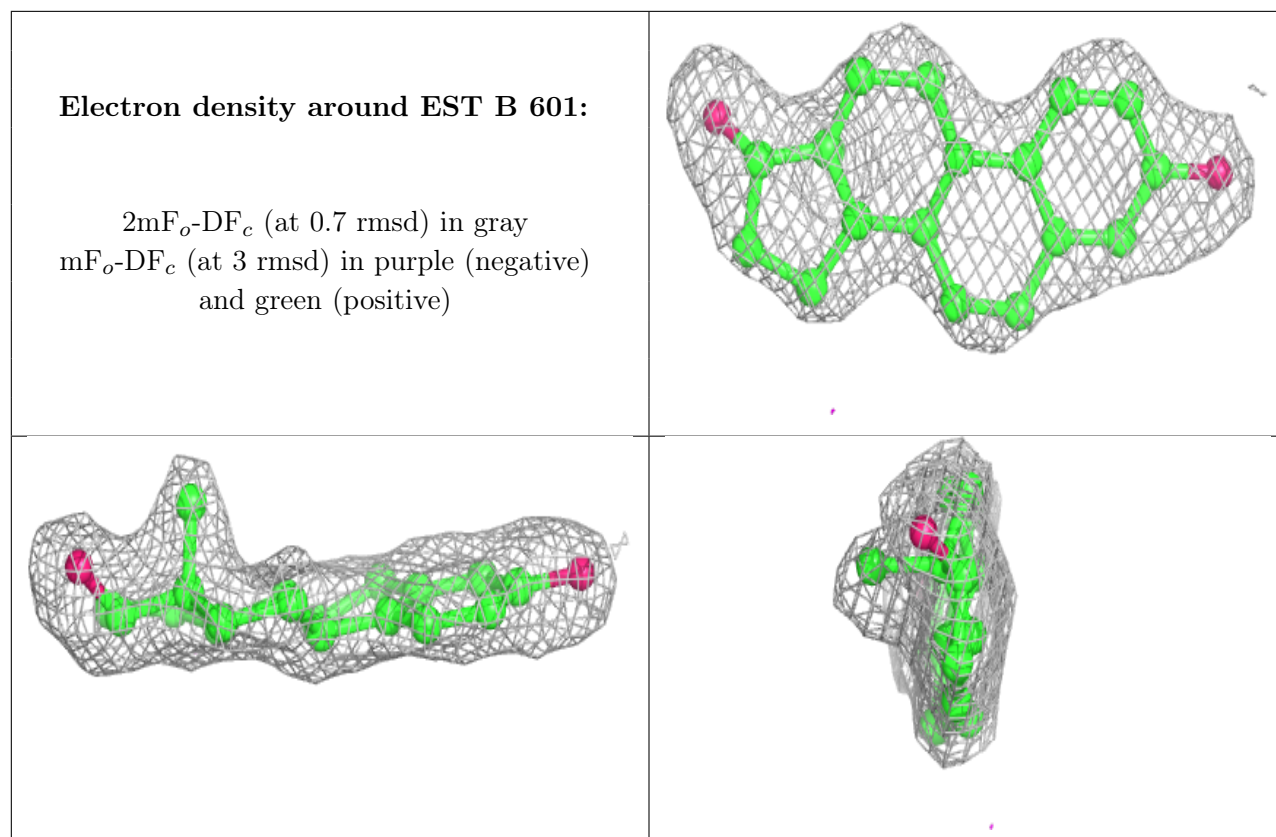## 6.3 Carbohydrates ⓘ

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

6.4 Ligands ⓘ

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

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around EST A 601:

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

Electron density around EST B 601:

$2mF_o$-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 ⓘ

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