



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 9, 2021 – 12:07 pm BST

PDB ID : 6ZYT
Title : Monomeric streptavidin with a conjugated biotinylated pyrrolidine
Authors : Nodling, A.R.; Lipka-Lloyd, M.; Tsai, Y.H.; Rizkallah, P.; Luk, L.Y.P.; Jin, Y.
Deposited on : 2020-08-03
Resolution : 1.80 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 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 : 2.23.1
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) : 2.23.1

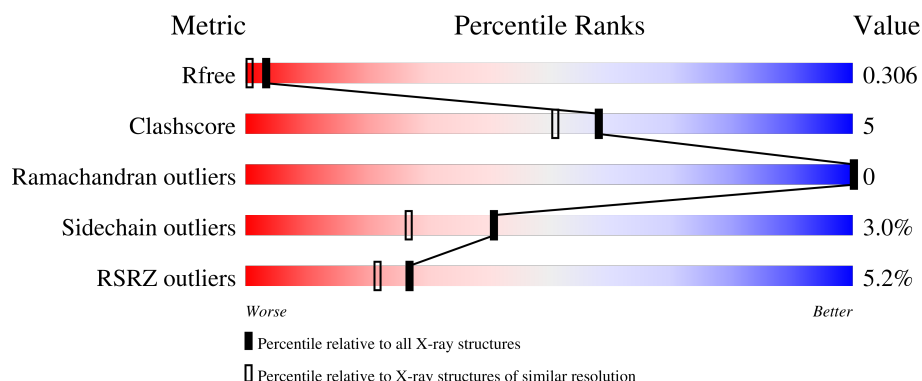
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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.



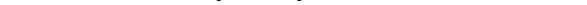
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	AAA	145	<div> <div>0%</div> <div>68% 7% 24%</div> </div>
1	BBB	145	<div> <div>3%</div> <div>68% 12% 21%</div> </div>
1	CCC	145	<div> <div>0%</div> <div>62% 14% 23%</div> </div>
1	DDD	145	<div> <div>0%</div> <div>72% 5% 23%</div> </div>
1	EEE	145	<div> <div>12%</div> <div>61% 14% 25%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	FFF	145	 <p>7% 66% 10% 23%</p>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5622 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 Streptavidin/Rhizavidin Hybrid.

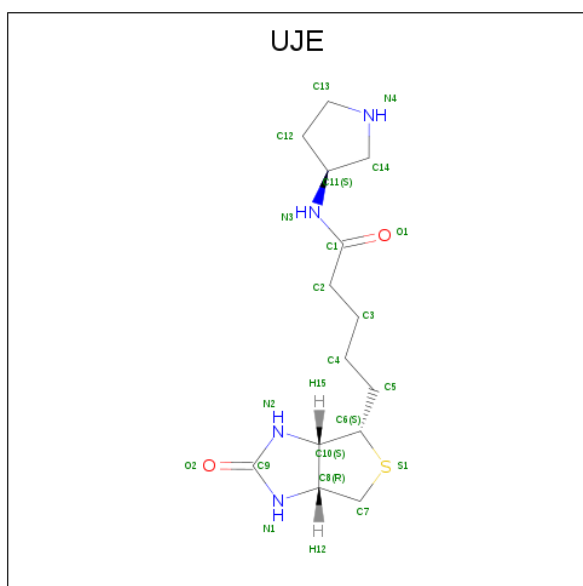
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	BBB	115	Total	C	N	O	S	0	2	0
			907	549	167	189	2			
1	AAA	110	Total	C	N	O	S	0	0	0
			860	523	157	178	2			
1	CCC	111	Total	C	N	O	S	0	0	0
			864	525	158	179	2			
1	FFF	111	Total	C	N	O	S	0	0	0
			864	525	158	179	2			
1	DDD	112	Total	C	N	O	S	0	1	0
			878	534	161	181	2			
1	EEE	109	Total	C	N	O	S	0	1	0
			858	521	158	177	2			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	BBB	1	Total	O	S	0	0
			5	4	1		
2	BBB	1	Total	O	S	0	0
			5	4	1		
2	AAA	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is 5-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-N-((S)-pyrrolidin-3-yl)pentanamide (three-letter code: UJE) (formula: C₁₄H₂₄N₄O₂S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	BBB	1	Total	C	N	O	S	0	0
			21	14	4	2	1		
3	AAA	1	Total	C	N	O	S	0	0
			21	14	4	2	1		
3	CCC	1	Total	C	N	O	S	0	0
			21	14	4	2	1		
3	FFF	1	Total	C	N	O	S	0	0
			21	14	4	2	1		
3	DDD	1	Total	C	N	O	S	0	0
			21	14	4	2	1		
3	EEE	1	Total	C	N	O	S	0	0
			21	14	4	2	1		

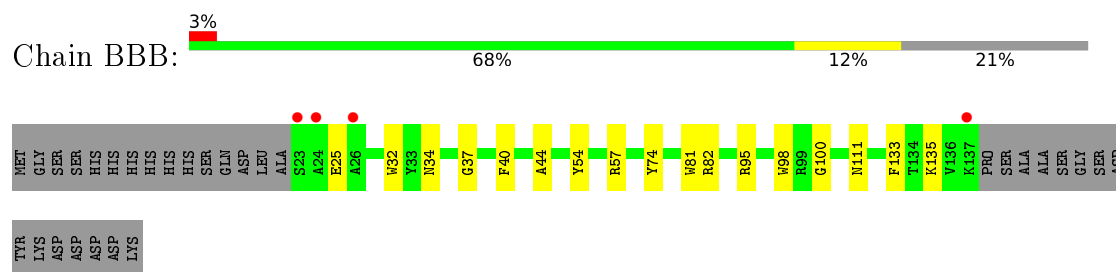
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	BBB	43	Total 43	O 43	0	0
4	AAA	67	Total 67	O 67	0	0
4	CCC	49	Total 49	O 49	0	0
4	FFF	22	Total 22	O 22	0	0
4	DDD	48	Total 48	O 48	0	0
4	EEE	21	Total 21	O 21	0	0

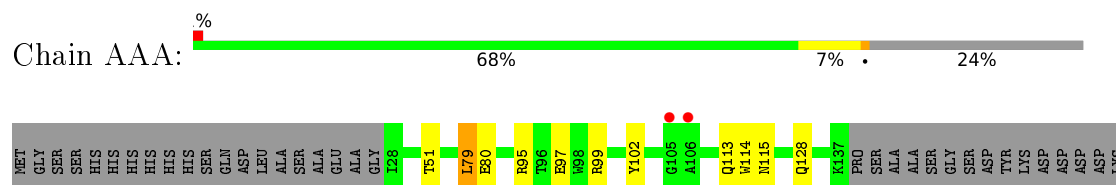
3 Residue-property plots [i](#)

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.

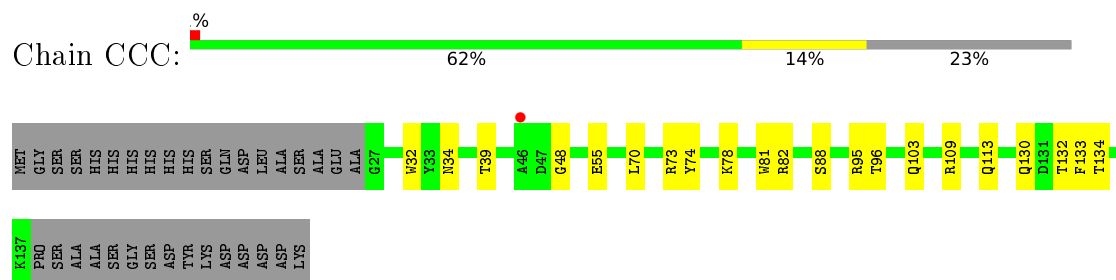
- Molecule 1: Streptavidin/Rhizavidin Hybrid



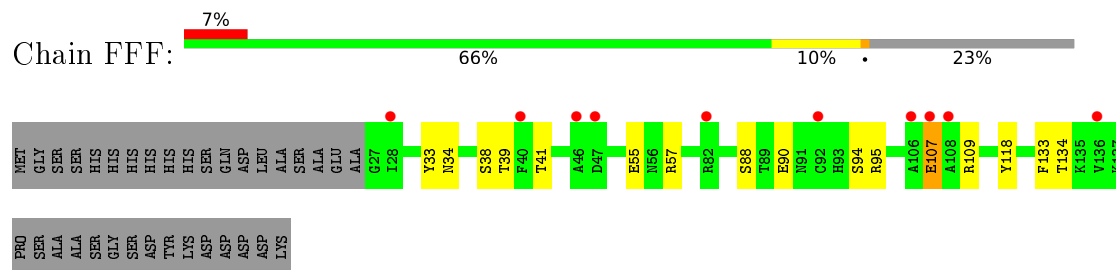
- Molecule 1: Streptavidin/Rhizavidin Hybrid



- Molecule 1: Streptavidin/Rhizavidin Hybrid



- Molecule 1: Streptavidin/Rhizavidin Hybrid



- Molecule 1: Streptavidin/Rhizavidin Hybrid

Chain DDD:

72%

5%

23%



● Molecule 1: Streptavidin/Rhizavidin Hybrid

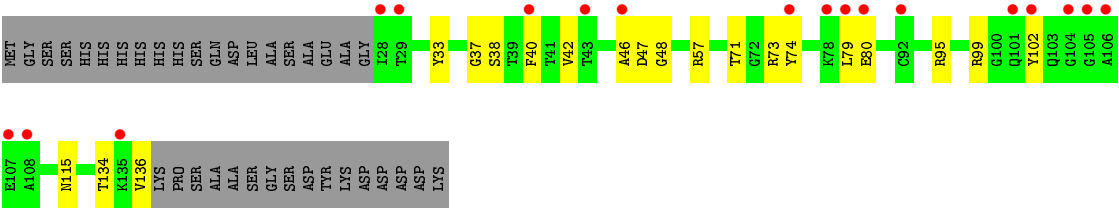
Chain EEE:

12%

61%

14%

25%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	41.79 Å 52.74 Å 89.36 Å 75.24° 78.24° 84.88°	Depositor
Resolution (Å)	25.48 – 1.80 25.48 – 1.80	Depositor EDS
% Data completeness (in resolution range)	95.7 (25.48-1.80) 95.8 (25.48-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.83 (at 1.80 Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.259 , 0.301 0.262 , 0.306	Depositor DCC
R_{free} test set	3144 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	27.9	Xtriage
Anisotropy	0.609	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 26.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5622	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.92% 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: SO4, UJE

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	AAA	0.63	0/880	0.81	0/1196
1	BBB	0.64	0/927	0.84	0/1258
1	CCC	0.64	0/884	0.81	0/1201
1	DDD	0.64	0/898	0.78	0/1219
1	EEE	0.66	0/878	0.78	0/1194
1	FFF	0.65	0/884	0.79	0/1201
All	All	0.64	0/5351	0.80	0/7269

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	AAA	860	0	775	6	0
1	BBB	907	0	812	9	0
1	CCC	864	0	778	11	0
1	DDD	878	0	795	3	0
1	EEE	858	0	768	12	0
1	FFF	864	0	778	11	0
2	AAA	5	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	BBB	10	0	0	0	0
3	AAA	21	0	0	0	0
3	BBB	21	0	0	0	0
3	CCC	21	0	0	0	0
3	DDD	21	0	0	0	0
3	EEE	21	0	0	0	0
3	FFF	21	0	0	0	0
4	AAA	67	0	0	0	0
4	BBB	43	0	0	1	0
4	CCC	49	0	0	0	0
4	DDD	48	0	0	0	0
4	EEE	21	0	0	0	0
4	FFF	22	0	0	1	0
All	All	5622	0	4706	48	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 48 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:AAA:51:THR:HG23	1:CCC:130:GLN:OE1	1.87	0.74
1:EEE:40:PHE:CE1	1:EEE:42:VAL:HG11	2.24	0.72
1:CCC:109:ARG:HD3	1:CCC:134:THR:HG22	1.77	0.66
1:FFF:57:ARG:HE	1:EEE:71:THR:HG21	1.63	0.64
1:FFF:33:TYR:CD2	1:FFF:39:THR:HG22	2.35	0.61

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	AAA	108/145 (74%)	107 (99%)	1 (1%)	0	100	100
1	BBB	115/145 (79%)	112 (97%)	3 (3%)	0	100	100
1	CCC	109/145 (75%)	107 (98%)	2 (2%)	0	100	100
1	DDD	111/145 (77%)	108 (97%)	3 (3%)	0	100	100
1	EEE	108/145 (74%)	101 (94%)	7 (6%)	0	100	100
1	FFF	109/145 (75%)	105 (96%)	4 (4%)	0	100	100
All	All	660/870 (76%)	640 (97%)	20 (3%)	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	AAA	88/115 (76%)	85 (97%)	3 (3%)	37	22
1	BBB	91/115 (79%)	90 (99%)	1 (1%)	73	68
1	CCC	88/115 (76%)	85 (97%)	3 (3%)	37	22
1	DDD	89/115 (77%)	85 (96%)	4 (4%)	27	13
1	EEE	87/115 (76%)	84 (97%)	3 (3%)	37	22
1	FFF	88/115 (76%)	84 (96%)	4 (4%)	27	13
All	All	531/690 (77%)	513 (97%)	18 (3%)	41	22

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	95	ARG
1	EEE	115	ASN
1	EEE	95[B]	ARG
1	FFF	88	SER
1	DDD	79	LEU

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

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 ⓘ

9 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	UJE	CCC	201	-	22,23,23	1.28	1 (4%)	28,31,31	2.23	6 (21%)
2	SO4	AAA	201	-	4,4,4	0.41	0	6,6,6	0.11	0
3	UJE	EEE	201	-	22,23,23	0.96	1 (4%)	28,31,31	1.13	4 (14%)
3	UJE	BBB	203	-	22,23,23	1.09	1 (4%)	28,31,31	2.38	11 (39%)
3	UJE	AAA	202	-	22,23,23	1.16	2 (9%)	28,31,31	2.70	10 (35%)
2	SO4	BBB	201	-	4,4,4	0.41	0	6,6,6	0.04	0
3	UJE	FFF	201	-	22,23,23	1.09	1 (4%)	28,31,31	1.69	7 (25%)
3	UJE	DDD	201	-	22,23,23	1.08	1 (4%)	28,31,31	2.10	6 (21%)
2	SO4	BBB	202	-	4,4,4	0.39	0	6,6,6	0.11	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
3	UJE	CCC	201	-	-	1/11/39/39	0/3/3/3
3	UJE	BBB	203	-	-	2/11/39/39	0/3/3/3
3	UJE	AAA	202	-	-	3/11/39/39	0/3/3/3
3	UJE	FFF	201	-	-	3/11/39/39	0/3/3/3
3	UJE	DDD	201	-	-	2/11/39/39	0/3/3/3
3	UJE	EEE	201	-	-	2/11/39/39	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	CCC	201	UJE	C6-S1	-5.12	1.74	1.82
3	FFF	201	UJE	C6-S1	-4.25	1.75	1.82
3	DDD	201	UJE	C6-S1	-4.02	1.76	1.82
3	BBB	203	UJE	C6-S1	-3.82	1.76	1.82
3	EEE	201	UJE	C6-S1	-3.68	1.76	1.82

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	AAA	202	UJE	C2-C1-N3	6.93	127.84	115.83
3	BBB	203	UJE	C11-N3-C1	6.38	132.33	122.92
3	DDD	201	UJE	C11-N3-C1	5.88	131.59	122.92
3	AAA	202	UJE	C11-N3-C1	5.69	131.31	122.92
3	CCC	201	UJE	C8-C7-S1	-5.59	101.51	106.31

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

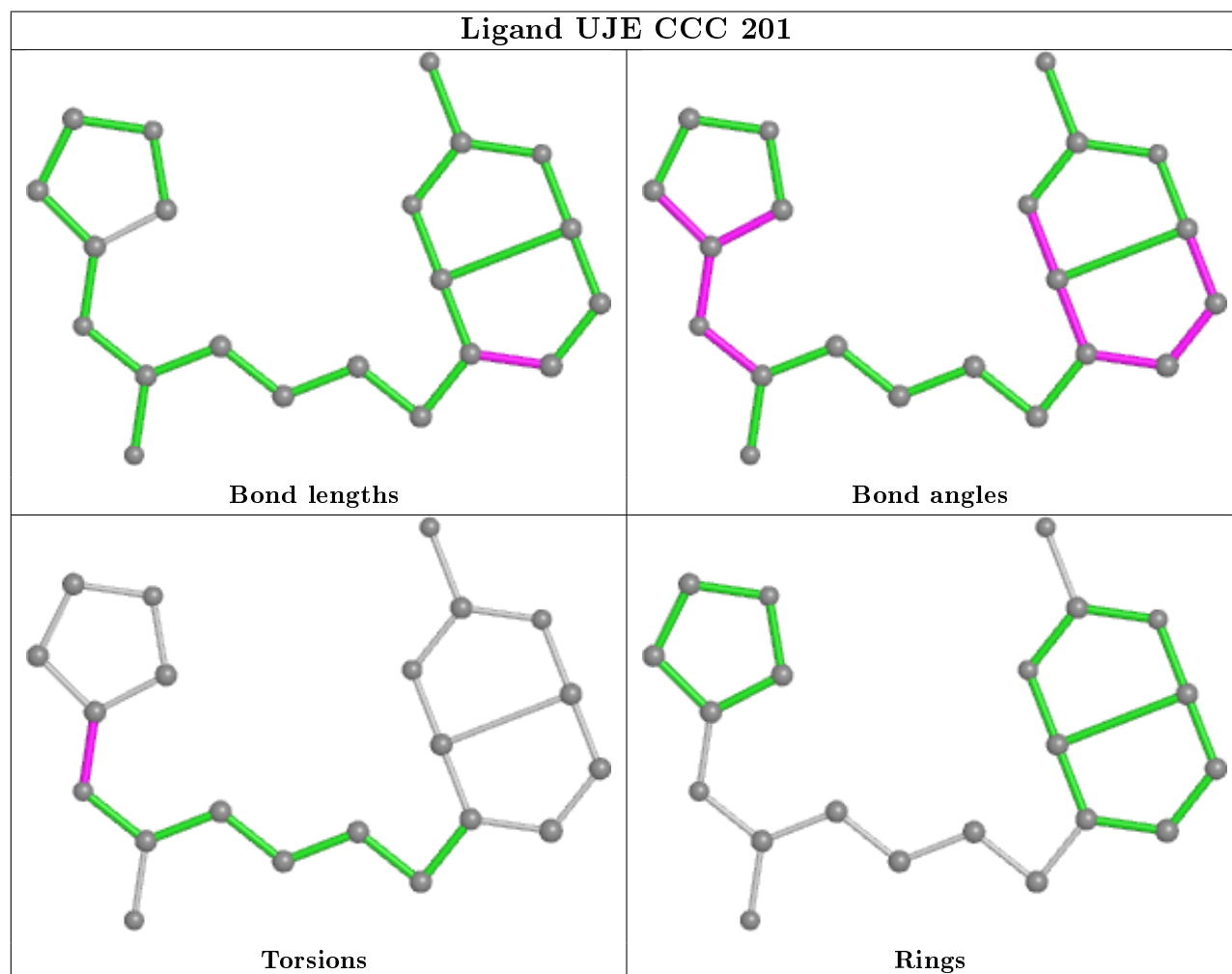
Mol	Chain	Res	Type	Atoms
3	BBB	203	UJE	C12-C11-N3-C1
3	CCC	201	UJE	C12-C11-N3-C1
3	FFF	201	UJE	C4-C5-C6-S1
3	FFF	201	UJE	C4-C5-C6-C10
3	FFF	201	UJE	C14-C11-N3-C1

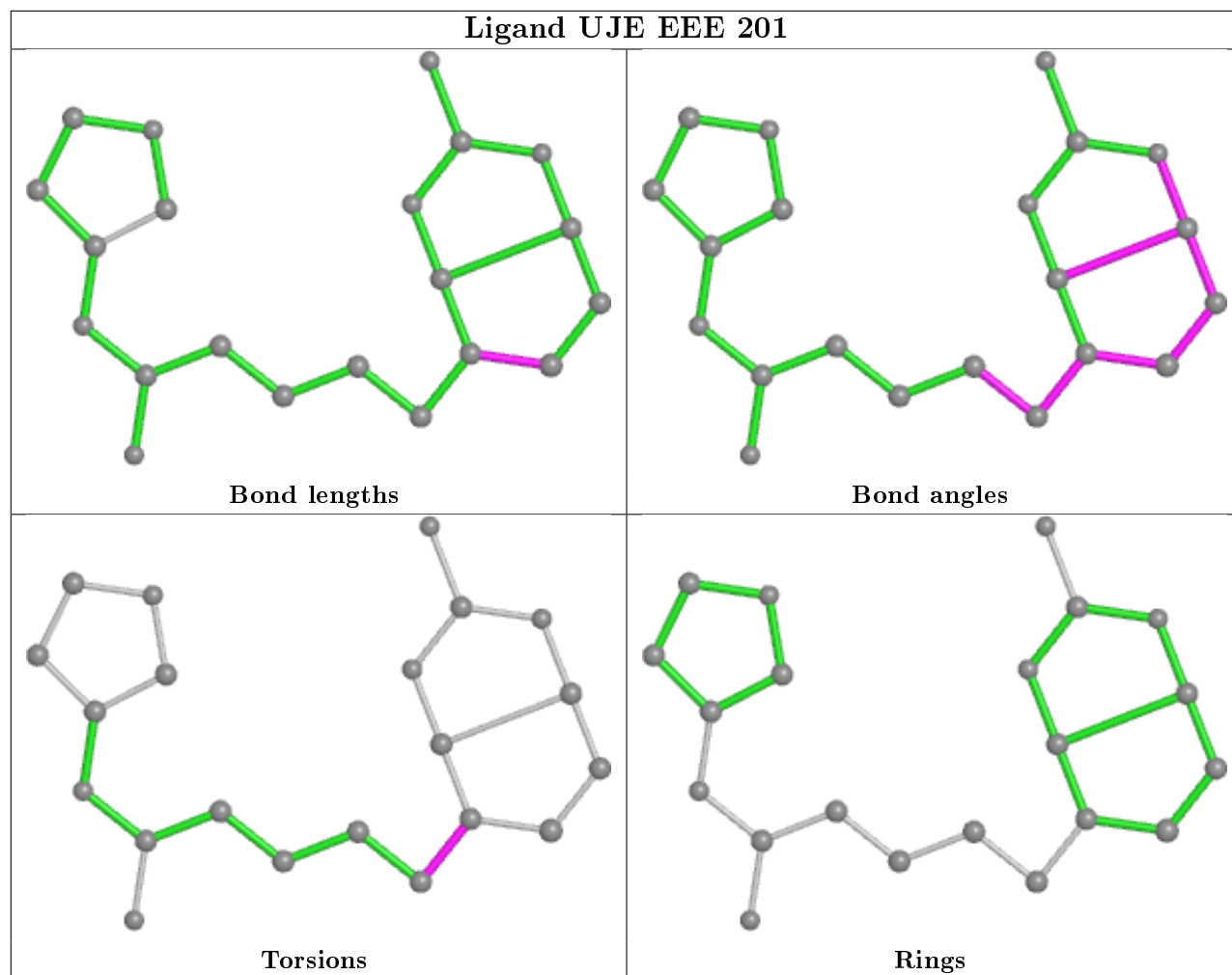
There are no ring outliers.

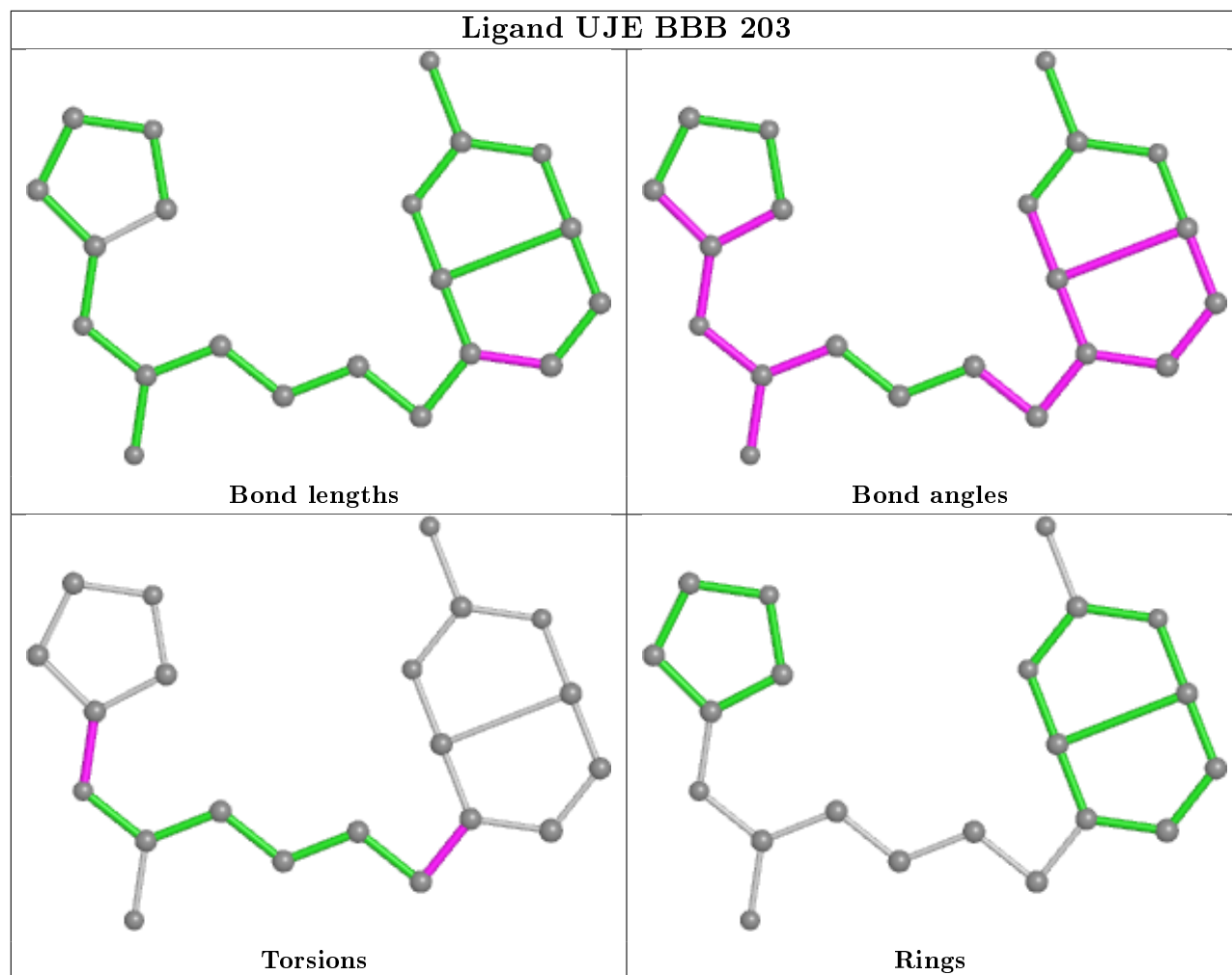
No monomer is involved in short contacts.

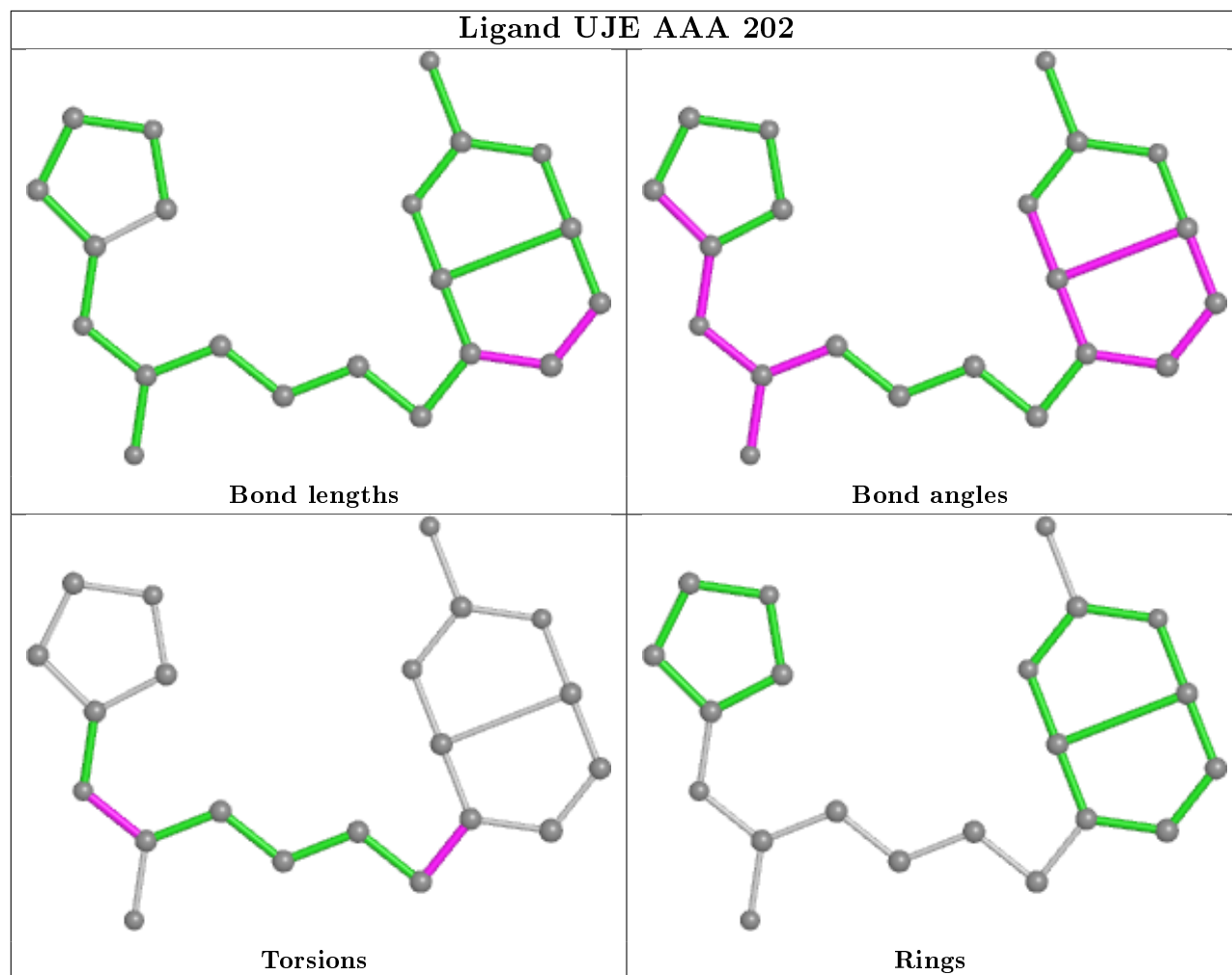
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is

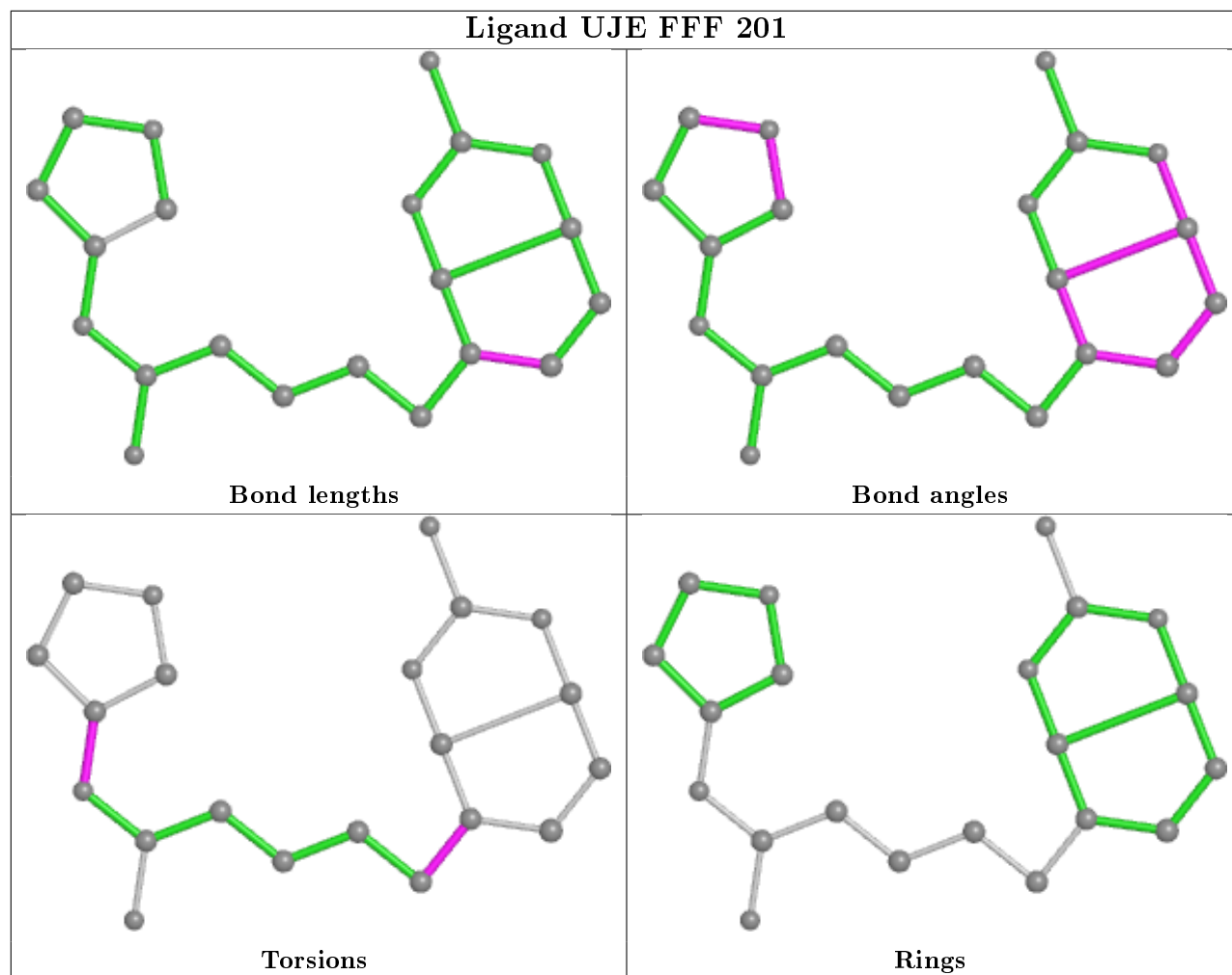
within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question 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.

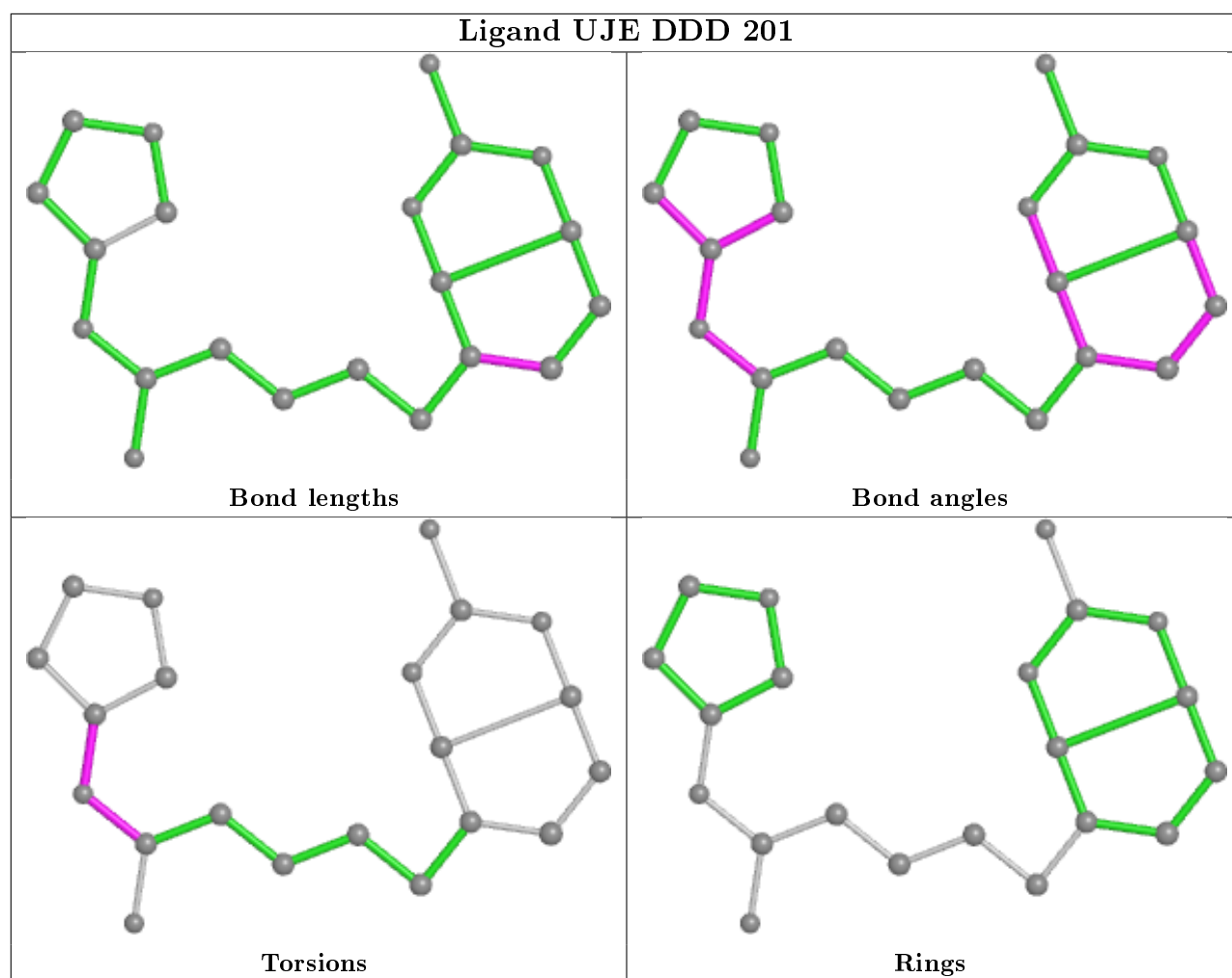












5.7 Other polymers [i](#)

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 ⓘ

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	AAA	110/145 (75%)	0.01	2 (1%) 68 64	22, 31, 45, 67	0
1	BBB	115/145 (79%)	0.13	4 (3%) 44 38	22, 31, 47, 58	0
1	CCC	111/145 (76%)	-0.07	1 (0%) 84 82	22, 32, 46, 64	0
1	DDD	112/145 (77%)	-0.07	0 100 100	21, 31, 40, 53	0
1	EEE	109/145 (75%)	0.93	18 (16%) 1 1	31, 49, 79, 87	0
1	FFF	111/145 (76%)	0.77	10 (9%) 9 7	33, 49, 69, 86	0
All	All	668/870 (76%)	0.28	35 (5%) 27 22	21, 35, 65, 87	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	EEE	106	ALA	7.2
1	AAA	106	ALA	6.7
1	FFF	46	ALA	6.3
1	BBB	26	ALA	5.3
1	FFF	136	VAL	4.9

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 [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(\AA^2)	Q<0.9
3	UJE	FFF	201	21/21	0.90	0.11	32,36,40,44	0
3	UJE	EEE	201	21/21	0.91	0.11	32,35,45,47	0
3	UJE	CCC	201	21/21	0.95	0.09	25,27,34,35	0
3	UJE	BBB	203	21/21	0.95	0.10	28,30,41,42	0
3	UJE	DDD	201	21/21	0.95	0.10	21,23,30,32	0
3	UJE	AAA	202	21/21	0.95	0.10	27,29,36,37	0
2	SO4	BBB	201	5/5	0.96	0.11	58,59,62,63	0
2	SO4	AAA	201	5/5	0.97	0.06	37,38,39,44	0
2	SO4	BBB	202	5/5	0.99	0.06	39,39,40,40	0

6.5 Other polymers [i](#)

There are no such residues in this entry.