

# wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2023 – 02:47 PM JST

PDB ID	:	5ZID
Title	:	Crystal Structure of human DPP-IV in complex with HL2
Authors	:	Zhu, L.; Li, H.; Wu, F.
Deposited on		
Resolution	:	3.00  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 (i) 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 (1)) were used in the production of this report:

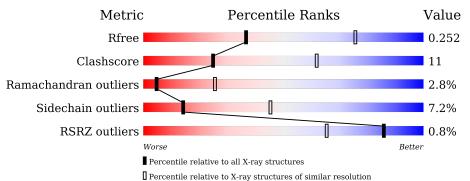
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
$\mathrm{EDS}$	:	2.36
buster-report	:	1.1.7(2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber $(2001)$
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

# 1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:  $X\text{-}RAY \, DIFFRACTION$ 

The reported resolution of this entry is 3.00 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	$\begin{array}{c} \textbf{Whole archive} \\ (\#\textbf{Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R <sub>free</sub>	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain						
1	А	734	% 71%	26%	•••				
1	В	734	% 67%	28%	•••				
2	С	4	75%	25%					



# 2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 12088 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.

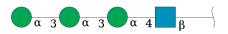
• Molecule 1 is a protein called Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	А	727	Total 5957	C 3824	N 981	O 1126	S 26	0	0	0
1	В	727	Total 5957	C 3824	N 981	O 1126	S 26	0	0	0

Chain	Residue	Modelled	Actual	Comment	Reference
A	767	GLY	-	expression tag	UNP P27487
А	768	HIS	-	expression tag	UNP P27487
A	769	HIS	-	expression tag	UNP P27487
A	770	HIS	-	expression tag	UNP P27487
А	771	HIS	-	expression tag	UNP P27487
A	772	HIS	-	expression tag	UNP P27487
A	773	HIS	-	expression tag	UNP P27487
В	767	GLY	-	expression tag	UNP P27487
В	768	HIS	-	expression tag	UNP P27487
В	769	HIS	-	expression tag	UNP P27487
В	770	HIS	-	expression tag	UNP P27487
В	771	HIS	-	expression tag	UNP P27487
В	772	HIS	-	expression tag	UNP P27487
В	773	HIS	-	expression tag	UNP P27487

There are 14 discrepancies between the modelled and reference sequences:

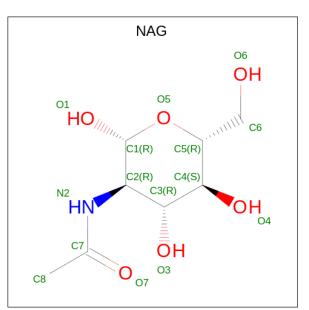
• Molecule 2 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-3)-alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.





Mol	Chain	Residues	Atoms		ZeroOcc	AltConf	Trace		
2	С	4	Total 47	C 26	N 1	O 20	0	0	0

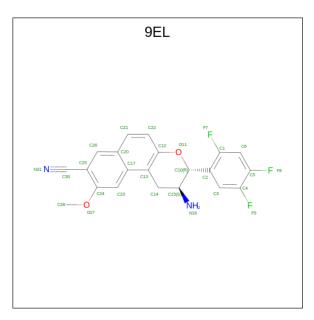
• Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	1	Total         C         N         O           14         8         1         5	0	0
3	В	1	Total         C         N         O           14         8         1         5	0	0

• Molecule 4 is (2S,3R)-2-amino-9-methoxy-3-(2,4,5-trifluorophenyl)-2,3-dihydro-1H-naphtho[2,1-b]pyran-8-carbonitrile (three-letter code: 9EL) (formula:  $C_{21}H_{15}F_3N_2O_2$ ).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
4	Δ	1	Total	С	F	Ν	Ο	0	0
4	Л	1	28	21	3	2	2	0	0
4	р	1	Total	С	F	Ν	Ο	0	0
4	D	1	28	21	3	2	2	0	0

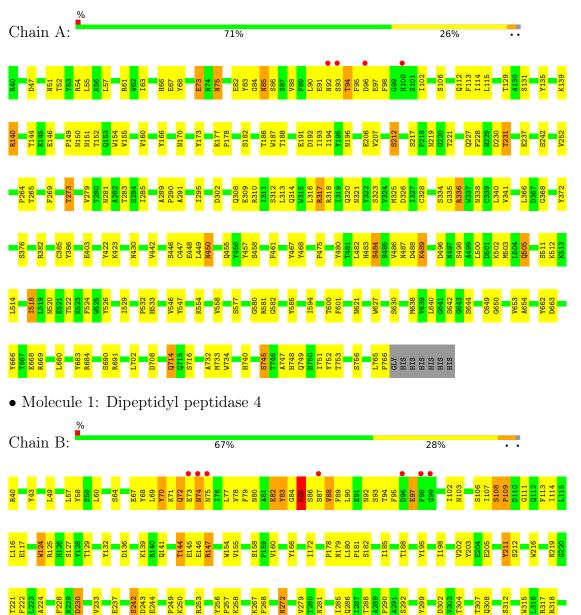
• Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	26	Total         O           26         26	0	0
5	В	17	Total         O           17         17	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: Dipeptidyl peptidase 4



# VT11 L657 L449 Y322 S720 1574 2438 325 V728 556 5458 332 V730 5583 5460 333 S744 H597 V465 343 V743 H597 V465 343 S744 H597 V465 333 S744 H597 V465 333 F601 H61 1470 1343 F601 R411 Y480 333 F618 R611 Y480 833 F618 R611 Y480 833 F618 R611 Y480 835 H15 R611 Y480 837 F618 R483 R343 835 F618 R483 R343 836 F618 F618 F438 837 F618 F618 F438 736 F618 F618 F438 736 F619

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

Chain C: 75% 25%

NAG 1 MAN2 MAN3 MAN3 MAN4



## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	66.07Å 126.19Å 118.62Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $99.30^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	85.82 - 3.00	Depositor
Resolution (A)	55.16 - 3.00	EDS
% Data completeness	91.2 (85.82-3.00)	Depositor
(in resolution range)	91.2 (55.16-3.00)	EDS
R <sub>merge</sub>	0.15	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.49 (at 3.01 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
D D.	0.191 , $0.253$	Depositor
$R, R_{free}$	0.201 , $0.252$	DCC
$R_{free}$ test set	1786 reflections $(5.08\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	39.1	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.36,54.1	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.48, < L^2>=0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	12088	wwPDB-VP
Average B, all atoms $(Å^2)$	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 3.81% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;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.



<sup>&</sup>lt;sup>1</sup>Intensities estimated from amplitudes.

# 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: MAN, NAG, 9EL

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		
	Unam	RMSZ	# Z  > 5	RMSZ	# Z  > 5	
1	А	0.73	0/6129	0.89	0/8336	
1	В	0.72	0/6129	0.89	0/8336	
All	All	0.72	0/12258	0.89	0/16672	

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	А	5957	0	5679	109	0
1	В	5957	0	5676	145	0
2	С	47	0	40	5	0
3	А	14	0	13	2	0
3	В	14	0	13	0	0
4	А	28	0	0	1	0
4	В	28	0	0	0	0
5	А	26	0	0	1	0
5	В	17	0	0	1	0
All	All	12088	0	11421	253	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 11.

The worst 5 of 253 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:91:GLU:HG2	1:A:94:THR:OG1	1.46	1.15
1:B:77:LEU:HD22	1:B:86:SER:HB3	1.44	0.97
1:B:397:ILE:HD12	1:B:439:TYR:OH	1.74	0.88
1:B:397:ILE:HD12	1:B:439:TYR:CZ	2.11	0.85
1:A:91:GLU:CG	1:A:94:THR:OG1	2.24	0.85

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	А	725/734~(99%)	646 (89%)	62~(9%)	17 (2%)	6 30
1	В	725/734~(99%)	621 (86%)	81 (11%)	23 (3%)	4 22
All	All	1450/1468~(99%)	1267 (87%)	143 (10%)	40 (3%)	5 25

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	85	ASN
1	А	93	SER
1	А	281	ASN
1	А	335	GLY
1	А	520	ASN



#### 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	Percer	ntiles
1	А	652/658~(99%)	614 (94%)	38~(6%)	20	55
1	В	652/658~(99%)	596 (91%)	56~(9%)	10	37
All	All	1304/1316~(99%)	1210~(93%)	94 (7%)	14	45

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	В	202	VAL
1	В	359	PRO
1	В	212	SER
1	В	279	VAL
1	В	448	GLU

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such side chains are listed below:

Mol	Chain	Res	Type
1	В	621	ASN
1	В	697	GLN
1	В	761	GLN
1	А	595	ASN
1	А	748	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)

4 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		Chain Res		Bo	ond leng	ths	В	ond ang	les
	туре	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z  > 2		
2	MAN	С	3	2	11,11,12	0.52	0	$15,\!15,\!17$	1.19	3 (20%)		
3	NAG	А	801	1	14,14,15	0.29	0	17,19,21	0.61	0		
3	NAG	В	801	1	14,14,15	0.58	0	17,19,21	1.83	4 (23%)		
2	MAN	С	2	2	11,11,12	0.95	0	$15,\!15,\!17$	1.72	2 (13%)		

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	MAN	С	3	2	-	2/2/19/22	0/1/1/1
3	NAG	А	801	1	-	4/6/23/26	0/1/1/1
3	NAG	В	801	1	-	2/6/23/26	0/1/1/1
2	MAN	С	2	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	С	2	MAN	C1-C2-C3	5.70	116.67	109.67
3	В	801	NAG	O5-C5-C6	4.56	114.35	107.20
3	В	801	NAG	C4-C3-C2	3.96	116.82	111.02
3	В	801	NAG	O5-C1-C2	-2.68	107.06	111.29
2	С	3	MAN	O5-C5-C6	2.58	111.24	107.20

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	801	NAG	C8-C7-N2-C2
3	А	801	NAG	O7-C7-N2-C2
3	В	801	NAG	O5-C5-C6-O6
3	А	801	NAG	O5-C5-C6-O6
2	С	2	MAN	C4-C5-C6-O6



There are no ring outliers.

1 monomer is involved in 2 short contacts:

$\mathbf{N}$	ſol	Chain	Res	Type	Clashes	Symm-Clashes
	3	А	801	NAG	2	0

#### 5.5 Carbohydrates (i)

4 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	Turne	Chain	Res	Link	Bo	ond leng	ths	В	ond ang	les
	Type	Chain	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
2	NAG	С	1	1,2	14,14,15	0.55	0	17,19,21	1.33	3 (17%)
2	MAN	С	2	2	11,11,12	0.95	0	$15,\!15,\!17$	1.72	2 (13%)
2	MAN	С	3	2	11,11,12	0.52	0	$15,\!15,\!17$	1.19	3 (20%)
2	MAN	С	4	2	11,11,12	0.62	0	$15,\!15,\!17$	1.19	3 (20%)

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	С	1	1,2	-	3/6/23/26	0/1/1/1
2	MAN	С	2	2	-	2/2/19/22	0/1/1/1
2	MAN	С	3	2	-	2/2/19/22	0/1/1/1
2	MAN	С	4	2	-	0/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
2	С	2	MAN	C1-C2-C3	5.70	116.67	109.67
2	С	3	MAN	O5-C5-C6	2.58	111.24	107.20

Continued on next page...



Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	С	3	MAN	C1-O5-C5	2.57	115.67	112.19
2	С	1	NAG	C8-C7-N2	2.56	120.44	116.10
2	С	4	MAN	C1-O5-C5	2.46	115.52	112.19

Continued from previous page...

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

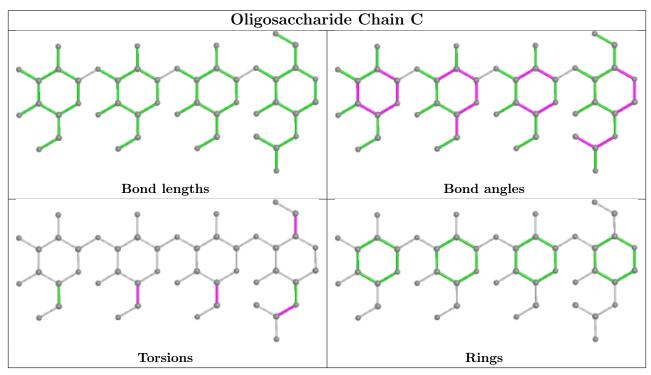
Mol	Chain	Res	Type	Atoms
2	С	2	MAN	C4-C5-C6-O6
2	С	1	NAG	C8-C7-N2-C2
2	С	1	NAG	O7-C7-N2-C2
2	С	2	MAN	O5-C5-C6-O6
2	С	3	MAN	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	С	1	NAG	5	0

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)

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	Turne	Chain	Res	Link	Bo	Bond lengths			ond ang	gles
10101	Type	Unam	nes	LIIIK	Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2
4	9EL	В	806	-	31,31,31	3.05	9 (29%)	43,46,46	1.83	7 (16%)
3	NAG	А	801	1	14,14,15	0.29	0	17,19,21	0.61	0
3	NAG	В	801	1	14,14,15	0.58	0	17,19,21	1.83	4 (23%)
4	9EL	А	802	-	31,31,31	2.99	8 (25%)	43,46,46	1.76	10 (23%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	9EL	В	806	-	-	1/8/20/20	0/4/4/4
3	NAG	А	801	1	-	4/6/23/26	0/1/1/1
3	NAG	В	801	1	-	2/6/23/26	0/1/1/1
4	9EL	А	802	-	-	1/8/20/20	0/4/4/4

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	$\mathrm{Ideal}(\mathrm{\AA})$
4	А	802	9EL	C25-C30	-12.92	1.24	1.44
4	В	806	9EL	C25-C30	-12.35	1.25	1.44
4	А	802	9EL	C12-C13	6.73	1.49	1.38
4	В	806	9EL	C12-C13	5.61	1.47	1.38
4	В	806	9EL	C17-C20	4.22	1.50	1.43

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
4	В	806	9EL	C3-C2-C1	5.69	121.21	116.48
4	В	806	9EL	C28-O27-C24	5.64	126.04	117.53

Continued on next page...



Mol	Chain	-	Type	Atoms	Ζ	$Observed(^{o})$	$Ideal(^{o})$
4	А	802	9EL	C13-C17-C20	-5.03	114.06	119.65
3	В	801	NAG	O5-C5-C6	4.56	114.35	107.20
4	В	806	9EL	C6-C1-C2	-4.23	118.77	123.83

Continued from previous page...

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	801	NAG	C8-C7-N2-C2
3	А	801	NAG	O7-C7-N2-C2
4	В	806	9EL	C24-C25-C30-N31
3	В	801	NAG	O5-C5-C6-O6
3	А	801	NAG	O5-C5-C6-O6

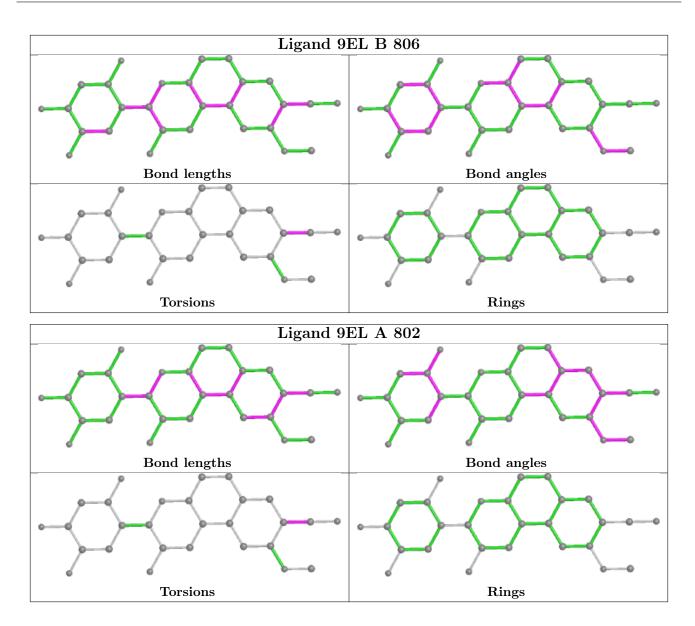
There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	А	801	NAG	2	0
4	А	802	9EL	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and sufficient 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,  $95^{th}$  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	А	727/734~(99%)	-0.37	4 (0%) 89 72	10, 27, 56, 92	2 (0%)
1	В	727/734~(99%)	-0.33	7 (0%) 82 59	11, 28, 62, 109	1 (0%)
All	All	1454/1468~(99%)	-0.35	11 (0%) 86 65	10, 27, 60, 109	3 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	В	87	SER	6.6
1	В	74	ASN	3.8
1	А	96	ASP	3.5
1	А	100	HIS	3.3
1	В	99	GLY	3.0

#### 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,  $95^{th}$  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	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
2	MAN	С	3	11/12	0.70	0.35	$54,\!56,\!57,\!59$	0
3	NAG	А	801	14/15	0.80	0.30	32,33,35,41	0
3	NAG	В	801	14/15	0.82	0.31	40,54,64,64	0
2	MAN	С	2	11/12	0.83	0.29	46,48,52,53	0

#### 6.3 Carbohydrates (i)

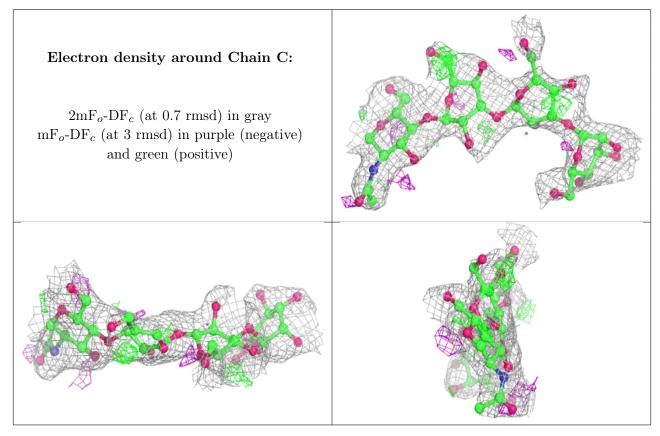
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum,



Mol	Type	Chain	Res	Atoms	RSCC	RSR	$B-factors(Å^2)$	Q<0.9
2	MAN	С	3	11/12	0.70	0.35	$54,\!56,\!57,\!59$	0
2	MAN	С	4	11/12	0.70	0.35	58,59,60,60	0
2	MAN	С	2	11/12	0.83	0.29	46,48,52,53	0
2	NAG	С	1	14/15	0.85	0.27	36,39,42,44	0

median,  $95^{th}$  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.

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.



#### Ligands (i) 6.4

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,  $95^{th}$  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	$\mathbf{RSR}$	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	А	801	14/15	0.80	0.30	$32,\!33,\!35,\!41$	0
3	NAG	В	801	14/15	0.82	0.31	40,54,64,64	0
4	9EL	В	806	28/28	0.97	0.17	19,21,23,25	0

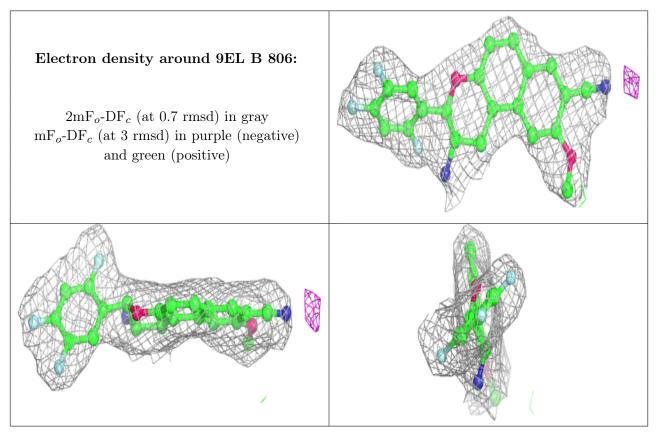
Continued on next page...



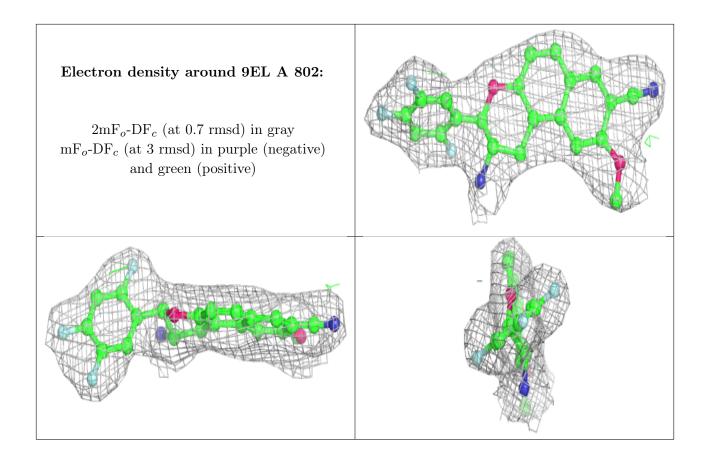
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-factors}(\mathrm{\AA}^2)$	Q<0.9
4	9EL	А	802	28/28	0.98	0.17	$16,\!19,\!25,\!30$	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.

