



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 12:15 PM BST

PDB ID : 4A5S  
Title : CRYSTAL STRUCTURE OF HUMAN DPP4 IN COMPLEX WITH A NO-VAL HETEROCYCLIC DPP4 INHIBITOR  
Authors : Ostermann, N.; Kroemer, M.; Zink, F.; Gerhartz, B.; Sutton, J.M.; Clark, D.E.; Dunsdon, S.J.; Fenton, G.; Fillmore, A.; Harris, N.V.; Higgs, C.; Hurley, C.A.; Krintel, S.L.; MacKenzie, R.E.; Duttaroy, A.; Gangl, E.; Maniara, W.; Sedrani, R.; Namoto, K.; Sirockin, F.; Trappe, J.; Hassiepen, U.; Baeschlin, D.K.  
Deposited on : 2011-10-28  
Resolution : 1.62 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 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 : 2.13.1  
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) : 2.13.1

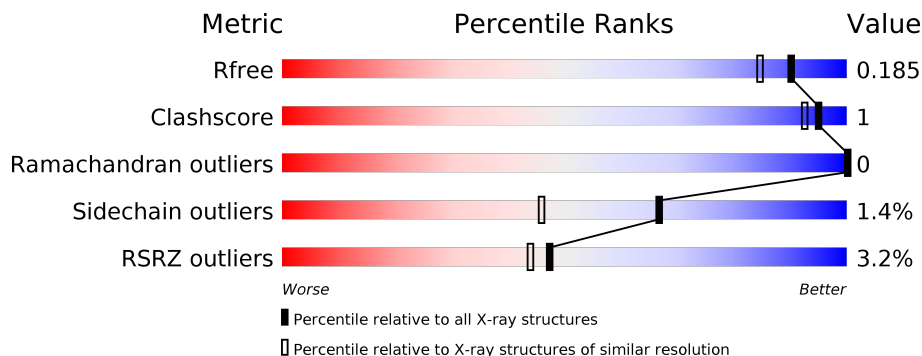
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.62 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	4693 (1.64-1.60)
Clashscore	141614	5002 (1.64-1.60)
Ramachandran outliers	138981	4888 (1.64-1.60)
Sidechain outliers	138945	4887 (1.64-1.60)
RSRZ outliers	127900	4609 (1.64-1.60)

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 on 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	740	<p>2% 94%</p>
1	B	740	<p>4% 95%</p>
2	C	5	<p>20% 80%</p>

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 13990 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 DIPEPTIDYL PEPTIDASE 4 SOLUBLE FORM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	728	6008	3862	985	1135	26	0	11	1
1	B	734	6050	3891	992	1141	26	0	11	1

There are 26 discrepancies between the modelled and reference sequences:

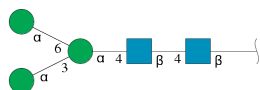
Chain	Residue	Modelled	Actual	Comment	Reference
A	767	ALA	-	expression tag	UNP P27487
A	768	ALA	-	expression tag	UNP P27487
A	769	ALA	-	expression tag	UNP P27487
A	770	SER	-	expression tag	UNP P27487
A	771	TRP	-	expression tag	UNP P27487
A	772	SER	-	expression tag	UNP P27487
A	773	HIS	-	expression tag	UNP P27487
A	774	PRO	-	expression tag	UNP P27487
A	775	GLN	-	expression tag	UNP P27487
A	776	PHE	-	expression tag	UNP P27487
A	777	GLU	-	expression tag	UNP P27487
A	778	LYS	-	expression tag	UNP P27487
A	437	ILE	SER	conflict	UNP P27487
B	767	ALA	-	expression tag	UNP P27487
B	768	ALA	-	expression tag	UNP P27487
B	769	ALA	-	expression tag	UNP P27487
B	770	SER	-	expression tag	UNP P27487
B	771	TRP	-	expression tag	UNP P27487
B	772	SER	-	expression tag	UNP P27487
B	773	HIS	-	expression tag	UNP P27487
B	774	PRO	-	expression tag	UNP P27487
B	775	GLN	-	expression tag	UNP P27487
B	776	PHE	-	expression tag	UNP P27487
B	777	GLU	-	expression tag	UNP P27487
B	778	LYS	-	expression tag	UNP P27487

*Continued on next page...*

Continued from previous page...

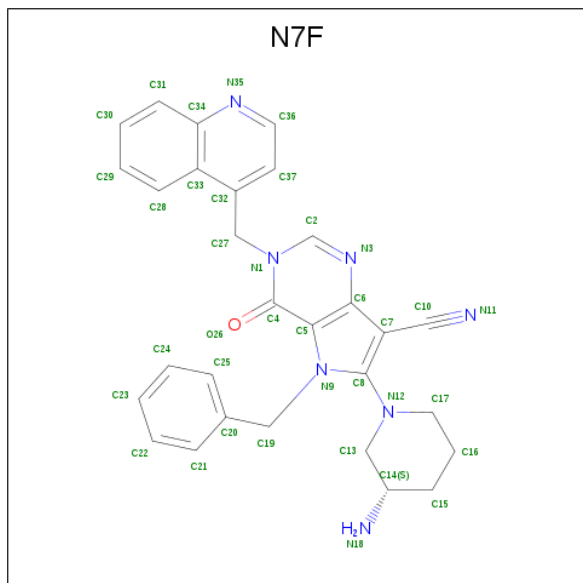
Chain	Residue	Modelled	Actual	Comment	Reference
B	437	ILE	SER	conflict	UNP P27487

- Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace	
			Total	C	N				O
2	C	5	61	34	2	25	0	0	0

- Molecule 3 is 6-[(3S)-3-AMINOPIPERIDIN-1-YL]-5-BENZYL-4-OXO-3-(QUINOLIN-4-YLMETHYL)-4,5-DIHYDRO-3H-PYRROLO[3,2-D]PYRIMIDINE-7-CARBONITRILE (three-letter code: N7F) (formula: C<sub>29</sub>H<sub>27</sub>N<sub>7</sub>O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	N			O
3	A	1	37	29	7	1	0	0
3	B	1	37	29	7	1	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		
4	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		

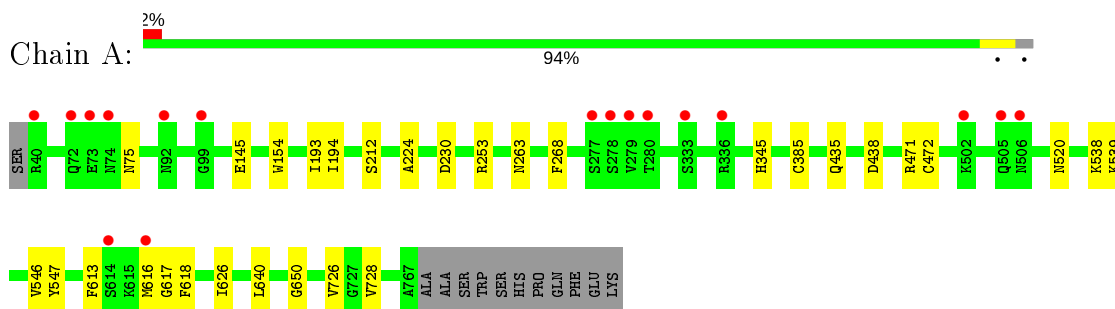
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	875	Total	O	0	5
			880	880		
6	B	715	Total	O	0	4
			719	719		

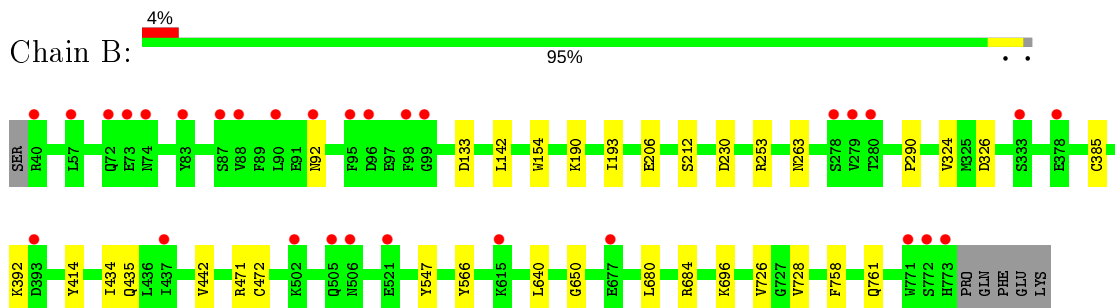
### 3 Residue-property plots [i](#)

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: DIPEPTIDYL PEPTIDASE 4 SOLUBLE FORM



- Molecule 1: DIPEPTIDYL PEPTIDASE 4 SOLUBLE FORM



- Molecule 2: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.18Å 121.42Å 190.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	68.06 – 1.62 51.49 – 1.62	Depositor EDS
% Data completeness (in resolution range)	99.7 (68.06-1.62) 99.7 (51.49-1.62)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.62 (at 1.62Å)	Xtrriage
Refinement program	BUSTER 2.11.2	Depositor
R, $R_{free}$	0.164 , 0.179 0.168 , 0.185	Depositor DCC
$R_{free}$ test set	5116 reflections (1.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.3	Xtrriage
Anisotropy	0.021	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 55.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	13990	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.84% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: SO4, N7F, NAG, MAN

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	0.55	0/6214	0.67	0/8452
1	B	0.50	0/6258	0.65	0/8513
All	All	0.53	0/12472	0.66	0/16965

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

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	6008	0	5749	15	0
1	B	6050	0	5794	17	0
2	C	61	0	52	0	0
3	A	37	0	27	0	0
3	B	37	0	27	1	0
4	A	84	0	78	2	0
4	B	84	0	78	0	0
5	A	15	0	0	0	0
5	B	15	0	0	0	0
6	A	880	0	0	2	0
6	B	719	0	0	0	0
All	All	13990	0	11805	30	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 1.

All (30) 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:A:253:ARG:HH21	1:B:253:ARG:HH21	1.23	0.87
1:B:434[A]:ILE:HG22	1:B:442:VAL:HG22	1.61	0.82
1:A:538:LYS:O	1:A:618:PHE:HA	1.84	0.77
1:A:520:ASN:HD22	4:A:1520:NAG:H83	1.62	0.63
1:B:726:VAL:HG23	1:B:728:VAL:HG23	1.82	0.61
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.86	0.58
1:B:434[A]:ILE:CG2	1:B:442:VAL:HG22	2.37	0.54
1:B:696:LYS:HG2	1:B:728:VAL:HG22	1.89	0.54
1:A:435:GLN:HE21	1:A:438:ASP:HB3	1.75	0.51
1:B:190:LYS:HD3	1:B:193:ILE:HD12	1.93	0.50
1:A:613:PHE:O	1:A:616:MET:HG2	2.12	0.50
1:A:640:LEU:HD11	1:A:650:GLY:HA3	1.94	0.50
1:A:193[B]:ILE:HG22	1:A:194:ILE:HG12	1.95	0.49
4:A:1520:NAG:H81	6:A:3309:HOH:O	2.13	0.49
1:A:253:ARG:HH21	1:B:253:ARG:NH2	2.03	0.47
1:B:696:LYS:CG	1:B:728:VAL:HG22	2.45	0.46
1:A:345:HIS:HD2	6:A:3523:HOH:O	1.97	0.46
1:A:546[B]:VAL:HG21	1:A:626:ILE:HD11	1.98	0.46
1:A:253:ARG:NH2	1:B:253:ARG:HH21	2.03	0.46
1:A:224:ALA:HB1	1:A:268:PHE:CZ	2.52	0.45
1:A:154:TRP:CE2	1:A:212[B]:SER:HB3	2.53	0.44
1:B:133:ASP:HB3	1:B:142:LEU:HD11	2.00	0.43
1:B:206:GLU:OE1	3:B:901:N7F:H152	2.18	0.43
1:B:640:LEU:HD11	1:B:650:GLY:HA3	2.00	0.43
1:B:154:TRP:CE2	1:B:212[B]:SER:HB3	2.53	0.42
1:B:414:TYR:CE1	1:B:435:GLN:HG2	2.54	0.42
1:B:680:LEU:HD11	1:B:684:ARG:CZ	2.49	0.42
1:A:539:LYS:HG2	1:A:617:GLY:O	2.19	0.42
1:B:758:PHE:O	1:B:761:GLN:HG3	2.21	0.41
1:B:290:PRO:HG3	1:B:324[B]:VAL:HG11	2.04	0.40

There are no symmetry-related clashes.

## 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	737/740 (100%)	714 (97%)	23 (3%)	0	100	100
1	B	743/740 (100%)	721 (97%)	22 (3%)	0	100	100
All	All	1480/1480 (100%)	1435 (97%)	45 (3%)	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	663/662 (100%)	655 (99%)	8 (1%)	71	52
1	B	666/662 (101%)	656 (98%)	10 (2%)	65	43
All	All	1329/1324 (100%)	1311 (99%)	18 (1%)	67	46

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	75	ASN
1	A	145	GLU
1	A	230	ASP
1	A	263	ASN
1	A	385	CYS
1	A	471	ARG
1	A	472	CYS
1	A	547	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	92	ASN
1	B	230	ASP
1	B	263	ASN
1	B	326	ASP
1	B	385	CYS
1	B	392	LYS
1	B	471	ARG
1	B	472	CYS
1	B	547	TYR
1	B	566	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	263	ASN
1	A	345	HIS
1	A	435	GLN
1	A	572	ASN
1	B	92	ASN
1	B	263	ASN
1	B	344	GLN
1	B	363	HIS
1	B	435	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](#)

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

## 5.5 Carbohydrates [i](#)

5 monosaccharides 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
2	NAG	C	1	1,2	14,14,15	0.29	0	17,19,21	0.63	0
2	NAG	C	2	2	14,14,15	0.37	0	17,19,21	1.04	2 (11%)
2	MAN	C	3	2	11,11,12	0.44	0	15,15,17	2.16	4 (26%)
2	MAN	C	4	2	11,11,12	0.33	0	15,15,17	0.80	1 (6%)
2	MAN	C	5	2	11,11,12	0.37	0	15,15,17	0.64	1 (6%)

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
2	NAG	C	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	C	2	2	-	2/6/23/26	0/1/1/1
2	MAN	C	3	2	-	0/2/19/22	1/1/1/1
2	MAN	C	4	2	-	0/2/19/22	0/1/1/1
2	MAN	C	5	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	3	MAN	C1-O5-C5	6.08	120.43	112.19
2	C	3	MAN	O2-C2-C1	-4.12	100.72	109.15
2	C	4	MAN	C1-O5-C5	3.02	116.28	112.19
2	C	2	NAG	O4-C4-C5	-2.87	102.17	109.30
2	C	3	MAN	C1-C2-C3	2.79	113.10	109.67
2	C	2	NAG	O5-C1-C2	-2.47	107.39	111.29
2	C	5	MAN	C1-O5-C5	2.07	115.00	112.19
2	C	3	MAN	O5-C1-C2	2.07	113.96	110.77

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	2	NAG	C8-C7-N2-C2

*Continued on next page...*

Continued from previous page...

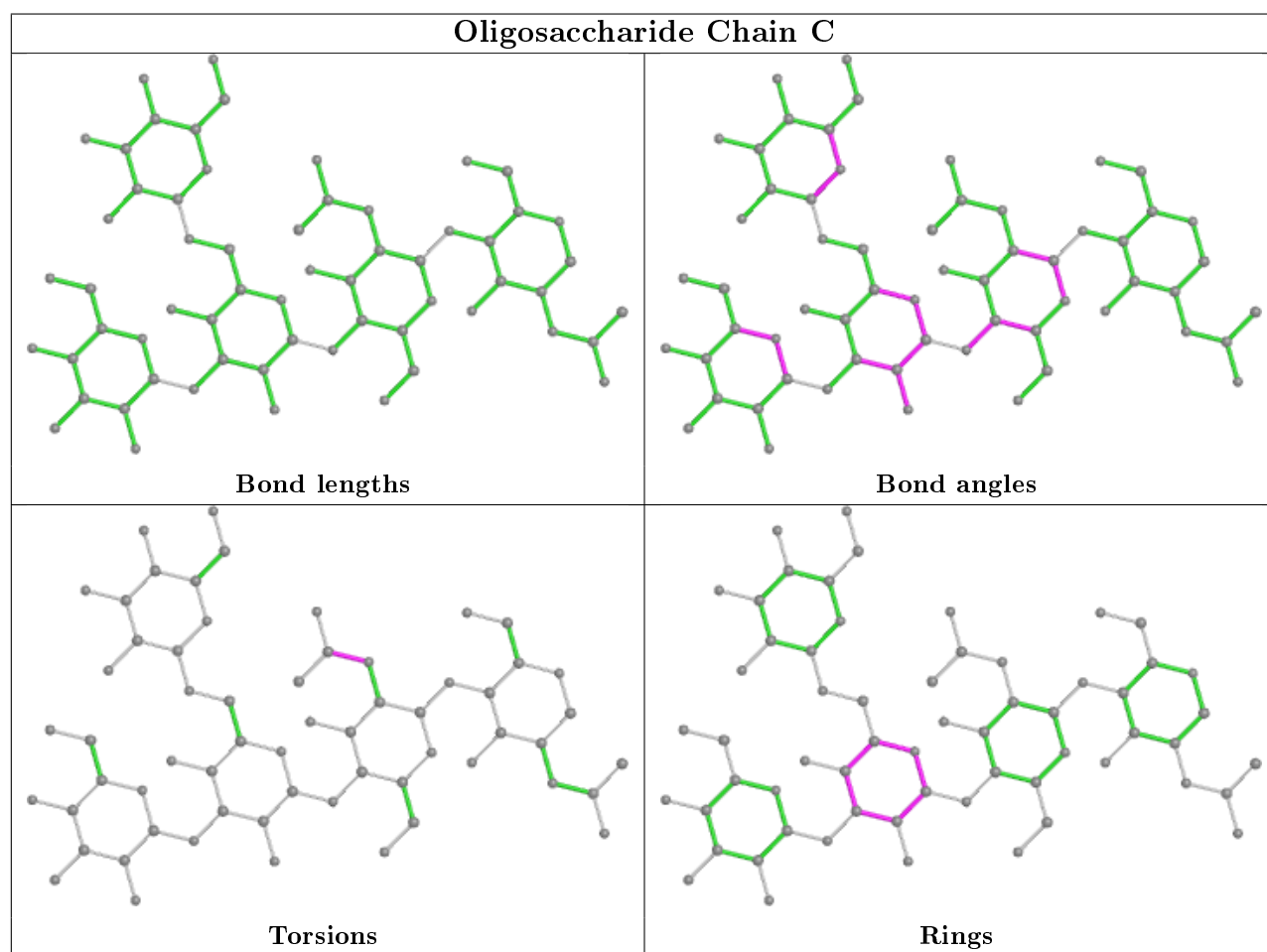
Mol	Chain	Res	Type	Atoms
2	C	2	NAG	O7-C7-N2-C2

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	3	MAN	C1-C2-C3-C4-C5-O5

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 oligosaccharide.



## 5.6 Ligand geometry [i](#)

20 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
5	SO4	A	1768	-	4,4,4	0.13	0	6,6,6	0.26	0
4	NAG	A	1281	1	14,14,15	0.29	0	17,19,21	0.62	0
4	NAG	A	1321	1	14,14,15	0.28	0	17,19,21	1.04	1 (5%)
5	SO4	B	1773	-	4,4,4	0.34	0	6,6,6	0.27	0
5	SO4	B	1774	-	4,4,4	0.18	0	6,6,6	0.10	0
4	NAG	B	2219	1	14,14,15	0.28	0	17,19,21	0.48	0
4	NAG	B	2229	1	14,14,15	0.29	0	17,19,21	0.48	0
4	NAG	A	1219	1	14,14,15	0.28	0	17,19,21	0.46	0
4	NAG	B	2085	1	14,14,15	0.28	0	17,19,21	0.65	1 (5%)
3	N7F	B	901	-	39,42,42	1.85	6 (15%)	43,60,60	2.14	11 (25%)
4	NAG	B	2150	1	14,14,15	0.31	0	17,19,21	0.68	1 (5%)
4	NAG	B	2281	1	14,14,15	0.30	0	17,19,21	0.57	0
4	NAG	B	2321	1	14,14,15	0.31	0	17,19,21	0.90	2 (11%)
4	NAG	A	1520	1	14,14,15	0.36	0	17,19,21	0.56	0
5	SO4	A	1769	-	4,4,4	0.21	0	6,6,6	0.11	0
5	SO4	A	1767	-	4,4,4	0.14	0	6,6,6	0.27	0
3	N7F	A	901	-	39,42,42	1.70	6 (15%)	43,60,60	2.24	11 (25%)
5	SO4	B	1775	-	4,4,4	0.17	0	6,6,6	0.17	0
4	NAG	A	1229	1	14,14,15	0.38	0	17,19,21	0.67	0
4	NAG	A	1150	1	14,14,15	0.28	0	17,19,21	0.43	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	NAG	B	2085	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1321	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2219	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2229	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1219	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1281	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2150	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2281	1	-	0/6/23/26	0/1/1/1
4	NAG	B	2321	1	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1520	1	-	2/6/23/26	0/1/1/1
3	N7F	B	901	-	-	5/8/24/24	0/6/6/6
3	N7F	A	901	-	-	5/8/24/24	0/6/6/6
4	NAG	A	1229	1	-	0/6/23/26	0/1/1/1
4	NAG	A	1150	1	-	0/6/23/26	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	901	N7F	C6-N3	5.76	1.44	1.37
3	A	901	N7F	C6-N3	5.69	1.44	1.37
3	B	901	N7F	C4-C5	5.04	1.49	1.41
3	A	901	N7F	C7-C8	-4.85	1.35	1.42
3	B	901	N7F	C7-C8	-4.78	1.35	1.42
3	A	901	N7F	C4-C5	3.69	1.47	1.41
3	B	901	N7F	C7-C6	2.85	1.45	1.41
3	B	901	N7F	C37-C32	2.47	1.42	1.37
3	A	901	N7F	C37-C32	2.36	1.41	1.37
3	B	901	N7F	C30-C31	2.10	1.41	1.36
3	A	901	N7F	C30-C31	2.08	1.41	1.36
3	A	901	N7F	C15-C14	2.01	1.56	1.51

All (27) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	901	N7F	C4-C5-C6	-6.25	115.95	119.96
3	A	901	N7F	C15-C14-C13	-5.96	103.00	110.65
3	B	901	N7F	C20-C19-N9	5.93	121.83	112.63
3	A	901	N7F	C5-C4-N1	-5.86	115.93	120.46
3	A	901	N7F	C4-C5-C6	-5.59	116.37	119.96
3	B	901	N7F	C5-C4-N1	-4.48	117.00	120.46
3	A	901	N7F	C16-C15-C14	-3.93	106.90	111.87
3	B	901	N7F	C15-C14-C13	-3.90	105.64	110.65
3	A	901	N7F	C14-C13-N12	3.79	114.42	109.64
3	A	901	N7F	C37-C36-N35	-3.76	118.76	124.58
3	A	901	N7F	C20-C19-N9	3.70	118.38	112.63
3	A	901	N7F	C36-N35-C34	3.57	122.45	116.93
3	B	901	N7F	C16-C15-C14	-3.38	107.60	111.87
3	A	901	N7F	C25-C20-C21	3.24	123.26	118.17
3	B	901	N7F	C36-N35-C34	3.14	121.79	116.93
4	A	1321	NAG	O5-C1-C2	-3.11	106.38	111.29

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	901	N7F	C2-N3-C6	2.77	119.35	116.76
3	B	901	N7F	C37-C36-N35	-2.74	120.34	124.58
3	B	901	N7F	C25-C20-C21	2.66	122.34	118.17
4	B	2321	NAG	C1-O5-C5	2.65	115.78	112.19
3	A	901	N7F	C19-C20-C21	-2.50	115.66	120.40
4	B	2085	NAG	C1-O5-C5	2.28	115.28	112.19
4	B	2321	NAG	O5-C1-C2	-2.24	107.75	111.29
3	A	901	N7F	C24-C25-C20	-2.22	117.22	120.63
3	B	901	N7F	C33-C34-N35	-2.17	120.52	122.83
3	B	901	N7F	C22-C21-C20	-2.17	117.31	120.63
4	B	2150	NAG	C1-O5-C5	2.07	115.00	112.19

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	901	N7F	C32-C27-N1-C4
3	B	901	N7F	N1-C27-C32-C33
3	A	901	N7F	C32-C27-N1-C4
3	A	901	N7F	N1-C27-C32-C33
4	A	1520	NAG	C8-C7-N2-C2
4	A	1520	NAG	O7-C7-N2-C2
3	B	901	N7F	N9-C19-C20-C21
3	B	901	N7F	C32-C27-N1-C2
3	A	901	N7F	C32-C27-N1-C2
3	A	901	N7F	N9-C19-C20-C25
3	B	901	N7F	N9-C19-C20-C25
3	A	901	N7F	N9-C19-C20-C21

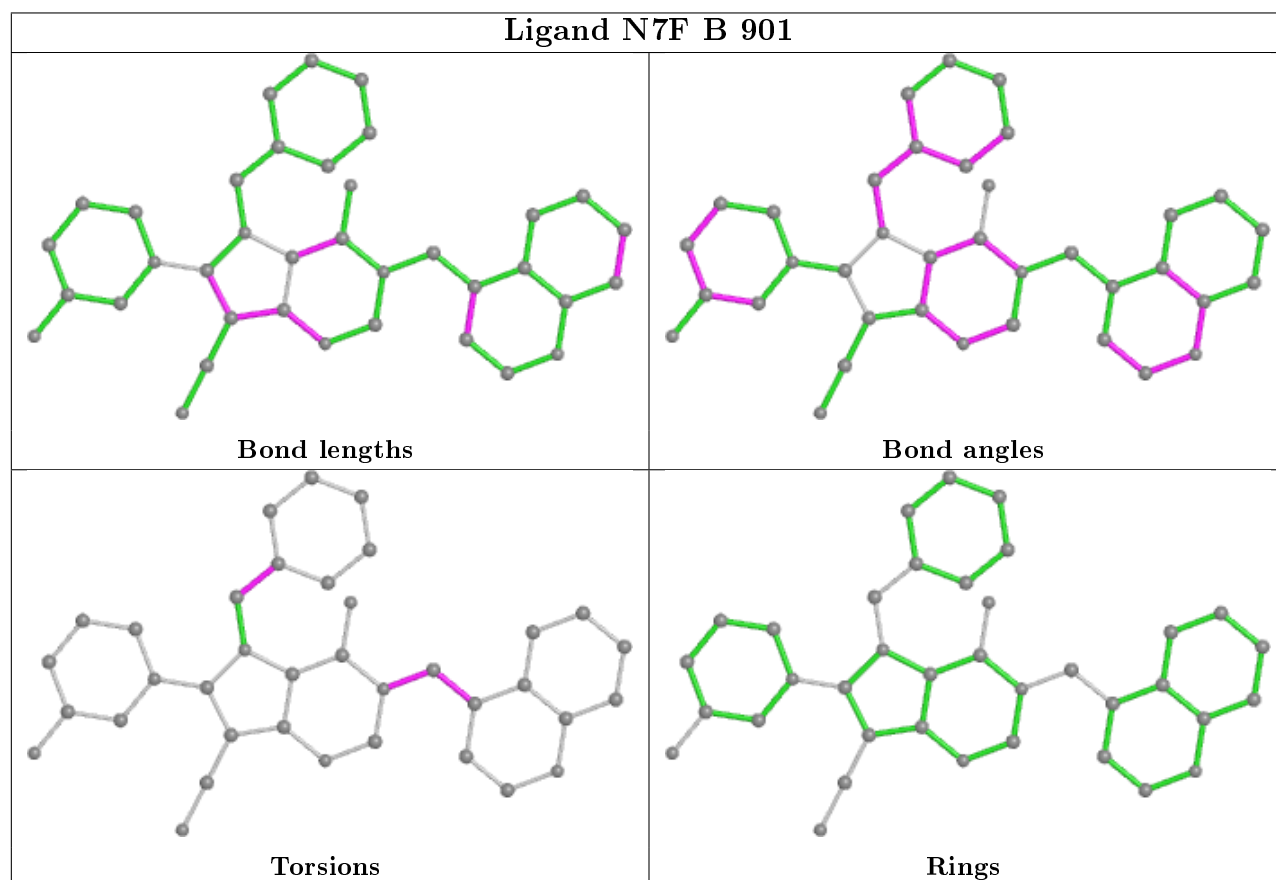
There are no ring outliers.

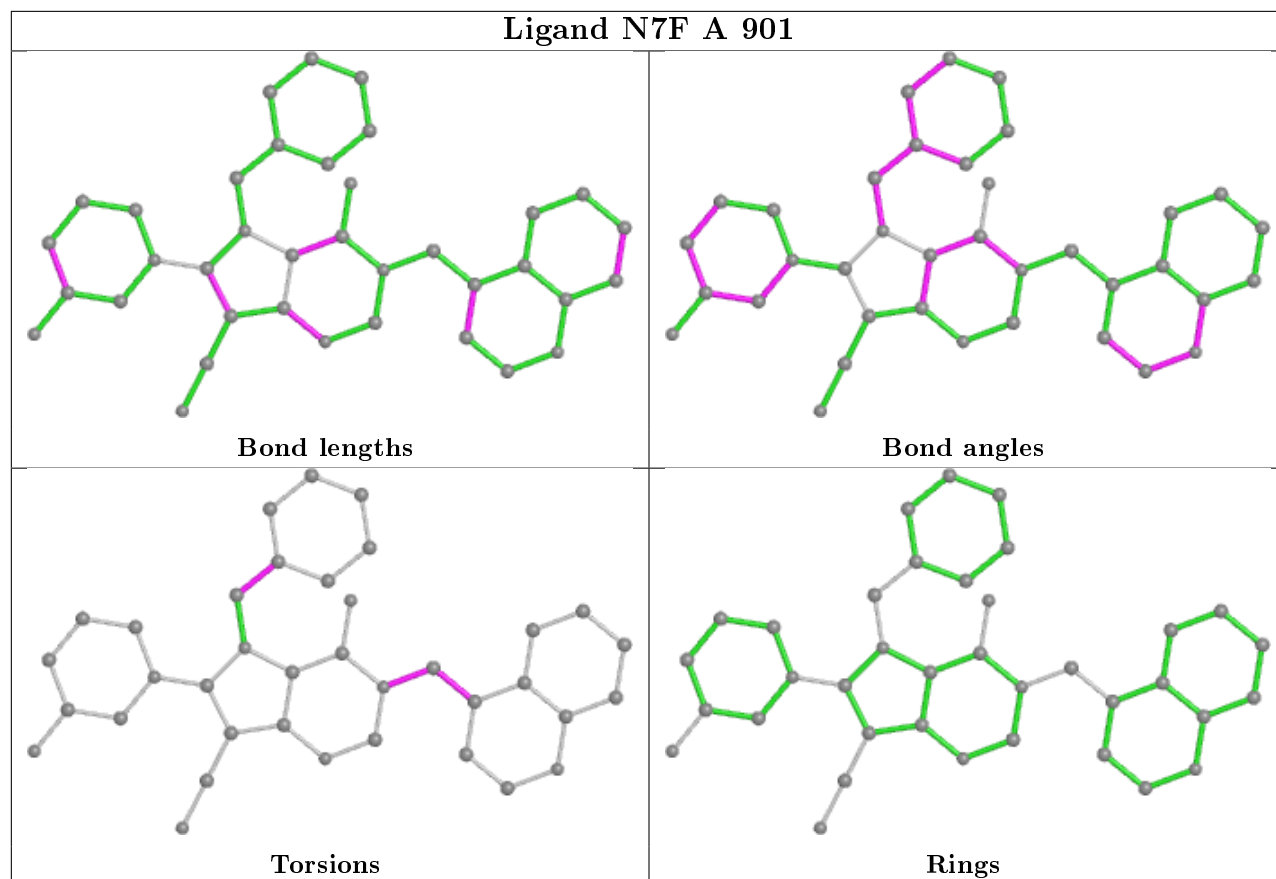
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	901	N7F	1	0
4	A	1520	NAG	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

### 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, 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	728/740 (98%)	-0.20	17 (2%) 60 58	11, 19, 40, 67	0
1	B	734/740 (99%)	-0.10	30 (4%) 37 33	12, 22, 49, 89	0
All	All	1462/1480 (98%)	-0.15	47 (3%) 47 44	11, 21, 46, 89	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	773	HIS	6.5
1	B	74	ASN	5.5
1	B	279	VAL	5.4
1	A	279	VAL	5.4
1	A	74	ASN	5.3
1	A	92	ASN	5.2
1	B	73	GLU	5.1
1	A	40	ARG	4.9
1	B	98	PHE	4.8
1	B	96	ASP	4.4
1	B	437	ILE	4.3
1	A	616	MET	4.3
1	B	40	ARG	4.0
1	B	95	PHE	3.9
1	B	72	GLN	3.8
1	B	278	SER	3.7
1	B	99	GLY	3.7
1	A	73	GLU	3.6
1	B	88	VAL	3.3
1	B	83	TYR	3.3
1	A	72	GLN	3.1
1	A	506	ASN	3.0
1	B	615	LYS	3.0
1	B	521	GLU	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	333	SER	2.9
1	B	772	SER	2.9
1	B	333	SER	2.7
1	B	771	TRP	2.7
1	A	336	ARG	2.7
1	A	505	GLN	2.6
1	A	280	THR	2.6
1	A	99	GLY	2.5
1	A	278	SER	2.5
1	B	92	ASN	2.5
1	B	506	ASN	2.4
1	A	502	LYS	2.4
1	B	87	SER	2.4
1	B	280	THR	2.4
1	B	393	ASP	2.3
1	B	378	GLU	2.3
1	A	614	SER	2.2
1	B	57	LEU	2.2
1	A	277	SER	2.1
1	B	505	GLN	2.1
1	B	502	LYS	2.1
1	B	677	GLU	2.1
1	B	90	LEU	2.0

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

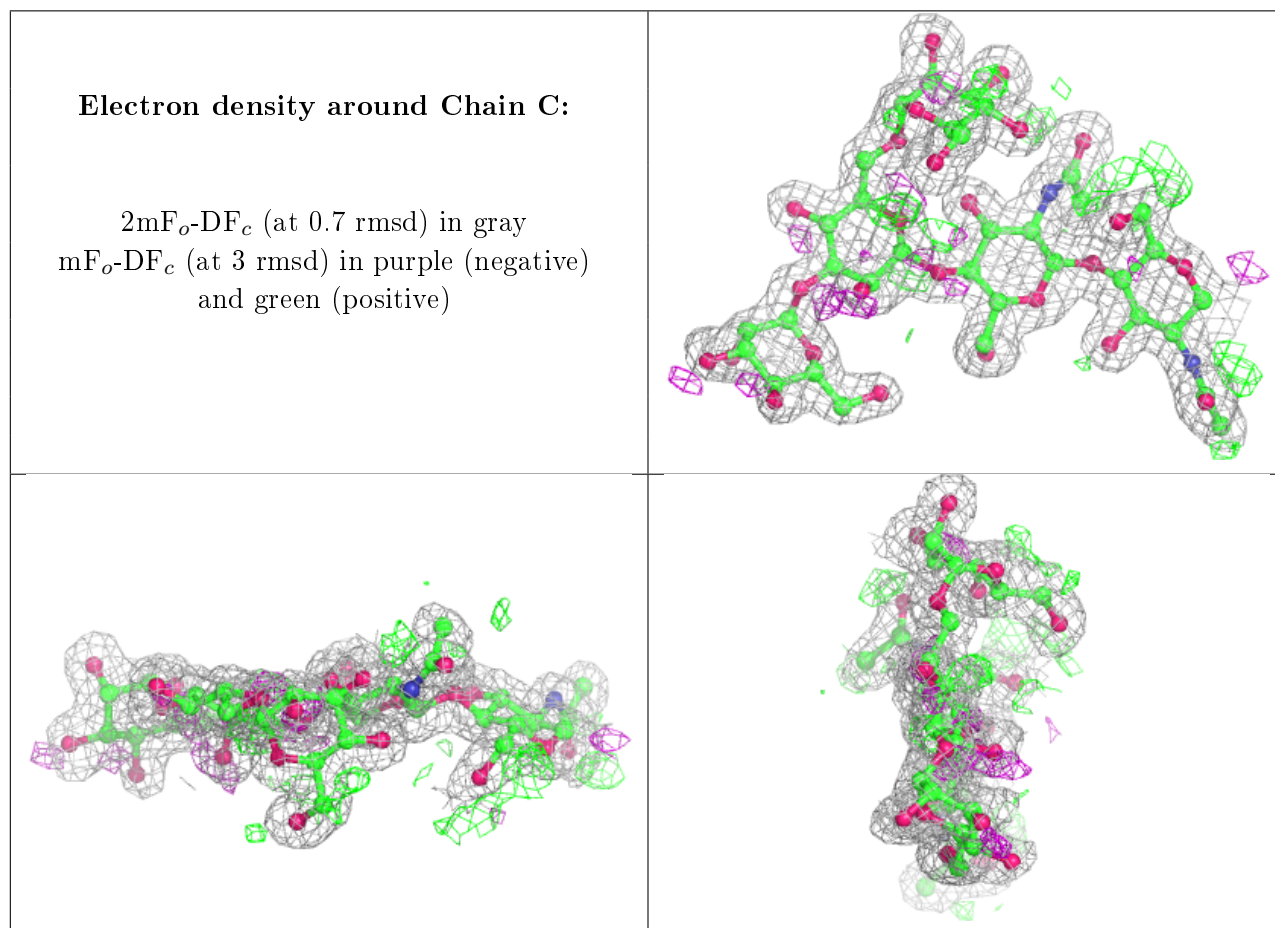
There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

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
2	MAN	C	3	11/12	0.88	0.14	21,26,31,32	0
2	NAG	C	2	14/15	0.90	0.12	25,27,36,36	0
2	MAN	C	4	11/12	0.90	0.21	35,40,43,43	0
2	MAN	C	5	11/12	0.94	0.12	21,24,27,29	0
2	NAG	C	1	14/15	0.95	0.10	25,26,30,35	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



## 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.

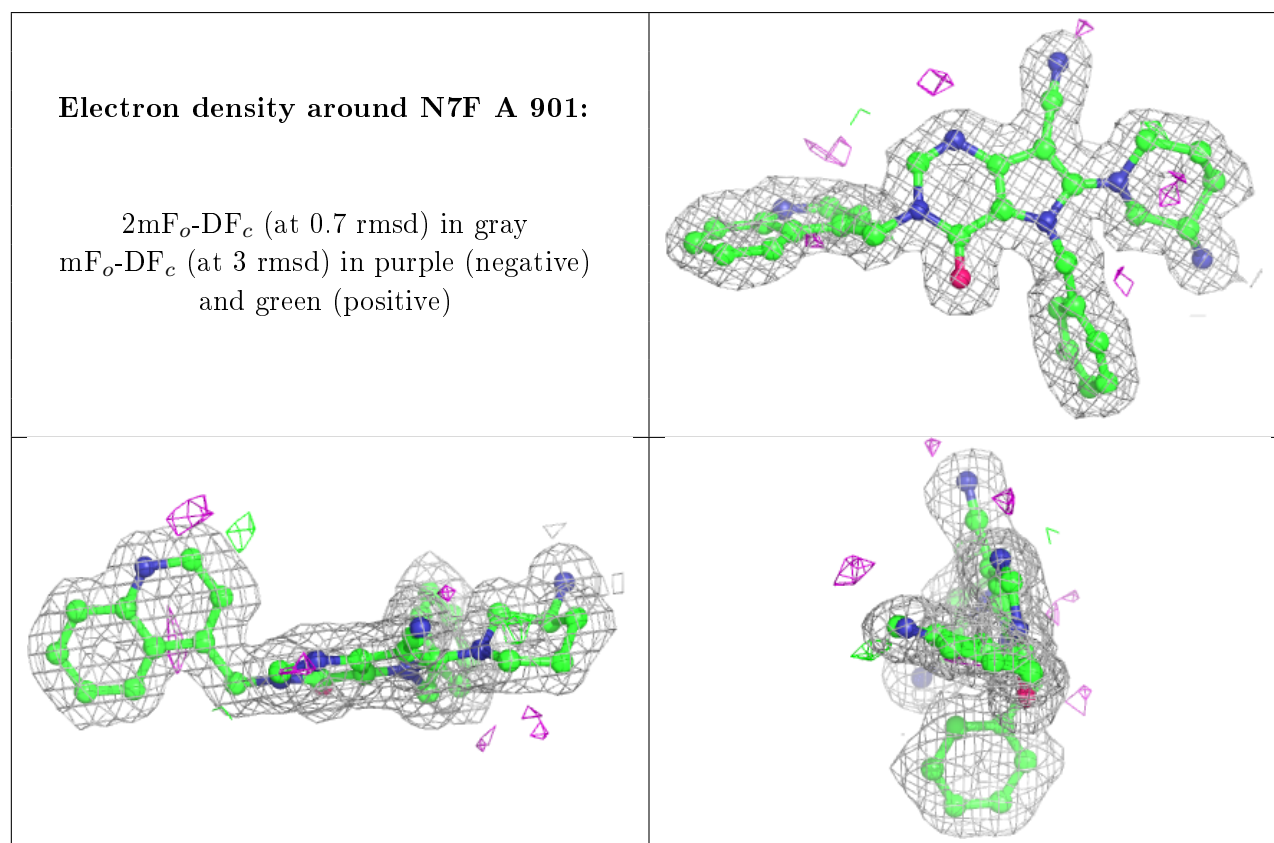
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	2321	14/15	0.79	0.22	30,37,44,46	0
4	NAG	B	2150	14/15	0.82	0.15	34,39,45,49	0
4	NAG	B	2085	14/15	0.84	0.23	51,55,60,61	0
4	NAG	A	1219	14/15	0.85	0.21	29,41,52,54	0
4	NAG	A	1150	14/15	0.87	0.23	35,39,44,44	0
4	NAG	B	2219	14/15	0.88	0.17	32,40,51,51	0
4	NAG	A	1281	14/15	0.88	0.17	36,43,46,47	0
4	NAG	A	1321	14/15	0.88	0.19	20,31,38,39	0
5	SO4	B	1773	5/5	0.89	0.18	45,51,54,57	0

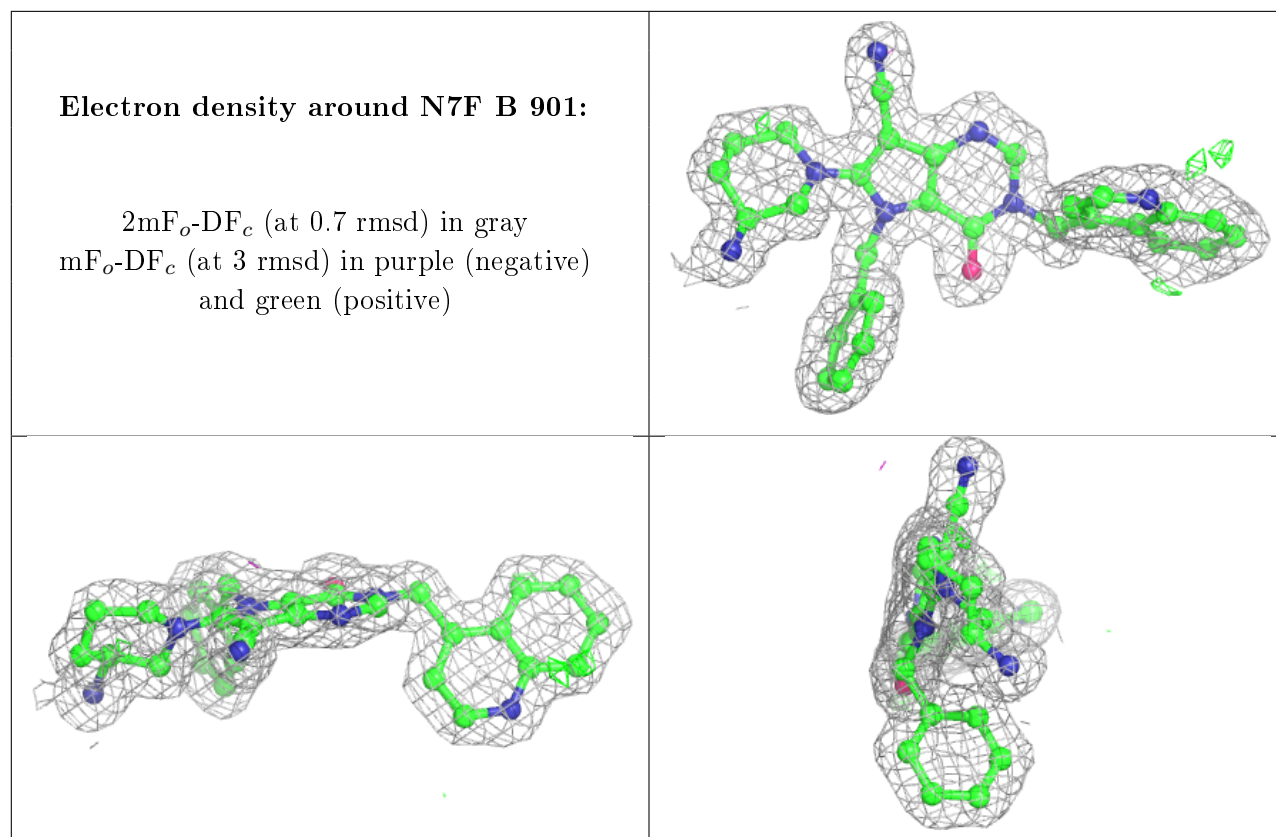
*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	B	1775	5/5	0.90	0.29	80,81,82,83	0
4	NAG	B	2281	14/15	0.92	0.18	31,39,43,43	0
5	SO4	B	1774	5/5	0.92	0.15	78,79,80,81	0
4	NAG	A	1520	14/15	0.93	0.13	17,26,32,35	0
4	NAG	B	2229	14/15	0.94	0.08	22,30,34,38	0
4	NAG	A	1229	14/15	0.94	0.13	23,31,36,38	0
5	SO4	A	1767	5/5	0.94	0.15	51,54,55,56	0
5	SO4	A	1769	5/5	0.95	0.26	63,65,65,67	0
3	N7F	A	901	37/37	0.96	0.09	11,15,19,24	0
3	N7F	B	901	37/37	0.96	0.09	14,18,23,28	0
5	SO4	A	1768	5/5	0.98	0.18	43,45,46,46	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.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.