



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2024 – 05:21 pm BST

PDB ID : 9F23  
Title : DARPIn eGFP complex DP2 (2G156)  
Authors : Mittl, P.R.; Hansen, S.  
Deposited on : 2024-04-22  
Resolution : 1.59 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](mailto: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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

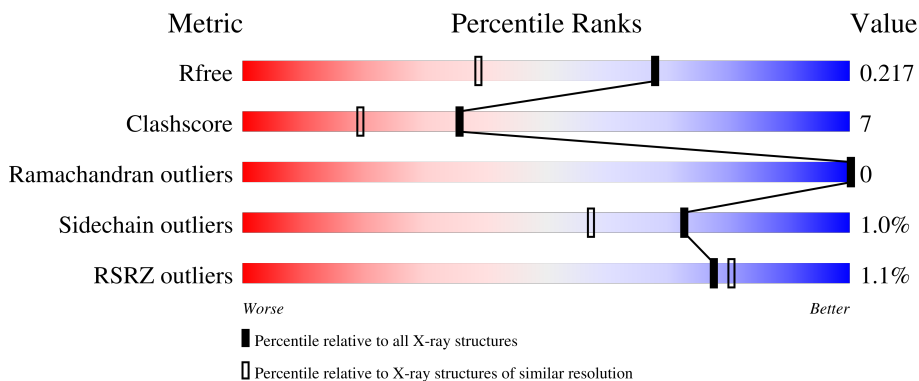
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.59 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))
$R_{free}$	164625	4274 (1.60-1.60)
Clashscore	180529	4682 (1.60-1.60)
Ramachandran outliers	177936	4583 (1.60-1.60)
Sidechain outliers	177891	4582 (1.60-1.60)
RSRZ outliers	164620	4272 (1.60-1.60)

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  $\geq 3$ , 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	243	
1	C	243	
2	B	128	
2	D	128	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
4	EDO	B	203	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 6867 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 Green fluorescent protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	227	1964	1246	330	382	6	0	18	0
1	C	228	2009	1272	336	395	6	0	22	0

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P42212
A	-3	PRO	-	expression tag	UNP P42212
A	-2	GLY	-	expression tag	UNP P42212
A	-1	SER	-	expression tag	UNP P42212
A	0	MET	-	expression tag	UNP P42212
A	1	VAL	-	expression tag	UNP P42212
A	64	LEU	PHE	engineered mutation	UNP P42212
A	66	CRO	SER	chromophore	UNP P42212
A	66	CRO	TYR	chromophore	UNP P42212
A	66	CRO	GLY	chromophore	UNP P42212
A	231	LEU	HIS	engineered mutation	UNP P42212
A	239	GLN	-	expression tag	UNP P42212
A	240	ALA	-	expression tag	UNP P42212
C	-4	GLY	-	expression tag	UNP P42212
C	-3	PRO	-	expression tag	UNP P42212
C	-2	GLY	-	expression tag	UNP P42212
C	-1	SER	-	expression tag	UNP P42212
C	0	MET	-	expression tag	UNP P42212
C	1	VAL	-	expression tag	UNP P42212
C	64	LEU	PHE	engineered mutation	UNP P42212
C	66	CRO	SER	chromophore	UNP P42212
C	66	CRO	TYR	chromophore	UNP P42212
C	66	CRO	GLY	chromophore	UNP P42212
C	231	LEU	HIS	engineered mutation	UNP P42212
C	239	GLN	-	expression tag	UNP P42212

*Continued on next page...*

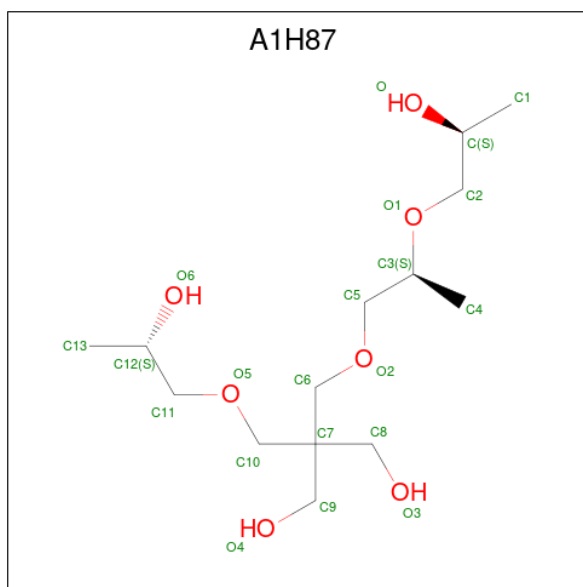
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	240	ALA	-	expression tag	UNP P42212

- Molecule 2 is a protein called DARPin DP2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
2	B	125	Total	C	N	O	S	0	7	0
			990	622	169	195	4			
2	D	124	Total	C	N	O	S	0	7	0
			983	618	169	192	4			

- Molecule 3 is 2-[[[(2S)-2-oxidanylpropoxy]methyl]-2-[[[(2S)-2-[(2S)-2-oxidanylpropoxy]propoxy]methyl]propane-1,3-diol (three-letter code: A1H87) (formula: C<sub>14</sub>H<sub>30</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			21	14	7		
3	C	1	Total	C	O	0	0
			21	14	7		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	C	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0
4	D	1	Total C O 4 2 2	0	0

*Continued on next page...*

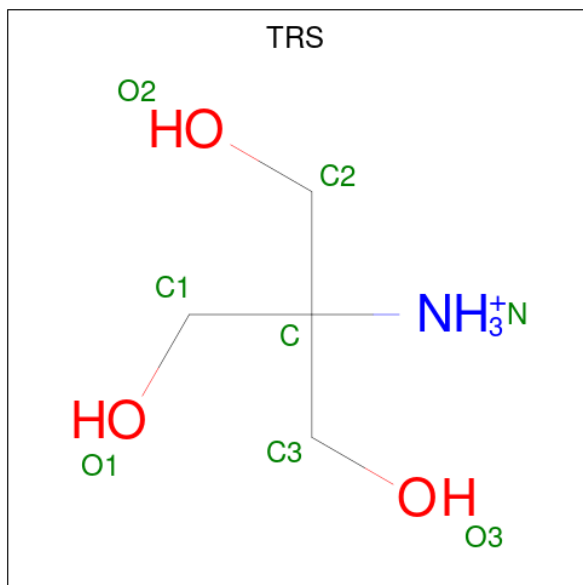
Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

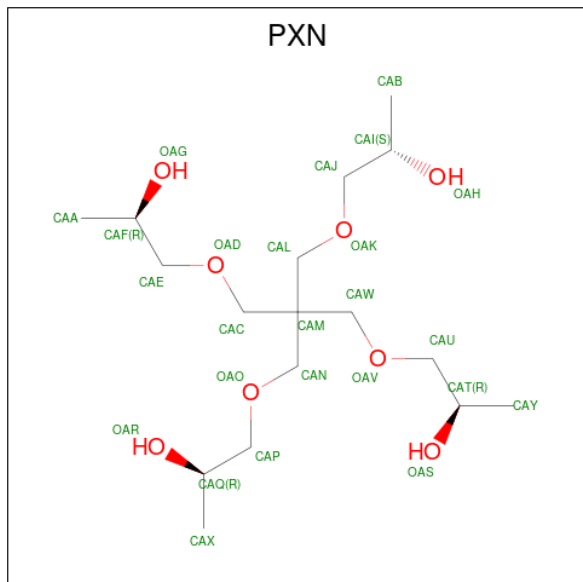
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Na	0	0
			1	1		
5	C	1	Total	Na	0	0
			1	1		

- Molecule 6 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	N	O	0	0
			8	4	1	3		
6	B	1	Total	C	N	O	0	0
			8	4	1	3		
6	B	1	Total	C	N	O	0	0
			8	4	1	3		
6	D	1	Total	C	N	O	0	0
			8	4	1	3		
6	D	1	Total	C	N	O	0	0
			8	4	1	3		
6	D	1	Total	C	N	O	0	0
			8	4	1	3		

- Molecule 7 is (2S)-1-[3-{{(2R)-2-hydroxypropyl}oxy}-2,2-bis({{(2R)-2-hydroxypropyl}oxy}methyl)propoxy]propan-2-ol (three-letter code: PXN) (formula: C<sub>17</sub>H<sub>36</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	1	Total	C O	0	0
			25	17 8		

- Molecule 8 is water.

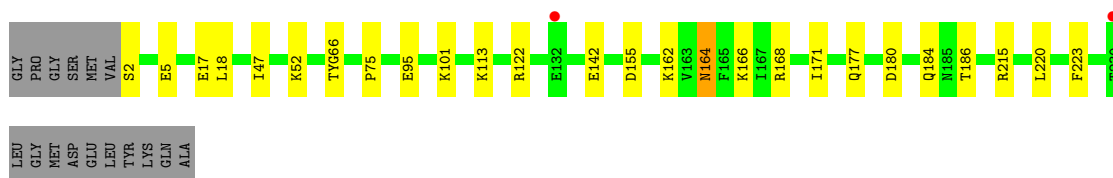
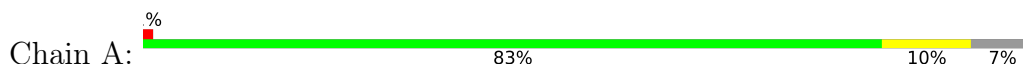
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	216	Total	O	0	0
			216	216		
8	B	143	Total	O	0	0
			143	143		
8	C	264	Total	O	0	0
			264	264		
8	D	121	Total	O	0	0
			121	121		



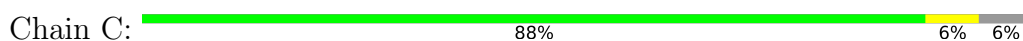
### 3 Residue-property plots [i](#)

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.

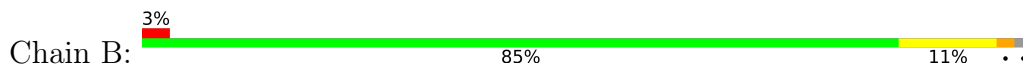
- Molecule 1: Green fluorescent protein



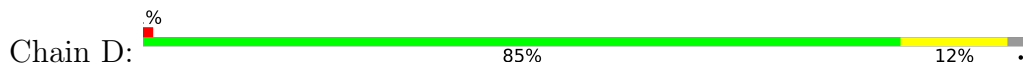
- Molecule 1: Green fluorescent protein



- Molecule 2: DARPin DP2



- Molecule 2: DARPin DP2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.87Å 95.16Å 69.47Å 90.00° 115.76° 90.00°	Depositor
Resolution (Å)	48.08 – 1.59 48.08 – 1.59	Depositor EDS
% Data completeness (in resolution range)	98.7 (48.08-1.59) 98.7 (48.08-1.59)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.43 (at 1.59Å)	Xtrriage
Refinement program	REFMAC 5.8.0352	Depositor
R, $R_{free}$	0.174 , 0.212 0.185 , 0.217	Depositor DCC
$R_{free}$ test set	4724 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.0	Xtrriage
Anisotropy	0.731	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 41.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.56$ , $\langle L^2 \rangle = 0.41$	Xtrriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	6867	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.18 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1873e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: A1H87, EDO, TRS, NA, CRO, PXN

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	0.51	1/1982 (0.1%)	0.80	2/2677 (0.1%)
1	C	0.55	0/2027	0.84	2/2736 (0.1%)
2	B	0.46	0/1001	0.79	0/1355
2	D	0.45	0/994	0.72	0/1346
All	All	0.50	1/6004 (0.0%)	0.80	4/8114 (0.0%)

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
2	B	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	142	GLU	CD-OE1	5.25	1.31	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	168	ARG	NE-CZ-NH1	7.43	124.01	120.30
1	C	168	ARG	NE-CZ-NH2	-7.26	116.67	120.30
1	A	168	ARG	NE-CZ-NH2	-6.48	117.06	120.30
1	A	168	ARG	NE-CZ-NH1	5.53	123.06	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	31	ARG	Sidechain

## 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	1964	0	1899	24	0
1	C	2009	0	1937	18	0
2	B	990	0	987	20	0
2	D	983	0	982	20	0
3	A	21	0	0	0	0
3	C	21	0	0	0	0
4	A	12	0	18	0	0
4	B	16	0	24	5	0
4	C	20	0	30	1	0
4	D	12	0	18	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	B	24	0	36	0	0
6	D	24	0	36	2	0
7	C	25	0	36	0	0
8	A	216	0	0	13	0
8	B	143	0	0	11	1
8	C	264	0	0	11	1
8	D	121	0	0	7	2
All	All	6867	0	6003	88	2

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 7.

The worst 5 of 88 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:113[A]:LYS:NZ	8:C:401:HOH:O	1.59	1.28
2:B:57[A]:VAL:HG21	2:B:59:HIS:CD2	1.92	1.04
1:C:223:PHE:CE1	8:D:301:HOH:O	2.20	0.95
2:B:57[A]:VAL:CG2	2:B:59:HIS:CD2	2.53	0.91

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:113[A]:LYS:NZ	8:A:401:HOH:O	1.75	0.86

All (2) 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)
8:C:594:HOH:O	8:D:321:HOH:O[2_245]	2.18	0.02
8:B:321:HOH:O	8:D:363:HOH:O[1_656]	2.19	0.01

## 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	240/243 (99%)	236 (98%)	4 (2%)	0	100	100
1	C	245/243 (101%)	241 (98%)	4 (2%)	0	100	100
2	B	130/128 (102%)	129 (99%)	1 (1%)	0	100	100
2	D	129/128 (101%)	128 (99%)	1 (1%)	0	100	100
All	All	744/742 (100%)	734 (99%)	10 (1%)	0	100	100

There are no Ramachandran outliers to report.

### 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	216/210 (103%)	214 (99%)	2 (1%)	75	62
1	C	221/210 (105%)	220 (100%)	1 (0%)	86	78
2	B	105/99 (106%)	101 (96%)	4 (4%)	28	8
2	D	104/99 (105%)	103 (99%)	1 (1%)	73	57
All	All	646/618 (104%)	638 (99%)	8 (1%)	73	50

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	135	LEU
1	C	1	VAL
2	B	61[A]	GLU
2	B	27	ASP
2	B	61[B]	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	204	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

2 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
1	CRO	A	66	1	23,23,24	1.09	1 (4%)	30,32,34	1.15	3 (10%)
1	CRO	C	66	1	23,23,24	0.87	1 (4%)	30,32,34	1.05	3 (10%)

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
1	CRO	A	66	1	-	0/12/31/32	0/2/2/2
1	CRO	C	66	1	-	0/12/31/32	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	66	CRO	C1-N2	4.06	1.38	1.32
1	C	66	CRO	C1-N2	2.79	1.36	1.32

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	66	CRO	CA2-C2-N3	3.45	105.00	103.37
1	A	66	CRO	CA2-C2-N3	3.33	104.94	103.37
1	A	66	CRO	C2-N3-C1	2.45	109.21	107.97
1	A	66	CRO	O3-C3-CA3	-2.37	119.23	126.39
1	C	66	CRO	O3-C3-CA3	-2.03	120.25	126.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 26 ligands modelled in this entry, 2 are monoatomic - leaving 24 for Mogul analysis.

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	EDO	B	201	-	3,3,3	0.12	0	2,2,2	0.35	0
6	TRS	B	206	-	7,7,7	0.19	0	9,9,9	0.35	0
7	PXN	C	302	-	20,24,24	0.25	0	22,30,30	0.46	0
3	A1H87	C	301	-	17,20,20	0.25	0	20,25,25	1.01	1 (5%)
6	TRS	B	207	-	7,7,7	0.24	0	9,9,9	0.27	0
4	EDO	C	306	-	3,3,3	0.22	0	2,2,2	0.40	0
6	TRS	D	206	-	7,7,7	0.21	0	9,9,9	0.37	0
3	A1H87	A	301	-	17,20,20	0.26	0	20,25,25	1.05	1 (5%)
4	EDO	D	201	-	3,3,3	0.73	0	2,2,2	0.39	0
4	EDO	C	307	-	3,3,3	0.19	0	2,2,2	0.20	0
4	EDO	B	203	-	3,3,3	0.50	0	2,2,2	0.04	0
4	EDO	B	202	-	3,3,3	0.34	0	2,2,2	1.04	0
4	EDO	A	302	-	3,3,3	0.24	0	2,2,2	0.05	0
4	EDO	C	305	-	3,3,3	0.27	0	2,2,2	0.34	0
4	EDO	A	303	-	3,3,3	0.32	0	2,2,2	0.27	0
6	TRS	D	204	-	7,7,7	0.42	0	9,9,9	0.44	0
4	EDO	D	203	-	3,3,3	0.06	0	2,2,2	0.22	0
4	EDO	D	202	-	3,3,3	0.08	0	2,2,2	0.22	0
6	TRS	B	205	-	7,7,7	0.38	0	9,9,9	0.60	0
4	EDO	B	204	-	3,3,3	0.30	0	2,2,2	0.57	0
4	EDO	C	303	-	3,3,3	0.22	0	2,2,2	0.75	0
4	EDO	A	304	-	3,3,3	0.14	0	2,2,2	0.05	0
6	TRS	D	205	-	7,7,7	0.33	0	9,9,9	0.55	0
4	EDO	C	304	-	3,3,3	0.37	0	2,2,2	0.04	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	EDO	B	201	-	-	1/1/1/1	-
6	TRS	B	206	-	-	3/9/9/9	-
7	PXN	C	302	-	-	15/28/28/28	-
3	A1H87	C	301	-	-	9/25/25/25	-
6	TRS	B	207	-	-	3/9/9/9	-
4	EDO	C	306	-	-	0/1/1/1	-
6	TRS	D	206	-	-	0/9/9/9	-

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	A1H87	A	301	-	-	8/25/25/25	-
4	EDO	D	201	-	-	0/1/1/1	-
4	EDO	C	307	-	-	1/1/1/1	-
4	EDO	B	203	-	-	1/1/1/1	-
4	EDO	B	202	-	-	1/1/1/1	-
4	EDO	A	302	-	-	1/1/1/1	-
4	EDO	C	305	-	-	1/1/1/1	-
4	EDO	A	303	-	-	1/1/1/1	-
6	TRS	D	204	-	-	0/9/9/9	-
4	EDO	D	203	-	-	1/1/1/1	-
4	EDO	D	202	-	-	0/1/1/1	-
6	TRS	B	205	-	-	0/9/9/9	-
4	EDO	B	204	-	-	1/1/1/1	-
4	EDO	C	303	-	-	0/1/1/1	-
4	EDO	A	304	-	-	1/1/1/1	-
6	TRS	D	205	-	-	3/9/9/9	-
4	EDO	C	304	-	-	1/1/1/1	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	301	A1H87	O1-C2-C	3.60	115.84	108.86
3	A	301	A1H87	O1-C2-C	3.45	115.55	108.86

There are no chirality outliers.

5 of 52 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	301	A1H87	C8-C7-C9-O4
3	A	301	A1H87	C6-C7-C8-O3
3	C	301	A1H87	C8-C7-C9-O4
3	C	301	A1H87	C6-C7-C9-O4
3	C	301	A1H87	C10-C7-C8-O3

There are no ring outliers.

4 monomers are involved in 8 short contacts:

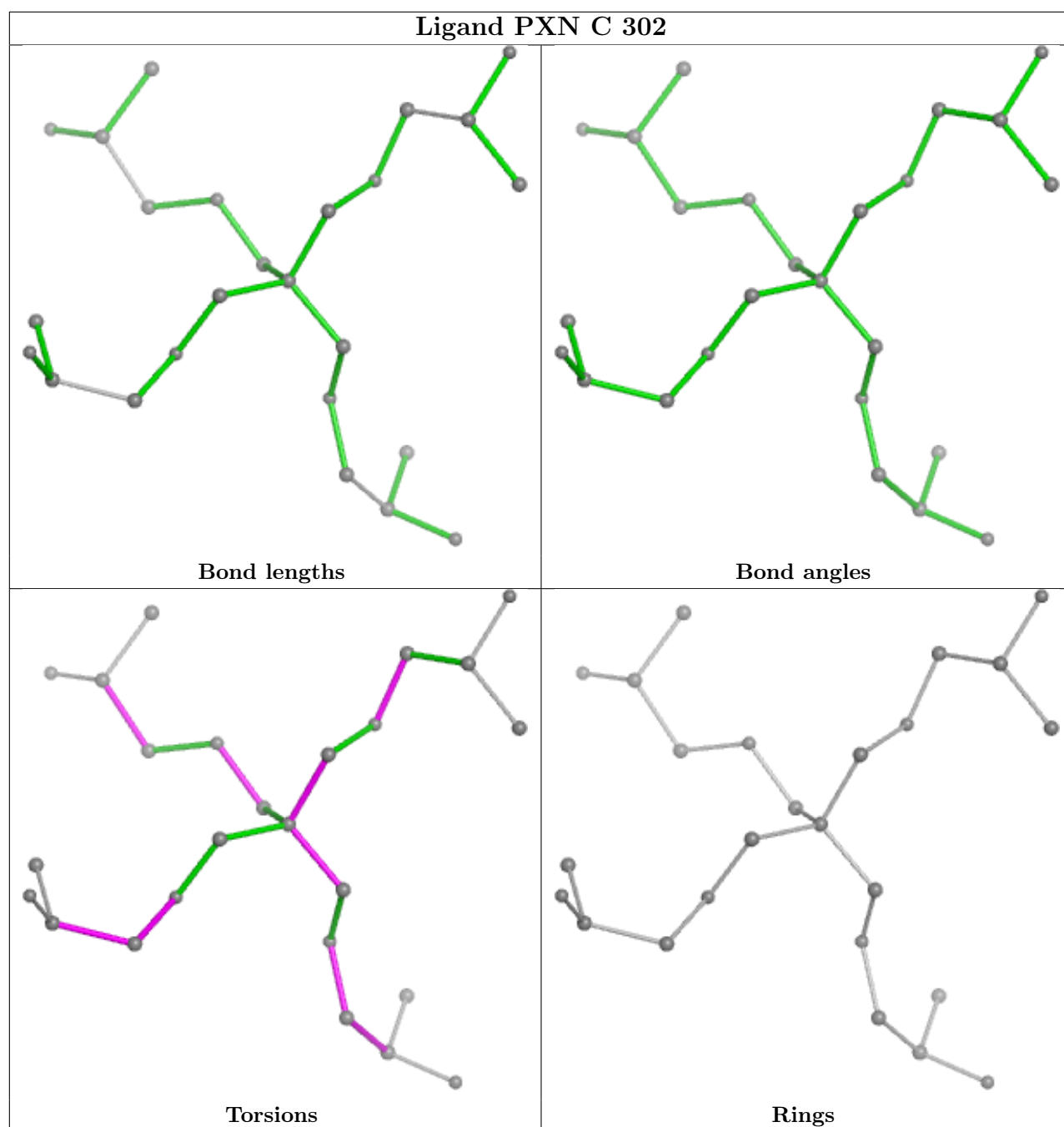
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	203	EDO	4	0

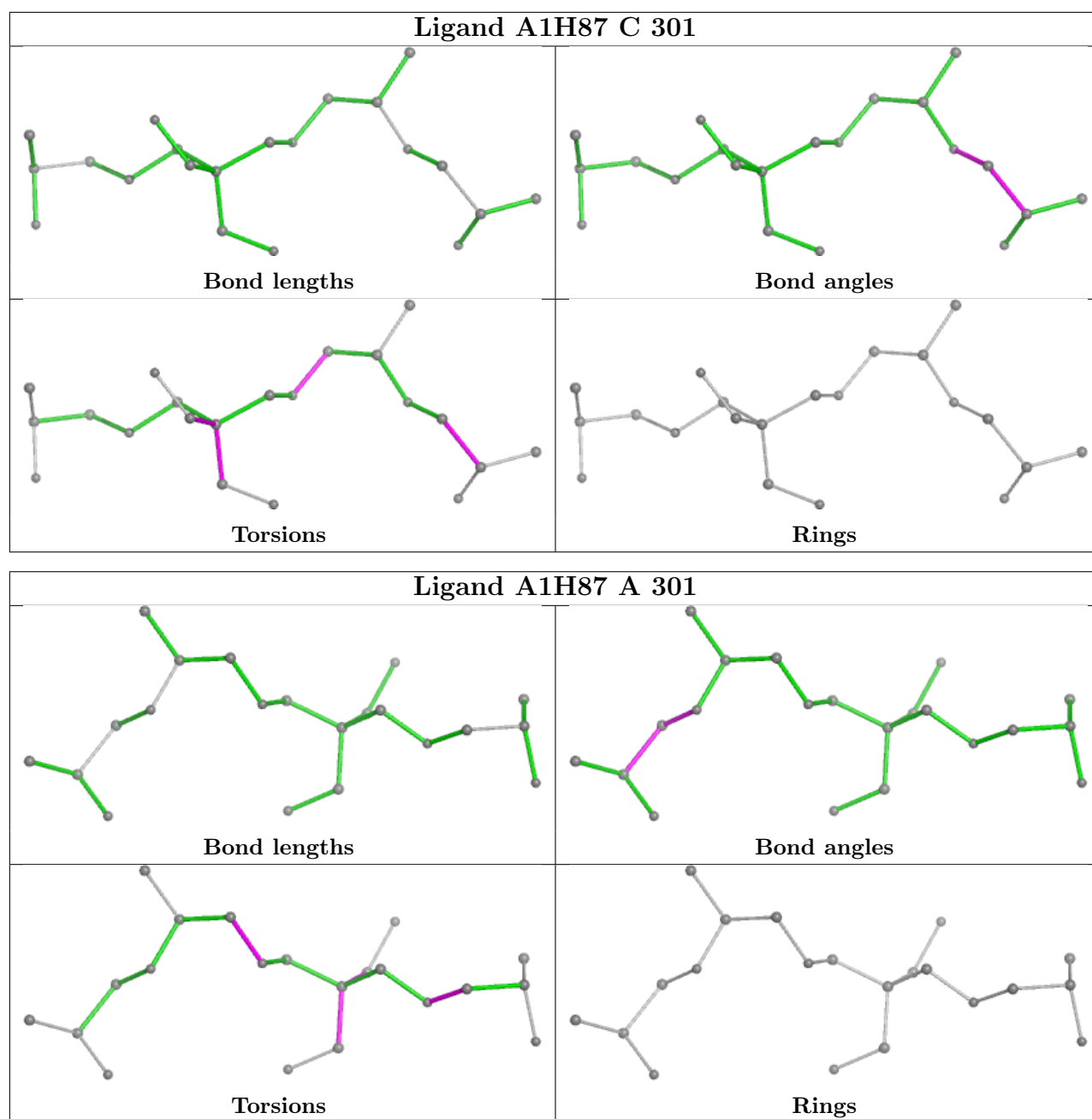
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	305	EDO	1	0
4	B	204	EDO	1	0
6	D	205	TRS	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 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	226/243 (93%)	-0.08	2 (0%) 81 83	6, 17, 33, 48	18 (7%)
1	C	227/243 (93%)	-0.27	1 (0%) 89 90	5, 14, 28, 45	22 (9%)
2	B	125/128 (97%)	0.12	4 (3%) 50 51	6, 18, 45, 54	7 (5%)
2	D	124/128 (96%)	0.11	1 (0%) 82 85	6, 19, 36, 52	7 (5%)
All	All	702/742 (94%)	-0.07	8 (1%) 77 80	5, 16, 34, 54	54 (7%)

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1	VAL	5.2
2	B	12	SER	2.9
2	B	135	LEU	2.6
1	A	230	THR	2.2
2	D	135	LEU	2.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	CRO	A	66	22/23	0.97	0.05	11,12,13,15	0
1	CRO	C	66	22/23	0.97	0.05	9,9,10,12	0

### 6.3 Carbohydrates [i](#)

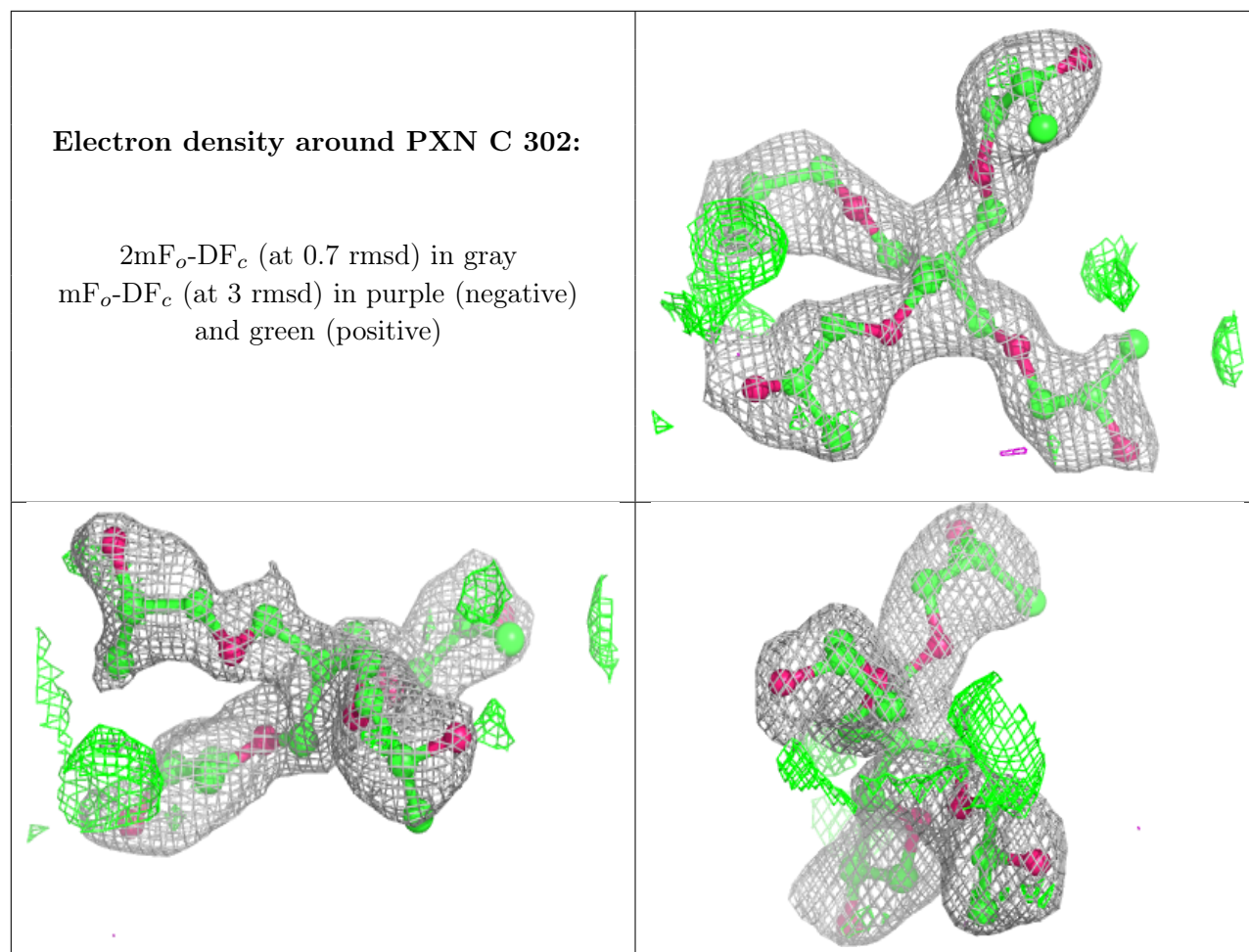
There are no monosaccharides in this entry.

## 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, 95<sup>th</sup> 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.

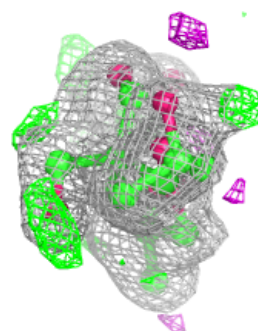
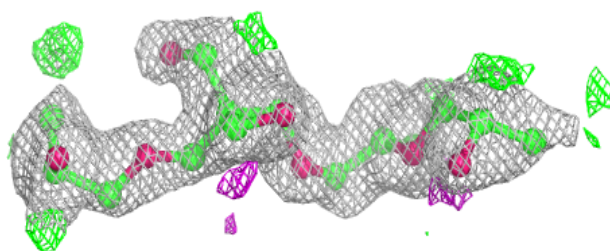
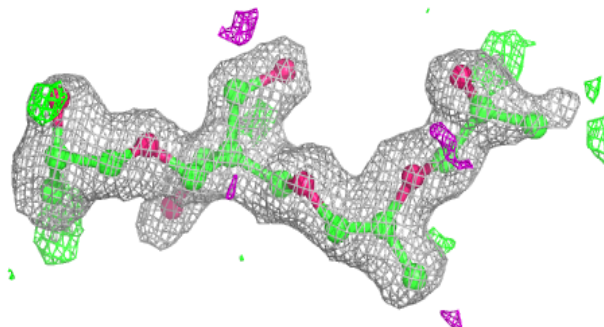
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	EDO	A	304	4/4	0.73	0.18	46,49,52,54	0
4	EDO	D	203	4/4	0.76	0.19	50,52,52,59	0
6	TRS	D	205	8/8	0.77	0.14	38,41,42,43	0
4	EDO	B	204	4/4	0.79	0.15	31,32,33,34	0
4	EDO	C	304	4/4	0.80	0.22	30,43,44,58	0
7	PXN	C	302	25/25	0.83	0.17	46,53,59,60	0
4	EDO	C	307	4/4	0.84	0.14	31,43,45,53	0
4	EDO	A	303	4/4	0.84	0.13	30,34,35,40	0
4	EDO	B	202	4/4	0.85	0.14	30,31,31,35	0
4	EDO	C	305	4/4	0.86	0.15	32,36,36,43	0
4	EDO	C	303	4/4	0.87	0.11	22,22,25,32	0
6	TRS	B	206	8/8	0.87	0.11	30,33,35,36	0
4	EDO	D	201	4/4	0.88	0.13	27,30,30,33	0
4	EDO	C	306	4/4	0.89	0.11	32,33,37,39	0
3	A1H87	C	301	21/21	0.89	0.11	22,32,39,40	0
3	A1H87	A	301	21/21	0.89	0.12	23,33,46,51	0
6	TRS	D	206	8/8	0.89	0.10	23,26,27,28	0
4	EDO	D	202	4/4	0.89	0.12	45,46,46,50	0
6	TRS	D	204	8/8	0.90	0.11	20,26,28,30	0
6	TRS	B	207	8/8	0.90	0.11	24,29,29,32	0
4	EDO	B	203	4/4	0.92	0.10	29,33,33,42	0
6	TRS	B	205	8/8	0.92	0.10	19,24,28,28	0
4	EDO	A	302	4/4	0.94	0.09	35,36,37,37	0
4	EDO	B	201	4/4	0.95	0.09	31,31,32,32	0
5	NA	C	308	1/1	0.99	0.06	13,13,13,13	0
5	NA	A	305	1/1	0.99	0.05	17,17,17,17	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.

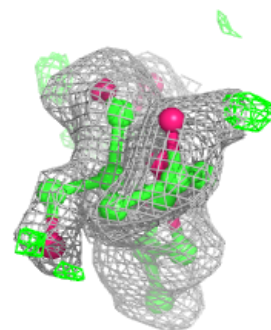
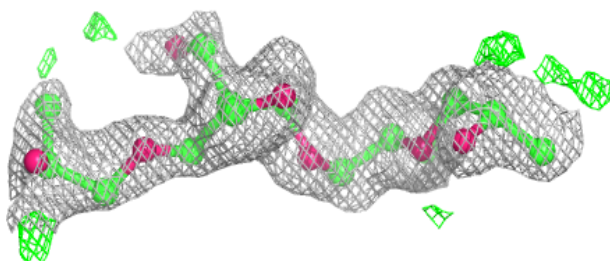
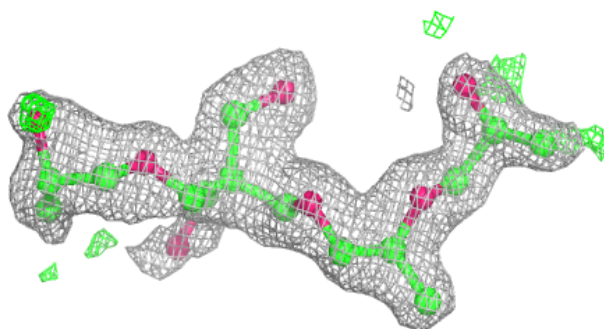


**Electron density around A1H87 C 301:**

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

**Electron density around A1H87 A 301:**

$2mF_o-DF_c$  (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.