wwPDB X-ray Structure Validation Summary Report (i)

Feb 17, 2024 – 10:26 PM EST

PDB ID	:	4DM8
Title	:	Crystal structure of RARb LBD in complex with 9cis retinoic acid
Authors	:	Osz, J.; Br livet, Y.; Peluso-Iltis, C.; Cura, V.; Eiler, S.; Ruff, M.; Bourguet, W.; Rochel, N.; Moras, D.
Deposited on	:	2012-02-07
Resolution	:	2.30 Å(reported)

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

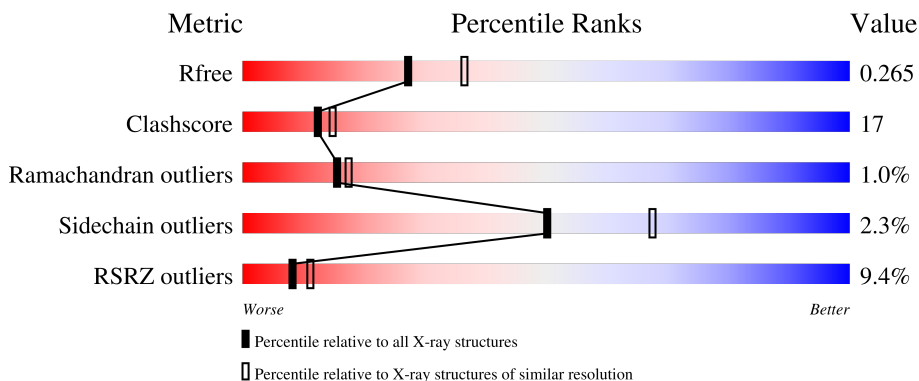|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 2.30 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		0.265
Clashscore		17
Ramachandran outliers		1.0%
Sidechain outliers		2.3%
RSRZ outliers		9.4%

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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	267	11% 62% 26% · 11%
1	B	267	6% 64% 24% · 10%
2	C	25	12% 12% 8% 68%
2	D	25	16% 16% 68%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 4083 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 Retinoic acid receptor beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	238	Total 1879	C 1196	N 315	O 353	S 15	36	0	0
1	B	239	Total 1890	C 1204	N 317	O 354	S 15	2	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	157	MET	-	initiating methionine	UNP P10826
A	158	GLY	-	expression tag	UNP P10826
A	159	SER	-	expression tag	UNP P10826
A	160	SER	-	expression tag	UNP P10826
A	161	HIS	-	expression tag	UNP P10826
A	162	HIS	-	expression tag	UNP P10826
A	163	HIS	-	expression tag	UNP P10826
A	164	HIS	-	expression tag	UNP P10826
A	165	HIS	-	expression tag	UNP P10826
A	166	HIS	-	expression tag	UNP P10826
A	167	SER	-	expression tag	UNP P10826
A	168	SER	-	expression tag	UNP P10826
A	169	GLY	-	expression tag	UNP P10826
A	170	LEU	-	expression tag	UNP P10826
A	171	VAL	-	expression tag	UNP P10826
A	172	PRO	-	expression tag	UNP P10826
A	173	ARG	-	expression tag	UNP P10826
A	174	GLY	-	expression tag	UNP P10826
A	175	SER	-	expression tag	UNP P10826
A	176	HIS	-	expression tag	UNP P10826
A	177	MET	-	expression tag	UNP P10826
B	1157	MET	-	initiating methionine	UNP P10826
B	1158	GLY	-	expression tag	UNP P10826
B	1159	SER	-	expression tag	UNP P10826
B	1160	SER	-	expression tag	UNP P10826

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1161	HIS	-	expression tag	UNP P10826
B	1162	HIS	-	expression tag	UNP P10826
B	1163	HIS	-	expression tag	UNP P10826
B	1164	HIS	-	expression tag	UNP P10826
B	1165	HIS	-	expression tag	UNP P10826
B	1166	HIS	-	expression tag	UNP P10826
B	1167	SER	-	expression tag	UNP P10826
B	1168	SER	-	expression tag	UNP P10826
B	1169	GLY	-	expression tag	UNP P10826
B	1170	LEU	-	expression tag	UNP P10826
B	1171	VAL	-	expression tag	UNP P10826
B	1172	PRO	-	expression tag	UNP P10826
B	1173	ARG	-	expression tag	UNP P10826
B	1174	GLY	-	expression tag	UNP P10826
B	1175	SER	-	expression tag	UNP P10826
B	1176	HIS	-	expression tag	UNP P10826
B	1177	MET	-	expression tag	UNP P10826

- Molecule 2 is a protein called Nuclear receptor coactivator 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	8	Total 71	C 47	N 15	O 9	0	0	0
2	D	8	Total 71	C 47	N 15	O 9	4	0	0

- Molecule 3 is RETINOIC ACID (three-letter code: REA) (formula: $C_{20}H_{28}O_2$).
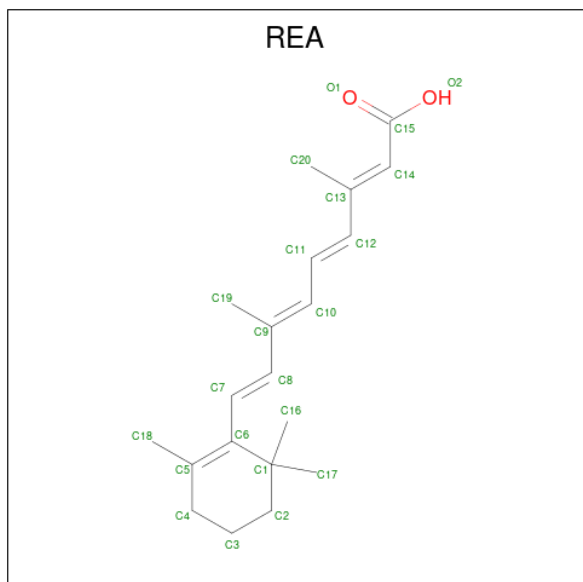
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 22	C 20	O 2	0	0
3	B	1	Total 22	C 20	O 2	0	0

- Molecule 4 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total 31	O 31	0	0
4	B	91	Total 91	O 91	0	0
4	C	2	Total 2	O 2	0	0
4	D	4	Total 4	O 4	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.
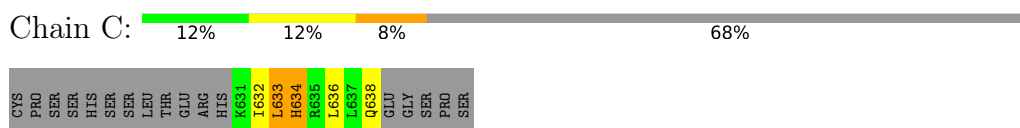
• Molecule 1: Retinoic acid receptor beta
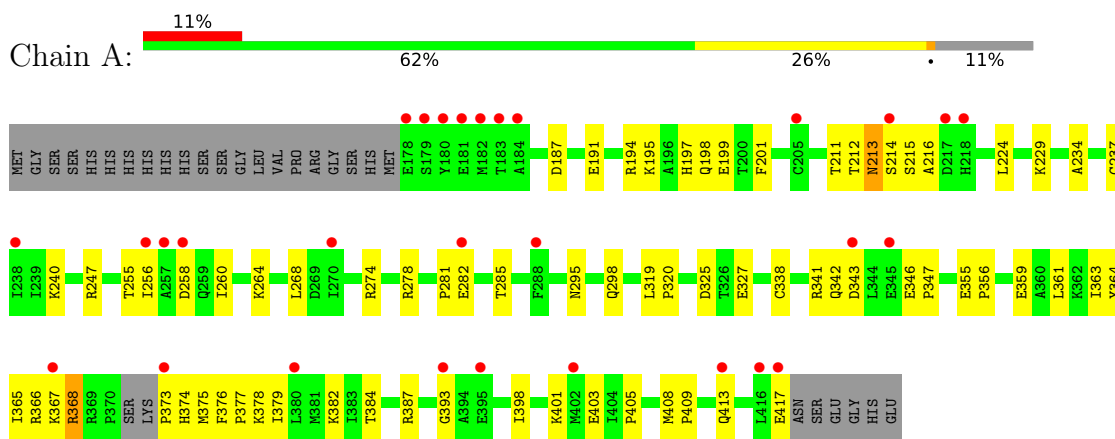
Chain A:

• Molecule 1: Retinoic acid receptor beta
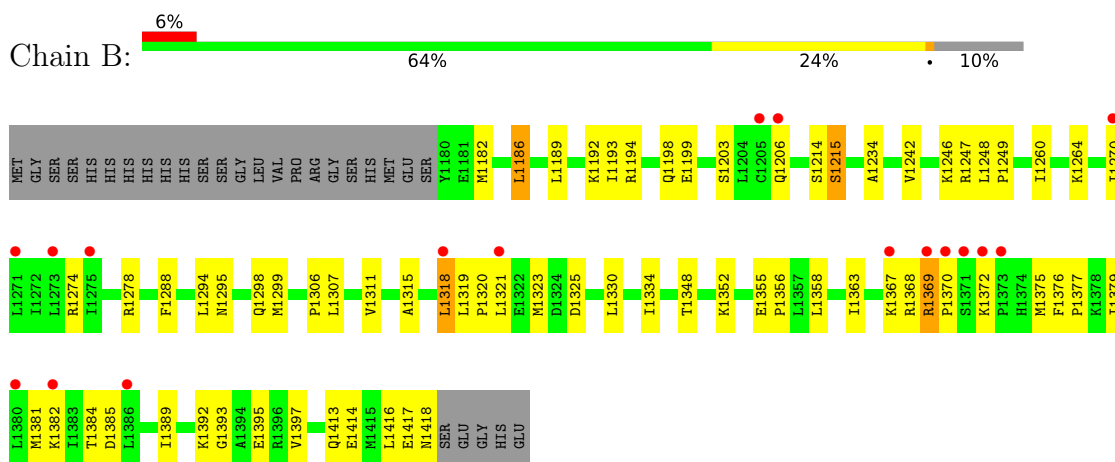
Chain B:

• Molecule 2: Nuclear receptor coactivator 1

Chain C:

• Molecule 2: Nuclear receptor coactivator 1

Chain D:

16%   16%   68%

CYS PRO SER SER HIS SER SER LEU THR GLU ARG HIS **K1631** **I1632** **L1633** **H1634** **R1635** **L1636** **L1637** **Q1638** GLU GLY SER PRO SER

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	58.48Å 83.62Å 108.78Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 29.24 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.30) 94.5 (29.24-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.22 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.224 , 0.277 0.212 , 0.265	Depositor DCC
R_{free} test set	1189 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 52.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4083	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% 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: REA

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	0.31	0/1908	0.56	1/2576 (0.0%)
1	B	0.35	0/1921	0.57	0/2595
2	C	0.24	0/71	0.48	0/93
2	D	0.28	0/71	0.63	0/93
All	All	0.33	0/3971	0.56	1/5357 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	373	PRO	N-CA-CB	5.65	110.08	103.30

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	1879	0	1937	57	0
1	B	1890	0	1956	76	0
2	C	71	0	84	4	0
2	D	71	0	84	6	0
3	A	22	0	27	1	0
3	B	22	0	27	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	31	0	0	3	0
4	B	91	0	0	7	0
4	C	2	0	0	0	0
4	D	4	0	0	0	0
All	All	4083	0	4115	135	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 17.

The worst 5 of 135 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)
4:B:1601:HOH:O	2:D:1631:LYS:HB3	1.69	0.91
1:A:191:GLU:HG3	1:A:195:LYS:HE3	1.60	0.83
1:A:341:ARG:HH22	1:A:387:ARG:HH22	1.25	0.81
1:A:375:MET:O	1:A:379:ILE:HG12	1.81	0.80
1:B:1376:PHE:HB3	1:B:1377:PRO:HD3	1.65	0.78

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	234/267 (88%)	212 (91%)	18 (8%)	4 (2%)	9	8
1	B	237/267 (89%)	229 (97%)	7 (3%)	1 (0%)	34	42
2	C	6/25 (24%)	6 (100%)	0	0	100	100
2	D	6/25 (24%)	6 (100%)	0	0	100	100
All	All	483/584 (83%)	453 (94%)	25 (5%)	5 (1%)	15	17

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	213	ASN
1	A	215	SER
1	A	405	PRO
1	A	214	SER
1	B	1215	SER

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	211/238 (89%)	208 (99%)	3 (1%)	67	81
1	B	213/238 (90%)	209 (98%)	4 (2%)	57	73
2	C	8/24 (33%)	5 (62%)	3 (38%)	0	0
2	D	8/24 (33%)	8 (100%)	0	100	100
All	All	440/524 (84%)	430 (98%)	10 (2%)	50	67

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

Mol	Chain	Res	Type
2	C	633	LEU
2	C	634	HIS
2	C	638	GLN
1	B	1186	LEU
1	B	1318	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1259	GLN
1	B	1298	GLN
2	C	634	HIS
1	B	1316	ASN
1	A	374	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 ⓘ

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	#$|Z| > 2$
3	REA	A	501	-	22,22,22	2.14	4 (18%)	30,30,30	1.59	6 (20%)
3	REA	B	1501	-	22,22,22	2.13	2 (9%)	30,30,30	1.73	7 (23%)

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	REA	A	501	-	-	2/15/32/32	0/1/1/1
3	REA	B	1501	-	-	2/15/32/32	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1501	REA	C1-C6	7.20	1.63	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	REA	C1-C6	7.12	1.63	1.53
3	A	501	REA	C5-C6	4.16	1.41	1.34
3	B	1501	REA	C5-C6	4.15	1.41	1.34
3	A	501	REA	C18-C5	2.05	1.54	1.50

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1501	REA	C18-C5-C6	3.90	128.91	124.53
3	A	501	REA	C18-C5-C6	3.81	128.80	124.53
3	A	501	REA	C15-C14-C13	3.46	133.90	128.50
3	B	1501	REA	C15-C14-C13	3.21	133.52	128.50
3	B	1501	REA	C10-C11-C12	-3.07	113.63	123.22

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	REA	C11-C10-C9-C19
3	B	1501	REA	C11-C10-C9-C19
3	A	501	REA	C11-C10-C9-C8
3	B	1501	REA	C11-C10-C9-C8

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	REA	1	0
3	B	1501	REA	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.
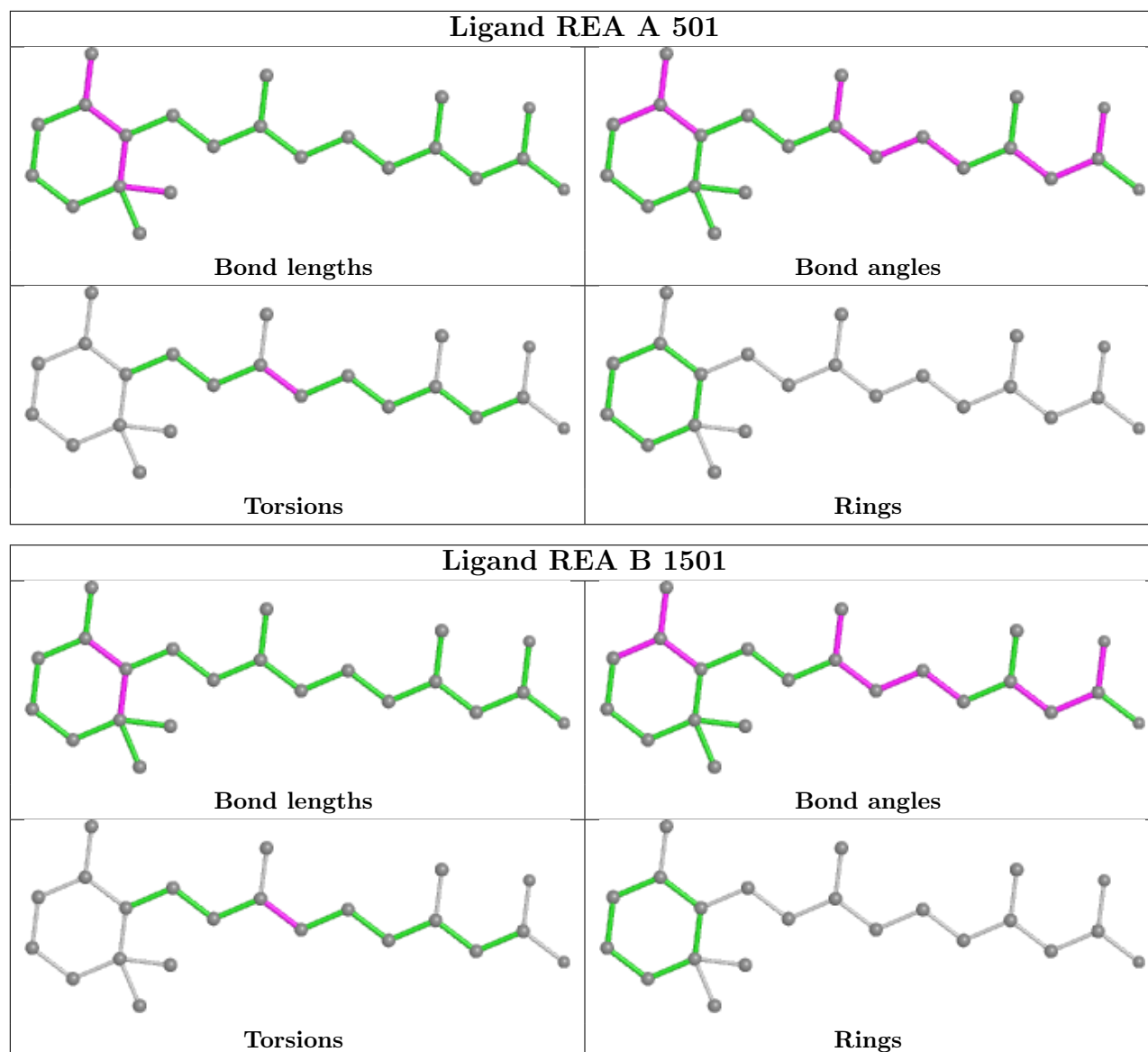
5.7 Other polymers ⓘ

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	235/267 (88%)	0.80	29 (12%)	4	6	24, 43, 70, 83	5 (2%)
1	B	239/267 (89%)	0.33	17 (7%)	16	21	15, 33, 56, 80	1 (0%)
2	C	8/25 (32%)	0.34	0	100	100	37, 42, 48, 57	0
2	D	8/25 (32%)	-0.05	0	100	100	27, 32, 40, 42	1 (12%)
All	All	490/584 (83%)	0.55	46 (9%)	8	11	15, 38, 66, 83	7 (1%)

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	205	CYS	7.5
1	B	1370	PRO	6.5
1	B	1205	CYS	5.9
1	A	178	GLU	5.5
1	B	1206	GLN	4.8

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 ⓘ
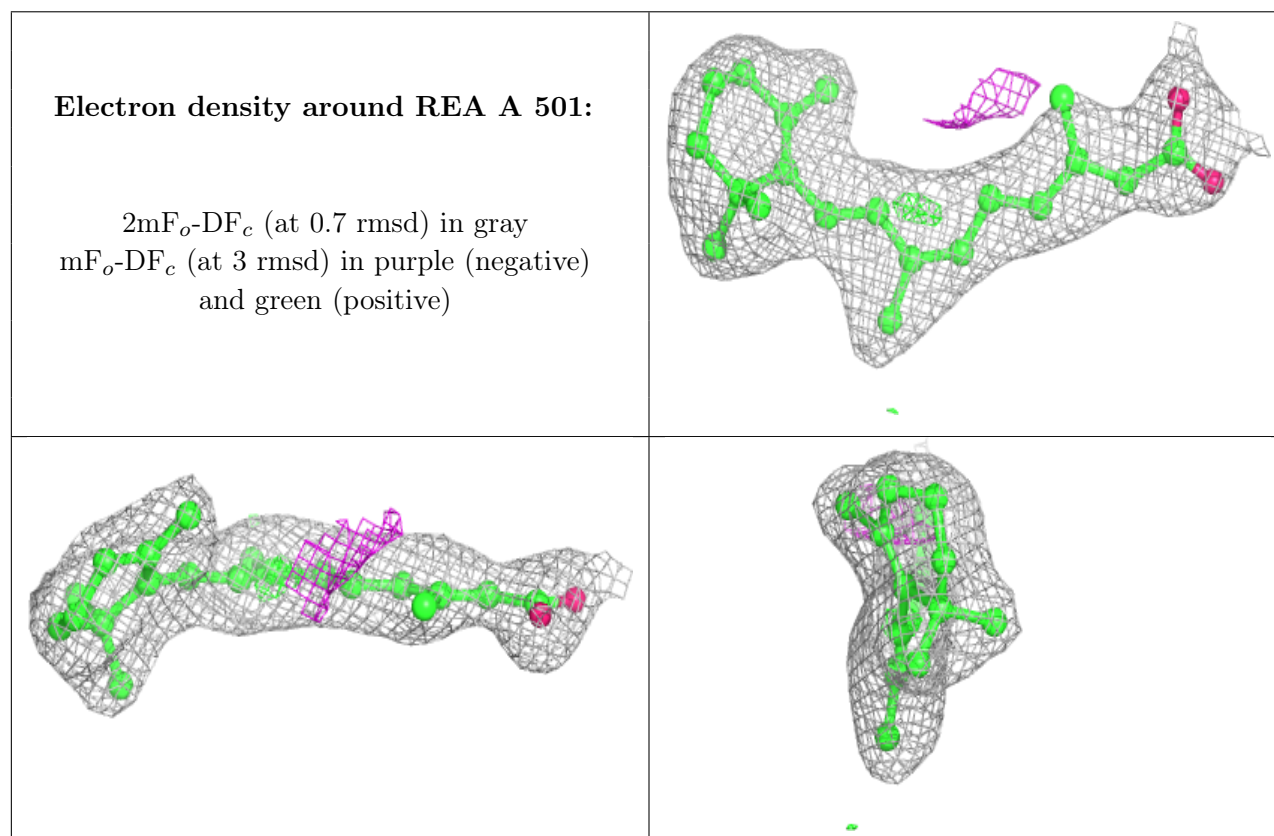
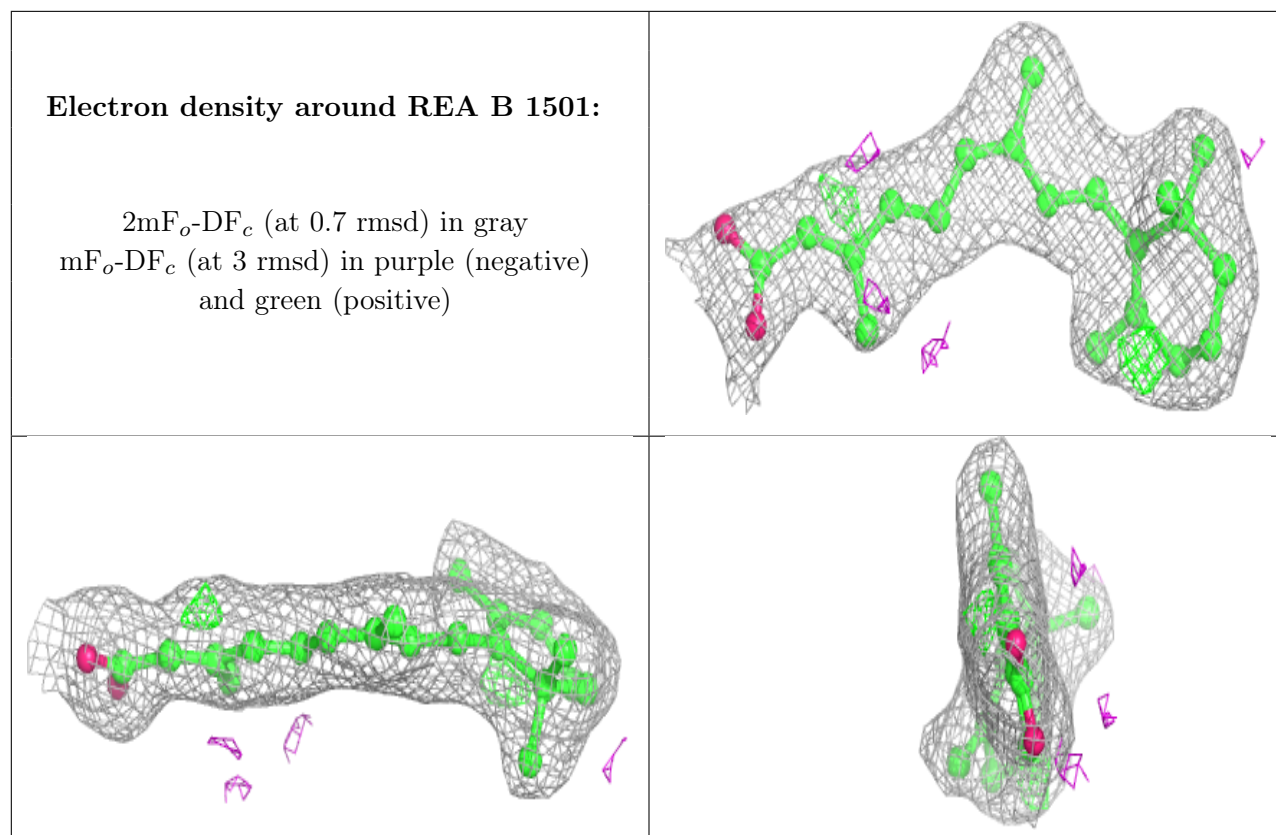There are no monosaccharides in this entry.

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
3	REA	A	501	22/22	0.88	0.33	33,37,40,40	0
3	REA	B	1501	22/22	0.88	0.31	25,33,41,47	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 REA A 501:

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

Electron density around REA B 1501:

$2mF_o$-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.