



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2024 – 04:01 PM EDT

PDB ID : 4HHH
Title : Structure of Pisum sativum Rubisco
Authors : Loewen, P.C.; Didychuk, A.L.; Switala, J.; Loewen, M.C.
Deposited on : 2012-10-09
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 2022.3.0, CSD as543be (2022)
Xtrriage (Phenix) : 1.20.1
EDS : 2.37.1
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) : 2.37.1

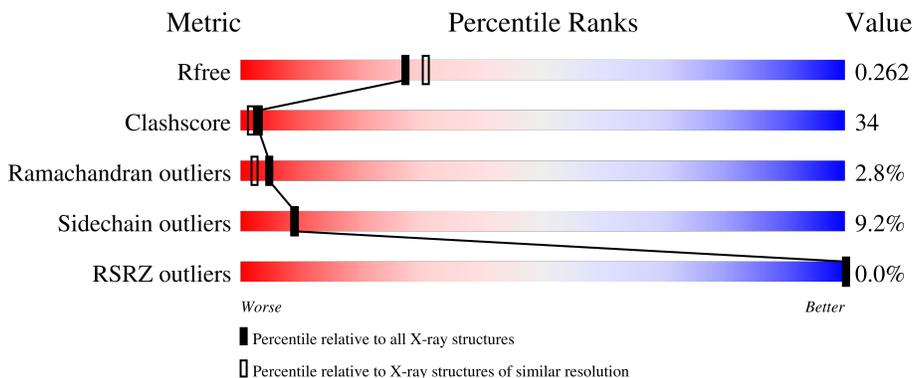
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 $\leq 5\%$. 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	475	
1	B	475	
1	C	475	
1	D	475	
2	S	123	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	T	123	
2	U	123	
2	V	123	

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	RUB	B	501	-	X	-	-

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 19680 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.

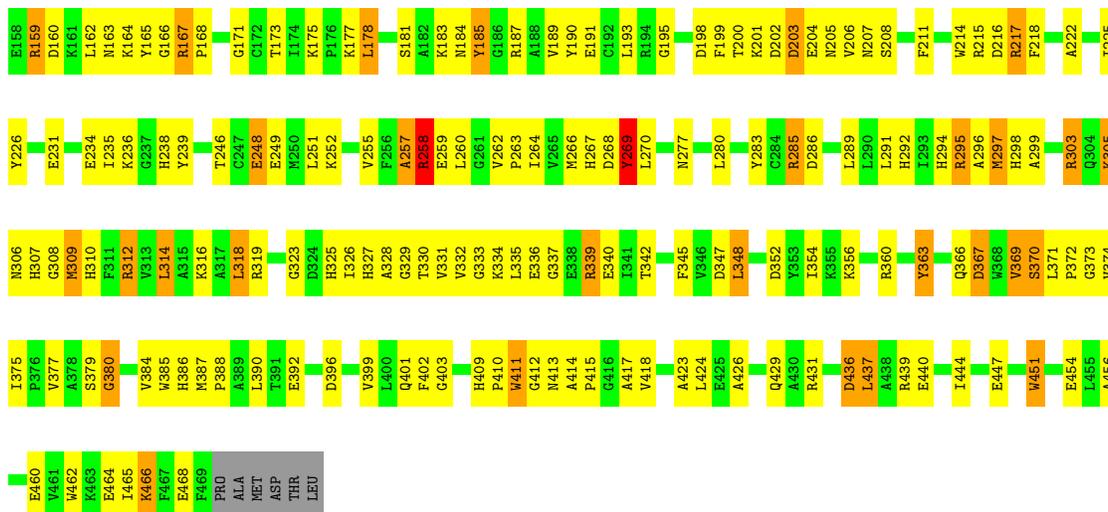
- Molecule 1 is a protein called Ribulose biphosphate carboxylase large chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	458	Total 3648	C 2332	N 634	O 664	S 18	0	11	0
1	B	458	Total 3653	C 2340	N 635	O 660	S 18	0	13	0
1	C	458	Total 3662	C 2338	N 644	O 662	S 18	0	12	0
1	D	458	Total 3647	C 2323	N 643	O 663	S 18	0	9	0

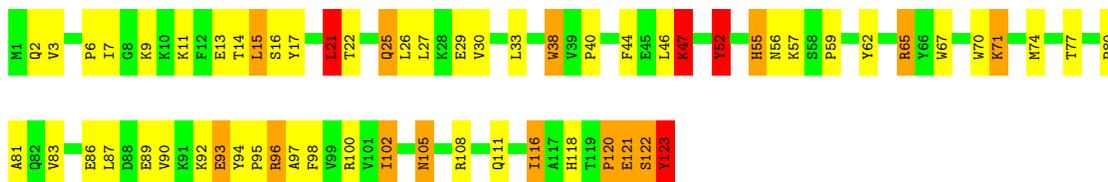
- Molecule 2 is a protein called Ribulose biphosphate carboxylase small chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	S	123	Total 1049	C 694	N 172	O 178	S 5	0	2	0
2	T	123	Total 1056	C 698	N 175	O 178	S 5	0	3	0
2	U	123	Total 1047	C 690	N 172	O 180	S 5	0	2	0
2	V	123	Total 1057	C 697	N 175	O 180	S 5	0	3	0

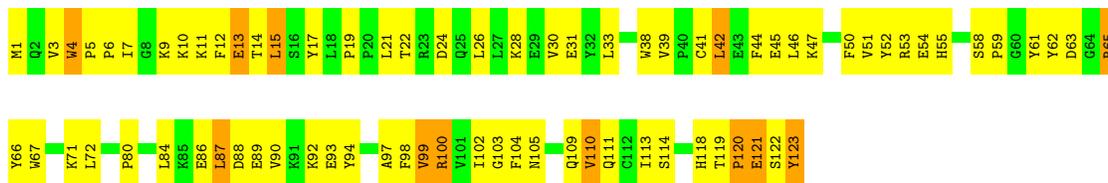
- Molecule 3 is RIBULOSE-1,5-DIPHOSPHATE (three-letter code: RUB) (formula: $C_5H_{12}O_{11}P_2$).



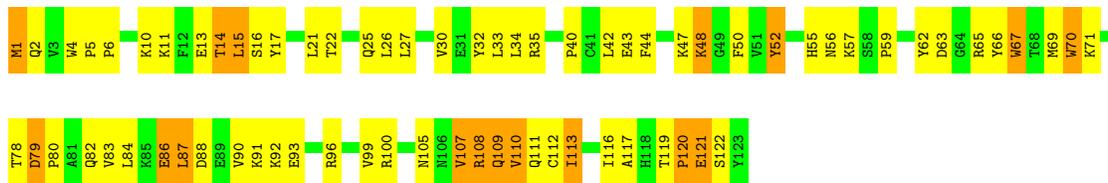
• Molecule 2: Ribulose biphosphate carboxylase small chain



• Molecule 2: Ribulose biphosphate carboxylase small chain

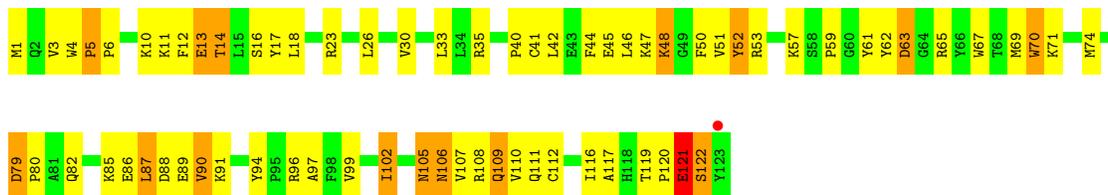


• Molecule 2: Ribulose biphosphate carboxylase small chain



• Molecule 2: Ribulose biphosphate carboxylase small chain





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, α , β , γ	109.79Å 109.95Å 201.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	109.95 – 2.20 48.20 – 2.20	Depositor EDS
% Data completeness (in resolution range)	93.0 (109.95-2.20) 93.1 (48.20-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.22	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.84 (at 2.20Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.197 , 0.279 0.198 , 0.262	Depositor DCC
R_{free} test set	5831 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	10.4	Xtriage
Anisotropy	0.893	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 3.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.37$, $\langle L^2 \rangle = 0.20$	Xtriage
Estimated twinning fraction	0.349 for k,h,-l	Xtriage
Reported twinning fraction	0.580 for H, K, L 0.420 for K, H, -L	Depositor
Outliers	0 of 116024 reflections	Xtriage
F_o, F_c correlation	0.85	EDS
Total number of atoms	19680	wwPDB-VP
Average B, all atoms (Å ²)	10.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.51% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: RUB

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	1.15	9/3765 (0.2%)	1.35	33/5104 (0.6%)
1	B	1.15	4/3783 (0.1%)	1.35	31/5128 (0.6%)
1	C	1.01	6/3790 (0.2%)	1.19	24/5135 (0.5%)
1	D	1.03	8/3758 (0.2%)	1.27	25/5090 (0.5%)
2	S	0.95	3/1089 (0.3%)	1.07	3/1474 (0.2%)
2	T	0.94	1/1099 (0.1%)	1.02	1/1487 (0.1%)
2	U	0.81	3/1087 (0.3%)	0.90	0/1472
2	V	0.84	2/1100 (0.2%)	0.94	3/1489 (0.2%)
All	All	1.04	36/19471 (0.2%)	1.23	120/26379 (0.5%)

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
1	C	0	3
2	S	0	1
2	V	0	1
All	All	0	6

The worst 5 of 36 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	214	TRP	NE1-CE2	-6.74	1.28	1.37
1	A	373	GLY	N-CA	6.64	1.56	1.46
1	C	204	GLU	CD-OE2	6.45	1.32	1.25
1	D	70	TRP	CD2-CE2	6.11	1.48	1.41
1	D	411	TRP	CD2-CE2	6.07	1.48	1.41

The worst 5 of 120 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	131	ARG	NE-CZ-NH1	11.31	125.95	120.30
1	D	367	ASP	CB-CG-OD2	9.96	127.27	118.30
1	D	295	ARG	NE-CZ-NH1	9.59	125.09	120.30
1	A	253	ARG	NE-CZ-NH1	9.30	124.95	120.30
1	A	160	ASP	CB-CG-OD2	8.97	126.38	118.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	199	PHE	Peptide
1	C	195	GLY	Peptide
1	C	199	PHE	Peptide
1	C	93	GLU	Peptide
2	S	47	LYS	Peptide

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	3648	0	3615	263	2
1	B	3653	0	3642	270	1
1	C	3662	0	3632	268	1
1	D	3647	0	3601	248	3
2	S	1049	0	1065	75	1
2	T	1056	0	1076	83	1
2	U	1047	0	1057	75	1
2	V	1057	0	1071	83	0
3	A	18	0	8	4	0
3	B	18	0	8	4	0
3	C	18	0	8	4	0
3	D	18	0	8	1	0
4	A	161	0	0	28	1
4	B	175	0	0	31	1
4	C	137	0	0	29	1
4	D	147	0	0	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	S	49	0	0	8	0
4	T	50	0	0	3	0
4	U	30	0	0	3	0
4	V	40	0	0	6	0
All	All	19680	0	18791	1296	7

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 34.

The worst 5 of 1296 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:C:201:LYS:NZ	1:C:294:HIS:NE2	1.70	1.35
1:C:201:LYS:HE3	1:C:202:ASP:O	1.40	1.16
1:D:201:LYS:NZ	1:D:294:HIS:NE2	1.98	1.11
1:A:60:GLU:HG3	1:A:127:PHE:HZ	1.04	1.09
1:A:409:HIS:HD2	1:A:458:ALA:HB2	1.16	1.08

The worst 5 of 7 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:S:105:ASN:O	4:A:727:HOH:O[2_555]	1.92	0.28
1:D:207:ASN:ND2	4:C:615:HOH:O[2_555]	1.94	0.26
1:D:75:THR:OG1	2:T:109:GLN:NE2[2_555]	2.00	0.20
1:A:106:ASP:OD2	1:A:370[B]:SER:OG[2_555]	2.05	0.15
1:A:65:THR:CG2	4:B:728:HOH:O[2_555]	2.07	0.13

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	467/475 (98%)	405 (87%)	51 (11%)	11 (2%)	6	3
1	B	469/475 (99%)	404 (86%)	48 (10%)	17 (4%)	3	1
1	C	468/475 (98%)	411 (88%)	47 (10%)	10 (2%)	7	4
1	D	465/475 (98%)	408 (88%)	47 (10%)	10 (2%)	6	4
2	S	123/123 (100%)	105 (85%)	12 (10%)	6 (5%)	2	1
2	T	124/123 (101%)	109 (88%)	12 (10%)	3 (2%)	6	3
2	U	123/123 (100%)	97 (79%)	20 (16%)	6 (5%)	2	1
2	V	124/123 (101%)	111 (90%)	9 (7%)	4 (3%)	4	2
All	All	2363/2392 (99%)	2050 (87%)	246 (10%)	67 (3%)	5	2

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	379	SER
1	B	212	MET
1	C	21	LYS
2	S	120	PRO
2	T	120	PRO

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	381/385 (99%)	353 (93%)	28 (7%)	14	15
1	B	383/385 (100%)	348 (91%)	35 (9%)	9	9
1	C	382/385 (99%)	347 (91%)	35 (9%)	9	9
1	D	379/385 (98%)	341 (90%)	38 (10%)	7	7
2	S	115/113 (102%)	101 (88%)	14 (12%)	5	4
2	T	116/113 (103%)	104 (90%)	12 (10%)	7	6
2	U	115/113 (102%)	98 (85%)	17 (15%)	3	2
2	V	116/113 (103%)	103 (89%)	13 (11%)	6	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1987/1992 (100%)	1795 (90%)	192 (10%)	9 7

5 of 192 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	203	ASP
2	S	93	GLU
1	D	262	VAL
1	D	401	GLN
2	T	1	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	96	GLN
1	D	292	HIS
1	D	115	ASN
1	D	205	ASN
1	D	383	HIS

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

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
3	RUB	D	501	-	17,17,17	1.07	0	17,25,25	1.84	5 (29%)
3	RUB	A	501	-	17,17,17	1.22	2 (11%)	17,25,25	2.05	7 (41%)
3	RUB	C	501	-	17,17,17	0.80	0	17,25,25	2.22	9 (52%)
3	RUB	B	501	-	17,17,17	1.61	4 (23%)	17,25,25	2.27	5 (29%)

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	RUB	D	501	-	-	10/20/20/20	-
3	RUB	A	501	-	-	14/20/20/20	-
3	RUB	C	501	-	-	7/20/20/20	-
3	RUB	B	501	-	-	16/20/20/20	-

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	501	RUB	O2-C2	3.65	1.27	1.21
3	B	501	RUB	P2-O5	2.75	1.69	1.60
3	A	501	RUB	O3-C3	2.67	1.47	1.42
3	B	501	RUB	O3-C3	2.47	1.47	1.42
3	A	501	RUB	C5-C4	2.42	1.55	1.51

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	501	RUB	O2P-P1-O1P	4.72	129.23	110.83
3	B	501	RUB	O5P-P2-O5	-4.01	96.20	106.67
3	D	501	RUB	O5-P2-O4P	-3.87	95.98	106.44
3	B	501	RUB	O6P-P2-O5	3.65	116.18	106.67
3	C	501	RUB	O5P-P2-O5	-3.53	97.46	106.67

There are no chirality outliers.

5 of 47 torsion outliers are listed below:

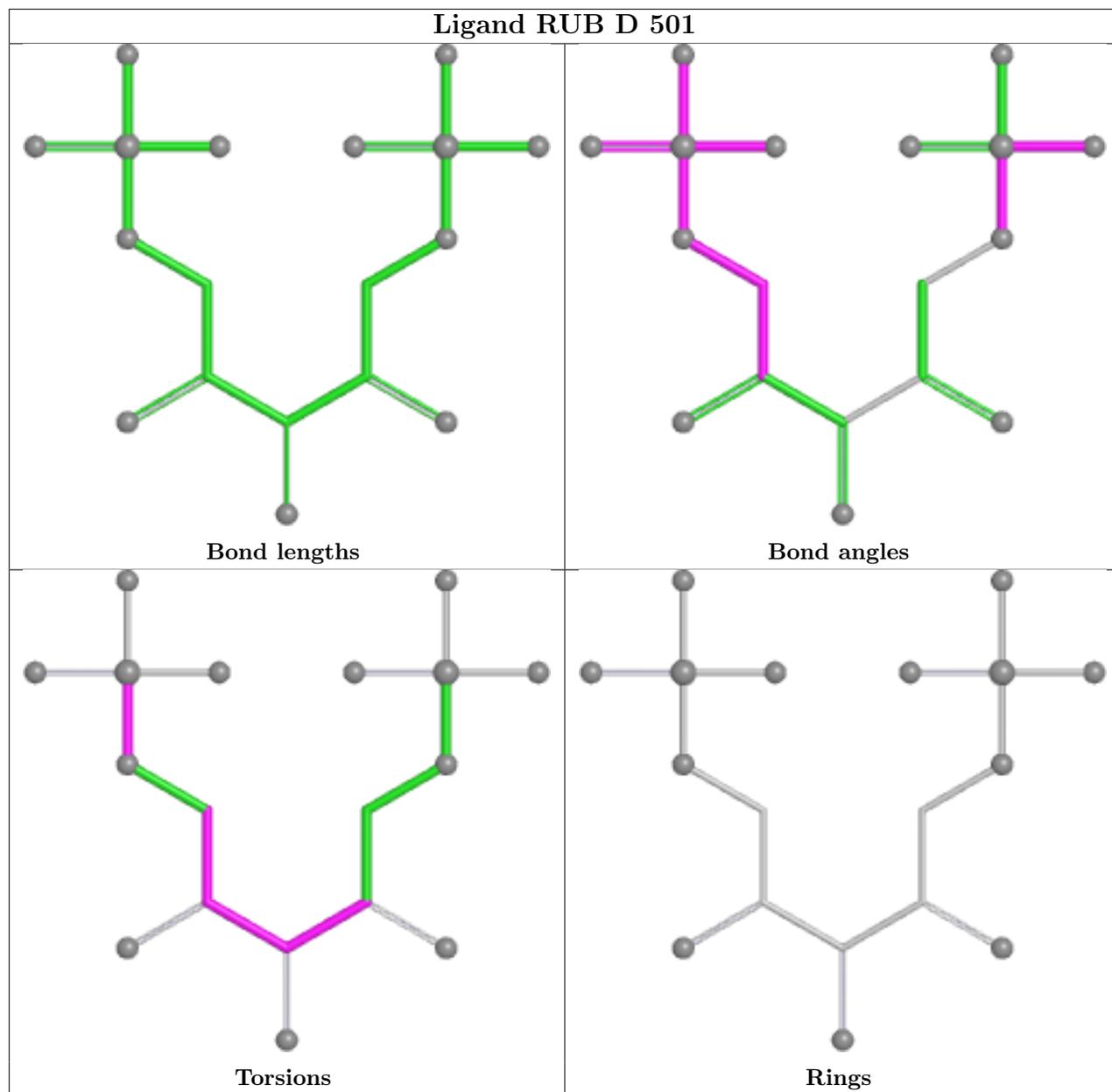
Mol	Chain	Res	Type	Atoms
3	A	501	RUB	O1-C1-C2-C3
3	A	501	RUB	O1-C1-C2-O2
3	A	501	RUB	O2-C2-C3-O3
3	A	501	RUB	C2-C3-C4-C5
3	A	501	RUB	C2-C3-C4-O4

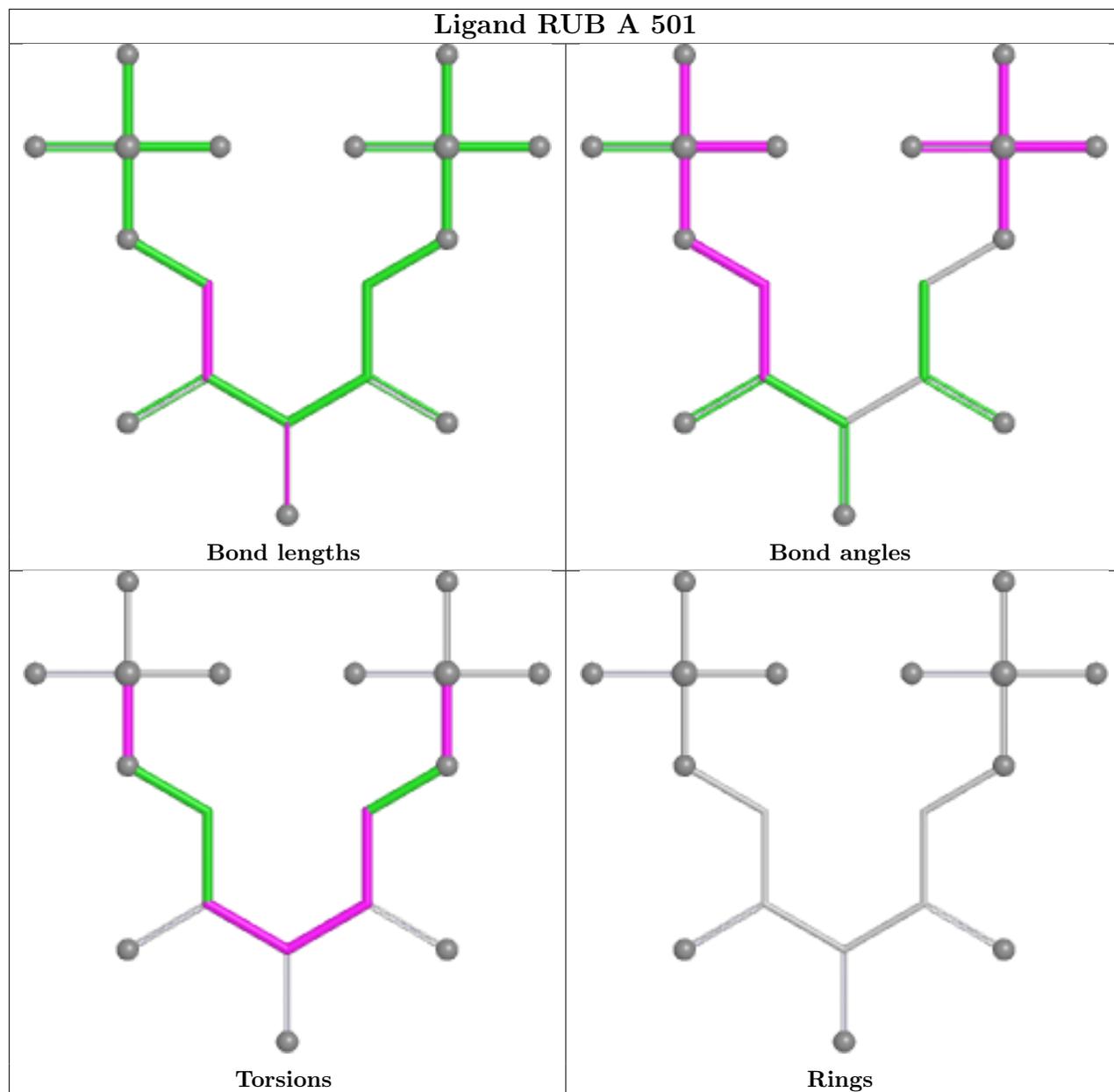
There are no ring outliers.

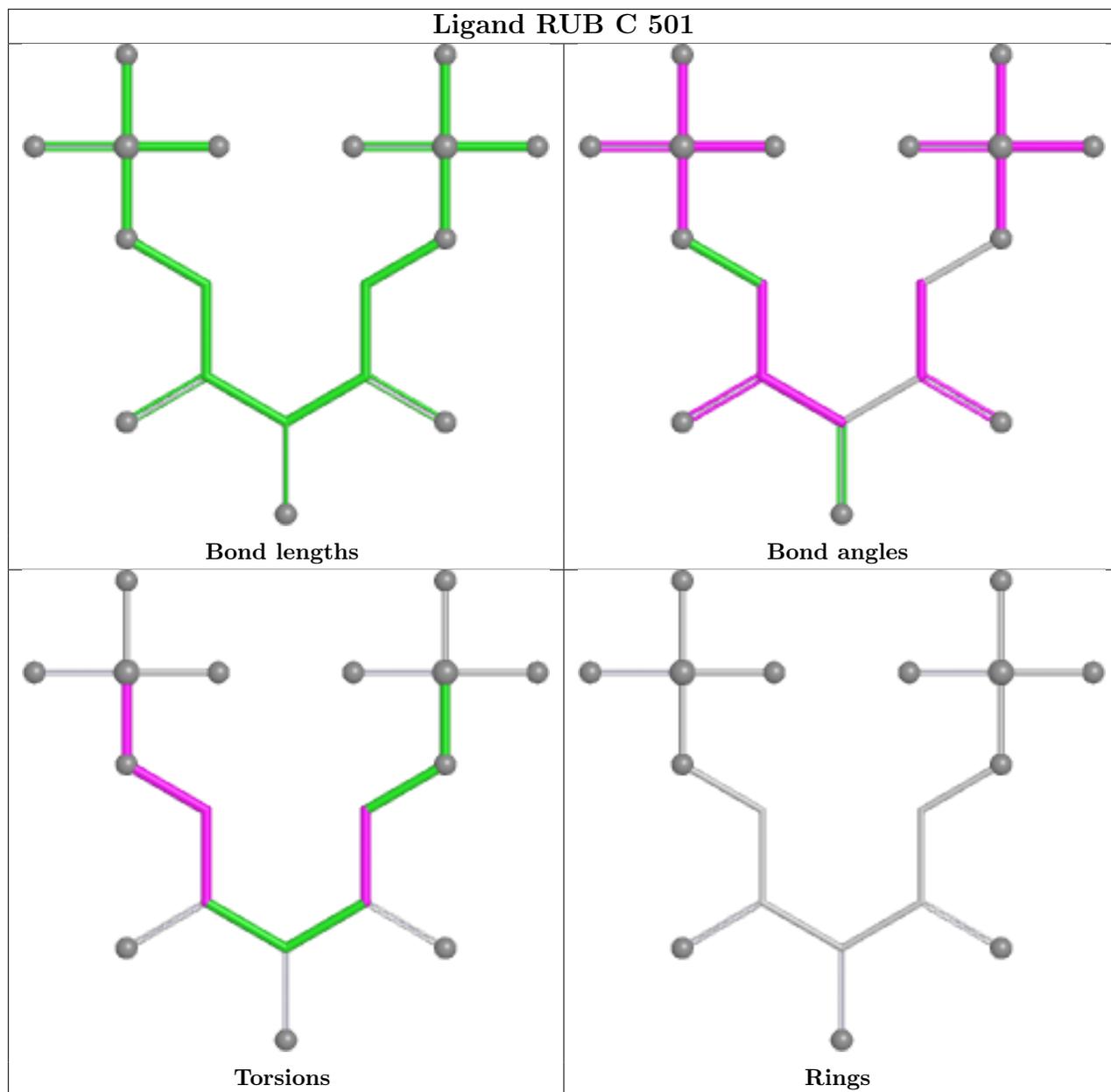
4 monomers are involved in 13 short contacts:

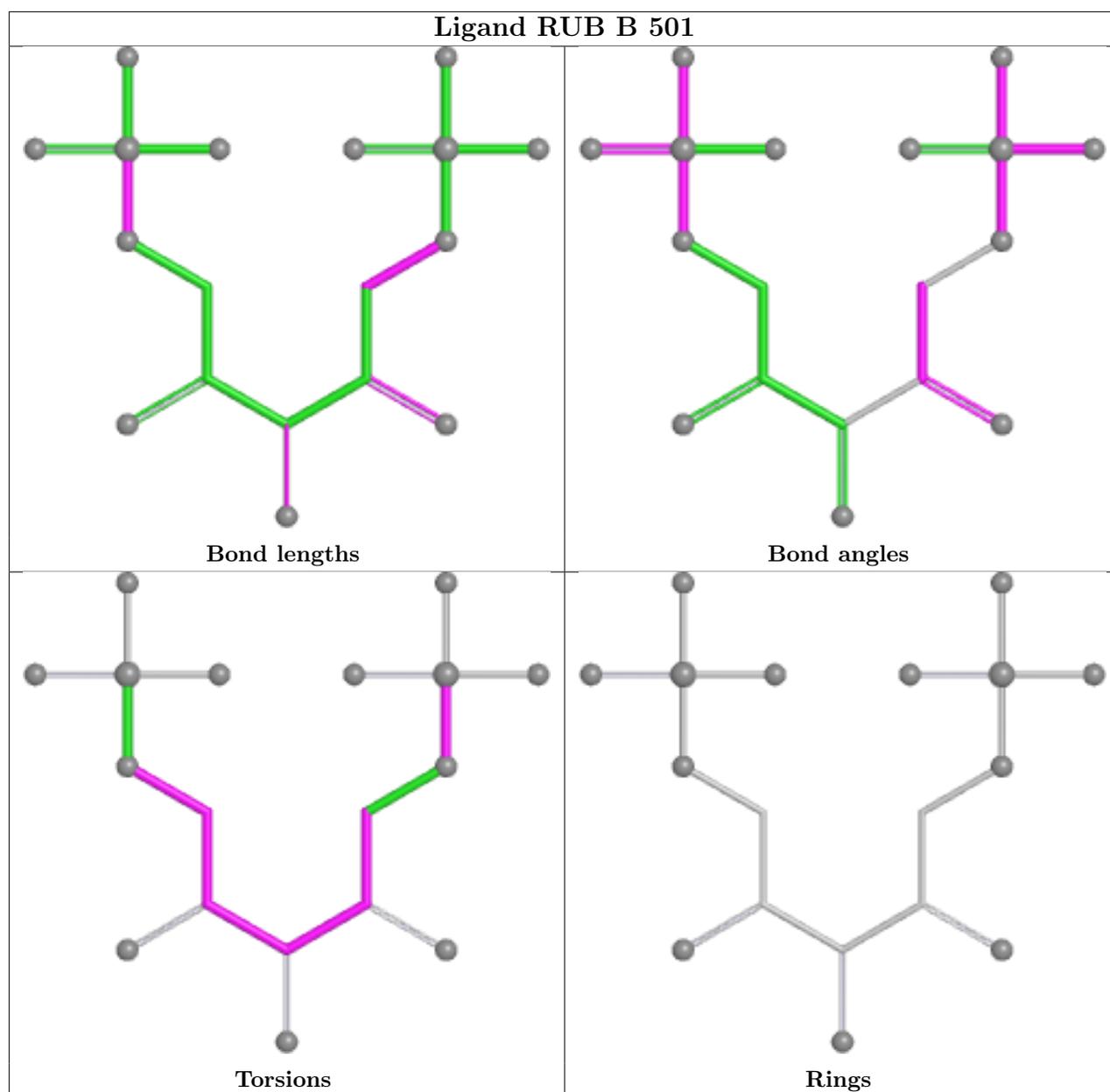
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	501	RUB	1	0
3	A	501	RUB	4	0
3	C	501	RUB	4	0
3	B	501	RUB	4	0

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 than 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 [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	458/475 (96%)	-0.96	0 100 100	2, 2, 20, 43	0
1	B	458/475 (96%)	-0.97	0 100 100	2, 2, 23, 46	0
1	C	458/475 (96%)	-0.86	0 100 100	2, 9, 28, 34	0
1	D	458/475 (96%)	-0.89	0 100 100	2, 10, 30, 37	0
2	S	123/123 (100%)	-0.80	0 100 100	2, 15, 24, 34	0
2	T	123/123 (100%)	-0.80	0 100 100	2, 14, 27, 52	0
2	U	123/123 (100%)	-0.60	0 100 100	7, 23, 37, 46	0
2	V	123/123 (100%)	-0.62	1 (0%) 86 85	5, 19, 31, 58	0
All	All	2324/2392 (97%)	-0.88	1 (0%) 100 100	2, 8, 29, 58	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	V	123	TYR	3.2

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](#)

There are no monosaccharides in this entry.

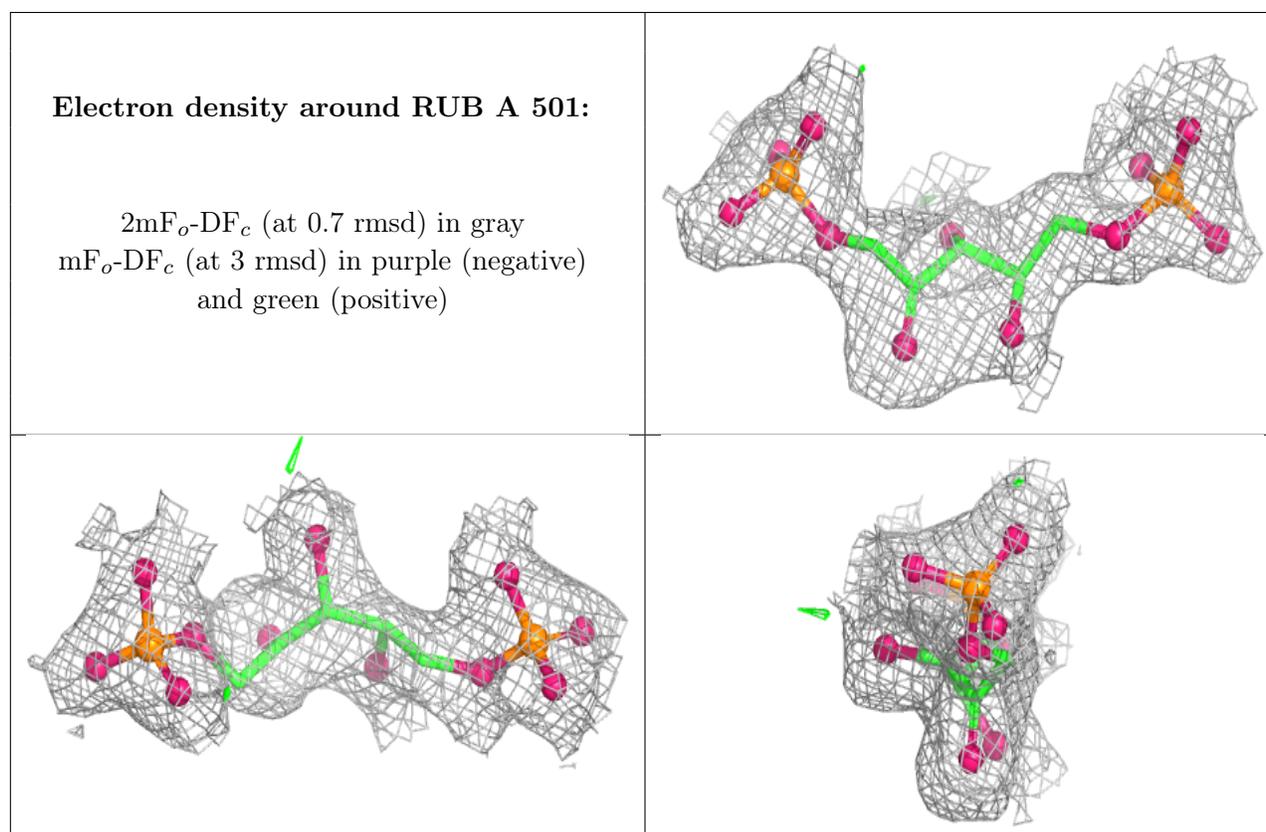
6.4 Ligands [i](#)

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.

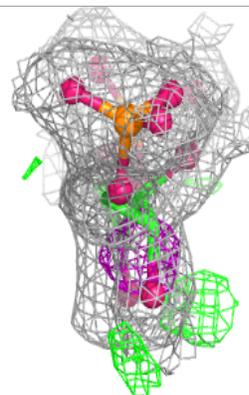
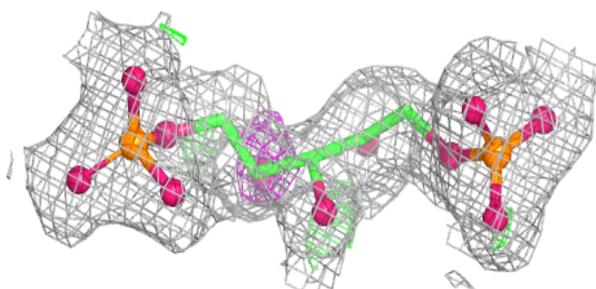
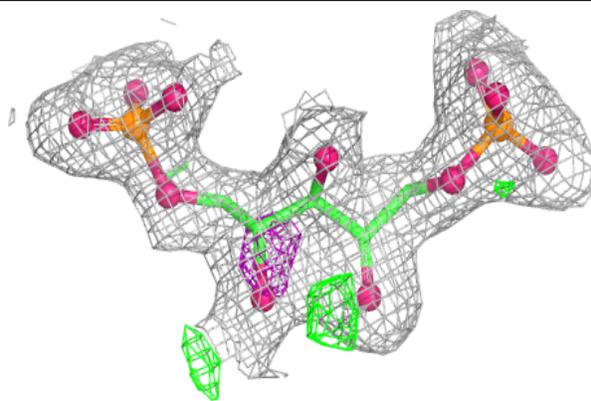
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	RUB	A	501	18/18	0.97	0.08	6,10,12,12	0
3	RUB	B	501	18/18	0.98	0.09	2,2,2,2	0
3	RUB	D	501	18/18	0.98	0.07	8,14,18,19	0
3	RUB	C	501	18/18	0.99	0.07	8,11,16,17	0

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.

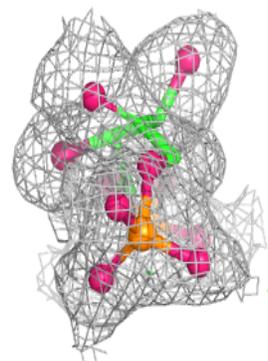
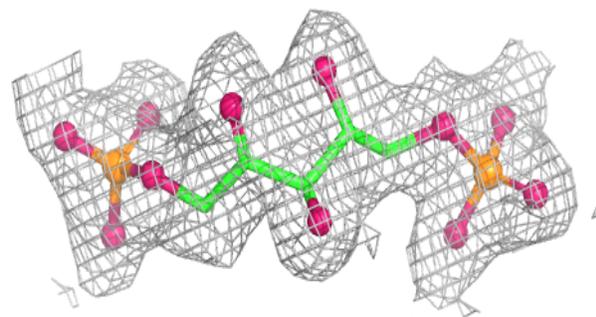
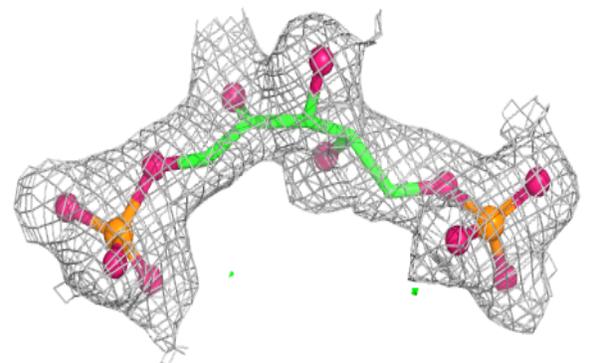


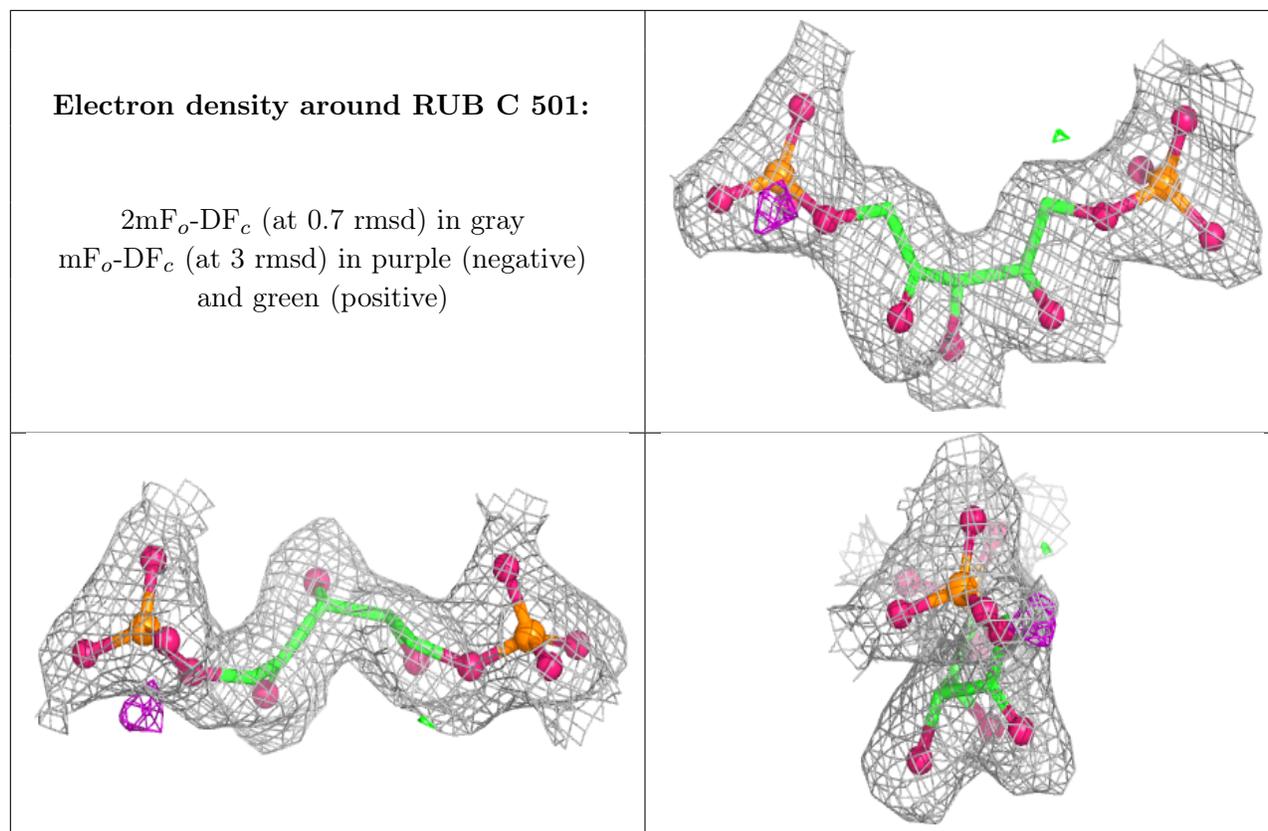
Electron density around RUB B 501:

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

**Electron density around RUB D 501:**

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





6.5 Other polymers [i](#)

There are no such residues in this entry.