



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 9, 2024 – 10:12 AM EST

PDB ID : 5WKI  
Title : Crystal structure of PG90 TCR-CD1b-PG complex  
Authors : Shahine, A.; Gras, S.; Rossjohn, J.  
Deposited on : 2017-07-25  
Resolution : 2.75 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 types of validation reports are described at

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

---

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

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

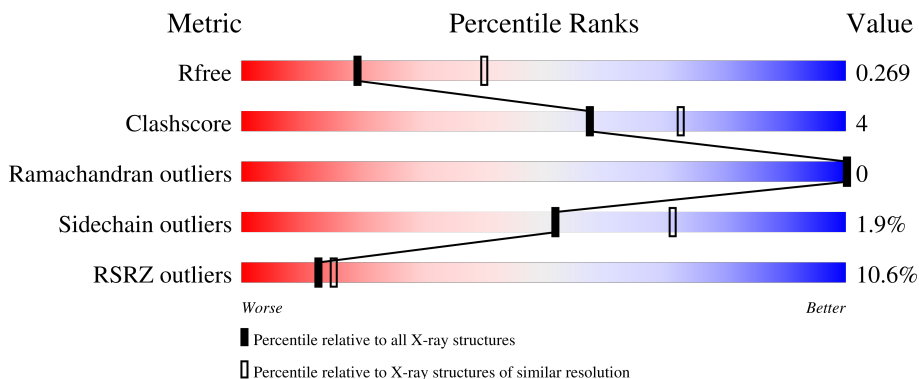
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1606 (2.78-2.74)
Clashscore	180529	1689 (2.78-2.74)
Ramachandran outliers	177936	1665 (2.78-2.74)
Sidechain outliers	177891	1665 (2.78-2.74)
RSRZ outliers	164620	1606 (2.78-2.74)

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	300	
2	B	99	
3	D	204	
4	E	249	
5	C	4	

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>
11	ACT	E	302	-	-	X	-

## 2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 6732 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 T-cell surface glycoprotein CD1b.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	267	2069	1328	351	380	10	0	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	279	GLY	-	expression tag	UNP P29016
A	280	SER	-	expression tag	UNP P29016
A	281	GLY	-	expression tag	UNP P29016
A	282	LEU	-	expression tag	UNP P29016
A	283	ASN	-	expression tag	UNP P29016
A	284	ASP	-	expression tag	UNP P29016
A	285	ILE	-	expression tag	UNP P29016
A	286	PHE	-	expression tag	UNP P29016
A	287	GLU	-	expression tag	UNP P29016
A	288	ALA	-	expression tag	UNP P29016
A	289	GLN	-	expression tag	UNP P29016
A	290	LYS	-	expression tag	UNP P29016
A	291	ILE	-	expression tag	UNP P29016
A	292	GLU	-	expression tag	UNP P29016
A	293	TRP	-	expression tag	UNP P29016
A	294	HIS	-	expression tag	UNP P29016
A	295	GLU	-	expression tag	UNP P29016
A	296	HIS	-	expression tag	UNP P29016
A	297	HIS	-	expression tag	UNP P29016
A	298	HIS	-	expression tag	UNP P29016
A	299	HIS	-	expression tag	UNP P29016
A	300	HIS	-	expression tag	UNP P29016
A	301	HIS	-	expression tag	UNP P29016

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	97	790	504	135	149	2	0	0	0

- Molecule 3 is a protein called T-cell receptor alpha variable 26-1, TRA@ protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	198	1539	958	263	308	10	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

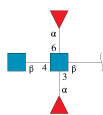
Chain	Residue	Modelled	Actual	Comment	Reference
D	8	PRO	THR	conflict	UNP A0A087WT03
D	91	ARG	-	linker	UNP A0A087WT03
D	92	VAL	-	linker	UNP A0A087WT03
D	93	ALA	-	linker	UNP A0A087WT03
D	94	TYR	-	linker	UNP A0A087WT03
D	95	ARG	-	linker	UNP A0A087WT03
D	96	GLN	-	linker	UNP A0A087WT03
D	97	LYS	-	linker	UNP A0A087WT03
D	98	VAL	-	linker	UNP A0A087WT03
D	99	THR	-	linker	UNP A0A087WT03
D	100	PHE	-	linker	UNP A0A087WT03
D	101	GLY	-	linker	UNP A0A087WT03
D	102	THR	-	linker	UNP A0A087WT03
D	103	GLY	-	linker	UNP A0A087WT03
D	104	THR	-	linker	UNP A0A087WT03
D	105	LYS	-	linker	UNP A0A087WT03
D	106	LEU	-	linker	UNP A0A087WT03
D	107	GLN	-	linker	UNP A0A087WT03
D	108	VAL	-	linker	UNP A0A087WT03
D	109	ILE	-	linker	UNP A0A087WT03
D	110	PRO	-	linker	UNP A0A087WT03
D	158	CYS	THR	conflict	UNP Q6P4G7

- Molecule 4 is a protein called PG90 TCR beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	248	1954	1229	345	375	5	0	0	0

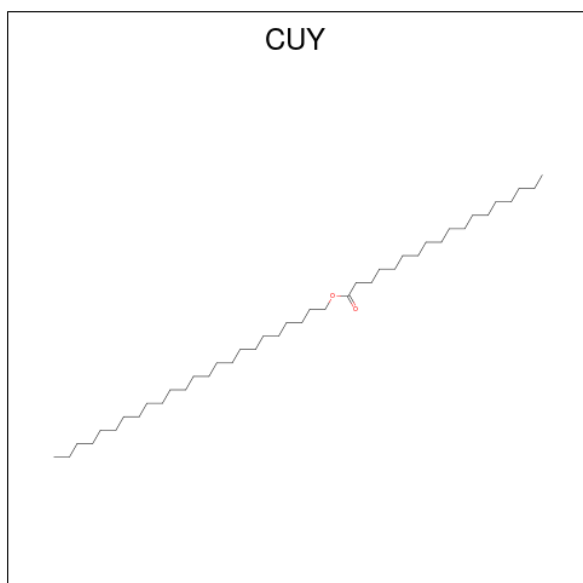
- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)][alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyr

anose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	C	4	48	28	2	18	13	0	0

- Molecule 6 is tetracosyl octadecanoate (three-letter code: CUY) (formula:  $C_{42}H_{84}O_2$ ).



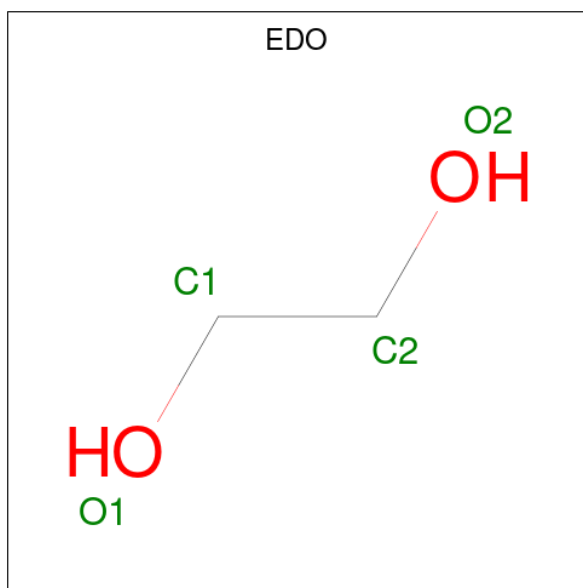
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	23	21	2	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



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

*Continued on next page...*

*Continued from previous page...*

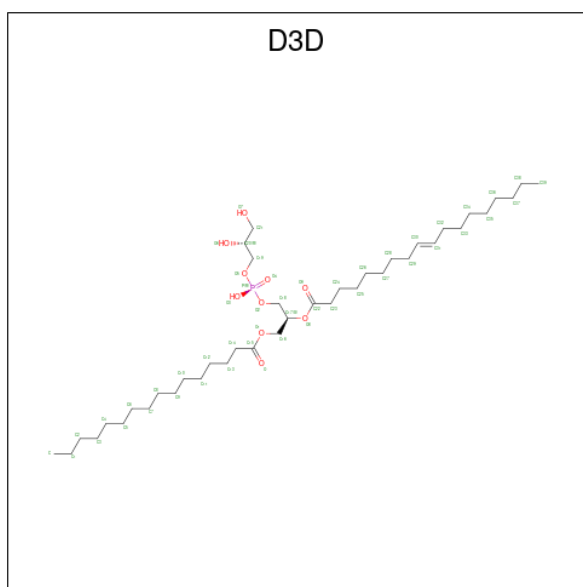
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 4 2 2	0	0
8	A	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0
8	D	1	Total C O 4 2 2	0	0

- Molecule 9 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	D	1	Total Na 1 1	0	0
9	E	1	Total Na 1 1	0	0

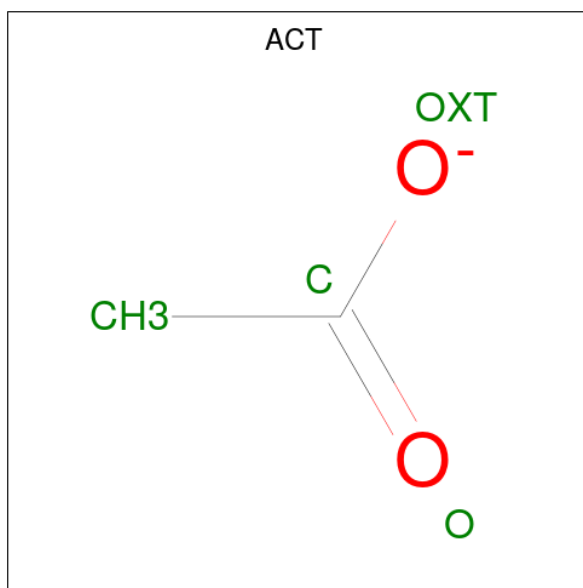
- Molecule 10 is (19S,22R,25R)-22,25,26-trihydroxy-16,22-dioxo-17,21,23-trioxa-22lambda a 5 -phosphahexacosan-19-yl (9E)-octadec-9-enoate (three-letter code: D3D) (formula: C<sub>40</sub>H<sub>77</sub>O<sub>10</sub>P).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	P		
10	D	1	49	38	10	1	0	0

- Molecule 11 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



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

- Molecule 12 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	E	1	Total Cl 1 1	0	0

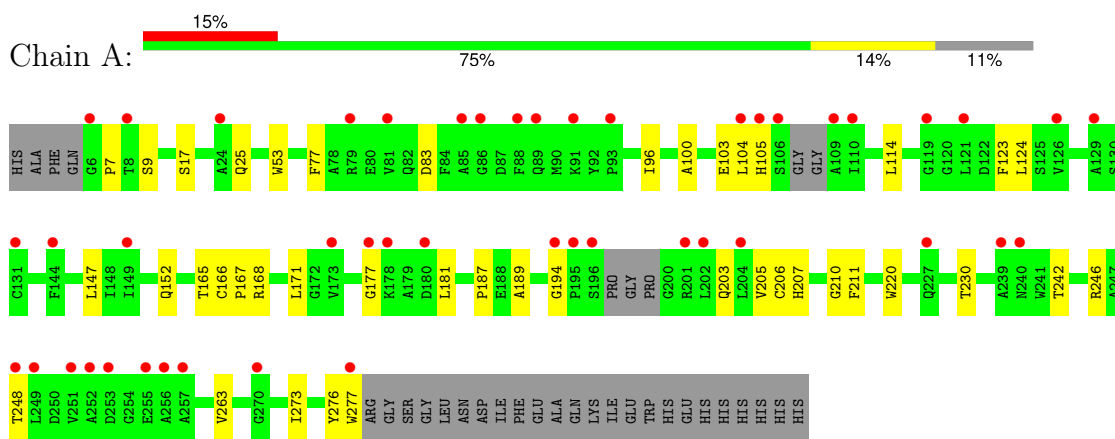
- Molecule 13 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	A	45	Total O 45 45	0	0
13	B	17	Total O 17 17	0	0
13	D	47	Total O 47 47	0	0
13	E	78	Total O 78 78	0	0

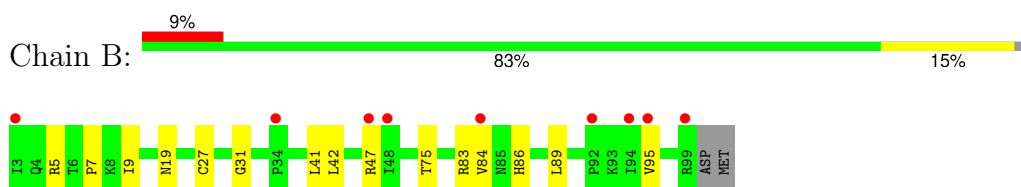
### 3 Residue-property plots [i](#)

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.

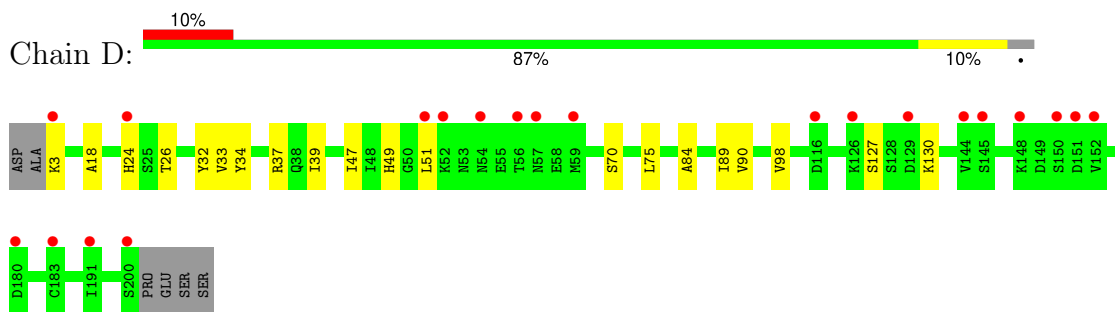
- Molecule 1: T-cell surface glycoprotein CD1b



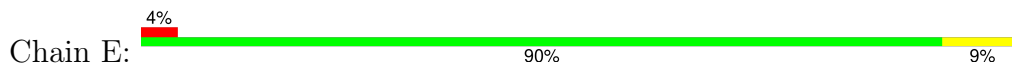
- Molecule 2: Beta-2-microglobulin



- Molecule 3: T-cell receptor alpha variable 26-1, TRA@ protein




- Molecule 4: PG90 TCR beta chain





- Molecule 5: alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)][alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:  25% 75%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.01Å 82.97Å 90.16Å 90.00° 94.88° 90.00°	Depositor
Resolution (Å)	75.73 – 2.75 75.73 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.8 (75.73-2.75) 99.8 (75.73-2.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.73Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.212 , 0.252 0.223 , 0.269	Depositor DCC
$R_{free}$ test set	1500 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.1	Xtrriage
Anisotropy	0.266	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 63.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6732	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: D3D, NAG, CL, CUY, NA, EDO, ACT, FUC

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	0.42	0/2122	0.63	0/2878
2	B	0.42	0/813	0.66	0/1106
3	D	0.46	0/1573	0.71	0/2144
4	E	0.43	0/2004	0.64	0/2724
All	All	0.43	0/6512	0.66	0/8852

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	A	2069	0	1988	24	0
2	B	790	0	730	10	0
3	D	1539	0	1451	11	0
4	E	1954	0	1861	12	0
5	C	48	0	43	0	0
6	A	23	0	0	0	0
7	A	14	0	13	0	0
8	A	20	0	30	1	0
8	B	4	0	6	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	D	24	0	36	0	0
9	D	1	0	0	0	0
9	E	1	0	0	0	0
10	D	49	0	0	0	0
11	E	8	0	6	3	0
12	E	1	0	0	0	0
13	A	45	0	0	1	0
13	B	17	0	0	0	0
13	D	47	0	0	0	0
13	E	78	0	0	1	0
All	All	6732	0	6164	56	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 4.

All (56) 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:100:ALA:HB2	1:A:114:LEU:HD23	1.61	0.80
1:A:187:PRO:HB3	1:A:211:PHE:HB3	1.69	0.74
1:A:104:LEU:HD21	13:A:540:HOH:O	1.94	0.68
2:B:7:PRO:HD2	2:B:89:LEU:HD21	1.81	0.63
3:D:37:ARG:HB2	3:D:47:ILE:HD11	1.83	0.60
1:A:276:TYR:O	1:A:277:TRP:HB2	2.01	0.59
4:E:248:ALA:O	4:E:249:ASP:HB2	2.00	0.59
1:A:220:TRP:H	1:A:230:THR:HG21	1.68	0.58
2:B:42:LEU:HD23	2:B:47:ARG:HA	1.85	0.56
3:D:39:ILE:HD13	3:D:84:ALA:HB2	1.92	0.52
1:A:206:CYS:O	1:A:246:ARG:HA	2.12	0.50
4:E:44:GLU:OE2	11:E:302:ACT:C	2.61	0.49
3:D:34:TYR:HB2	3:D:89:ILE:HG13	1.94	0.48
1:A:263:VAL:HB	1:A:273:ILE:HB	1.96	0.48
2:B:42:LEU:HD21	2:B:83:ARG:HH11	1.79	0.47
1:A:205:VAL:HG22	1:A:248:THR:HG22	1.97	0.47
1:A:220:TRP:HB2	1:A:230:THR:HG22	1.97	0.46
2:B:42:LEU:HD23	2:B:47:ARG:CA	2.45	0.46
3:D:3:LYS:HB3	3:D:24:HIS:CE1	2.50	0.46
1:A:165:THR:HG22	1:A:168:ARG:NH2	2.31	0.46
1:A:194:GLY:HA3	1:A:203:GLN:HB3	1.98	0.46
4:E:159:HIS:HB3	4:E:220:TYR:HB2	1.99	0.46
3:D:3:LYS:HB3	3:D:24:HIS:HE1	1.82	0.45

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:189:ALA:HA	1:A:207:HIS:O	2.15	0.45
1:A:7:PRO:HG2	1:A:181:LEU:HD22	1.97	0.45
3:D:32:TYR:CD1	3:D:51:LEU:HB3	2.52	0.45
4:E:44:GLU:OE2	11:E:302:ACT:CH3	2.65	0.45
4:E:46:LEU:HD22	4:E:61:PRO:HD3	1.97	0.45
2:B:5:ARG:HB3	2:B:31:GLY:O	2.17	0.45
1:A:166:CYS:HB3	1:A:167:PRO:HD3	1.98	0.45
1:A:17:SER:HB2	1:A:25:GLN:HG3	1.99	0.44
1:A:165:THR:HG22	1:A:168:ARG:HH21	1.82	0.44
2:B:27:CYS:HB2	2:B:41:LEU:HD21	1.99	0.44
3:D:89:ILE:HD13	3:D:98:VAL:HG13	2.00	0.44
1:A:100:ALA:HB2	1:A:114:LEU:CD2	2.41	0.44
4:E:29:HIS:HE1	4:E:109:PHE:CD1	2.36	0.44
2:B:86:HIS:HB3	2:B:89:LEU:HG	1.99	0.43
4:E:214:ARG:HD3	13:E:404:HOH:O	2.18	0.43
1:A:53:TRP:CD2	1:A:177:GLY:HA3	2.52	0.43
1:A:152:GLN:HB3	3:D:51:LEU:HD21	2.00	0.43
4:E:14:LYS:HB3	4:E:17:GLN:HG3	2.00	0.43
2:B:19:ASN:OD1	2:B:75:THR:HA	2.17	0.43
2:B:9:ILE:HG12	2:B:84:VAL:HG21	2.01	0.42
1:A:207:HIS:HE1	8:A:407:EDO:C2	2.32	0.42
4:E:125:VAL:HG12	4:E:235:PRO:HB2	2.00	0.42
1:A:9:SER:HB3	1:A:103:GLU:HG3	2.02	0.42
4:E:30:VAL:HB	4:E:97:ALA:HB3	2.00	0.42
3:D:18:ALA:HB3	3:D:75:LEU:HB2	2.01	0.42
1:A:210:GLY:HA2	1:A:242:THR:HB	2.03	0.41
3:D:33:VAL:O	3:D:49:HIS:HA	2.21	0.41
4:E:44:GLU:OE2	11:E:302:ACT:H2	2.20	0.41
1:A:77:PHE:HZ	1:A:96:ILE:HG21	1.86	0.41
3:D:127:SER:OG	3:D:130:LYS:HG3	2.21	0.41
2:B:9:ILE:HB	2:B:95:VAL:HG21	2.03	0.41
1:A:123:PHE:HD1	1:A:124:LEU:HD23	1.86	0.40
4:E:162:LEU:HG	4:E:217:VAL:HG22	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	261/300 (87%)	252 (97%)	9 (3%)	0	100	100
2	B	95/99 (96%)	93 (98%)	2 (2%)	0	100	100
3	D	196/204 (96%)	187 (95%)	9 (5%)	0	100	100
4	E	246/249 (99%)	237 (96%)	9 (4%)	0	100	100
All	All	798/852 (94%)	769 (96%)	29 (4%)	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	215/247 (87%)	211 (98%)	4 (2%)	52	71
2	B	86/94 (92%)	86 (100%)	0	100	100
3	D	174/184 (95%)	171 (98%)	3 (2%)	56	74
4	E	209/210 (100%)	203 (97%)	6 (3%)	37	59
All	All	684/735 (93%)	671 (98%)	13 (2%)	52	71

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	83	ASP
1	A	105	HIS

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
1	A	147	LEU
1	A	171	LEU
3	D	26	THR
3	D	70	SER
3	D	90	VAL
4	E	63	ASP
4	E	80	GLN
4	E	148	LEU
4	E	176	CYS
4	E	198	ARG
4	E	249	ASP

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

### 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](#)

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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NAG	C	1	5,1	14,14,15	0.39	0	17,19,21	1.40	3 (17%)
5	FUC	C	2	5	10,10,11	0.44	0	14,14,16	0.59	0
5	NAG	C	3	5	14,14,15	0.36	0	17,19,21	1.16	2 (11%)
5	FUC	C	4	5	10,10,11	0.43	0	14,14,16	1.01	1 (7%)

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
5	NAG	C	1	5,1	-	5/6/23/26	0/1/1/1
5	FUC	C	2	5	-	-	0/1/1/1
5	NAG	C	3	5	-	3/6/23/26	0/1/1/1
5	FUC	C	4	5	-	-	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	1	NAG	O5-C1-C2	-3.41	106.02	111.29
5	C	3	NAG	C1-C2-N2	3.21	115.48	110.43
5	C	3	NAG	C2-N2-C7	2.56	126.33	122.90
5	C	1	NAG	C1-O5-C5	2.47	115.49	112.19
5	C	1	NAG	C2-N2-C7	2.17	125.80	122.90
5	C	4	FUC	C3-C4-C5	2.11	113.02	109.81

There are no chirality outliers.

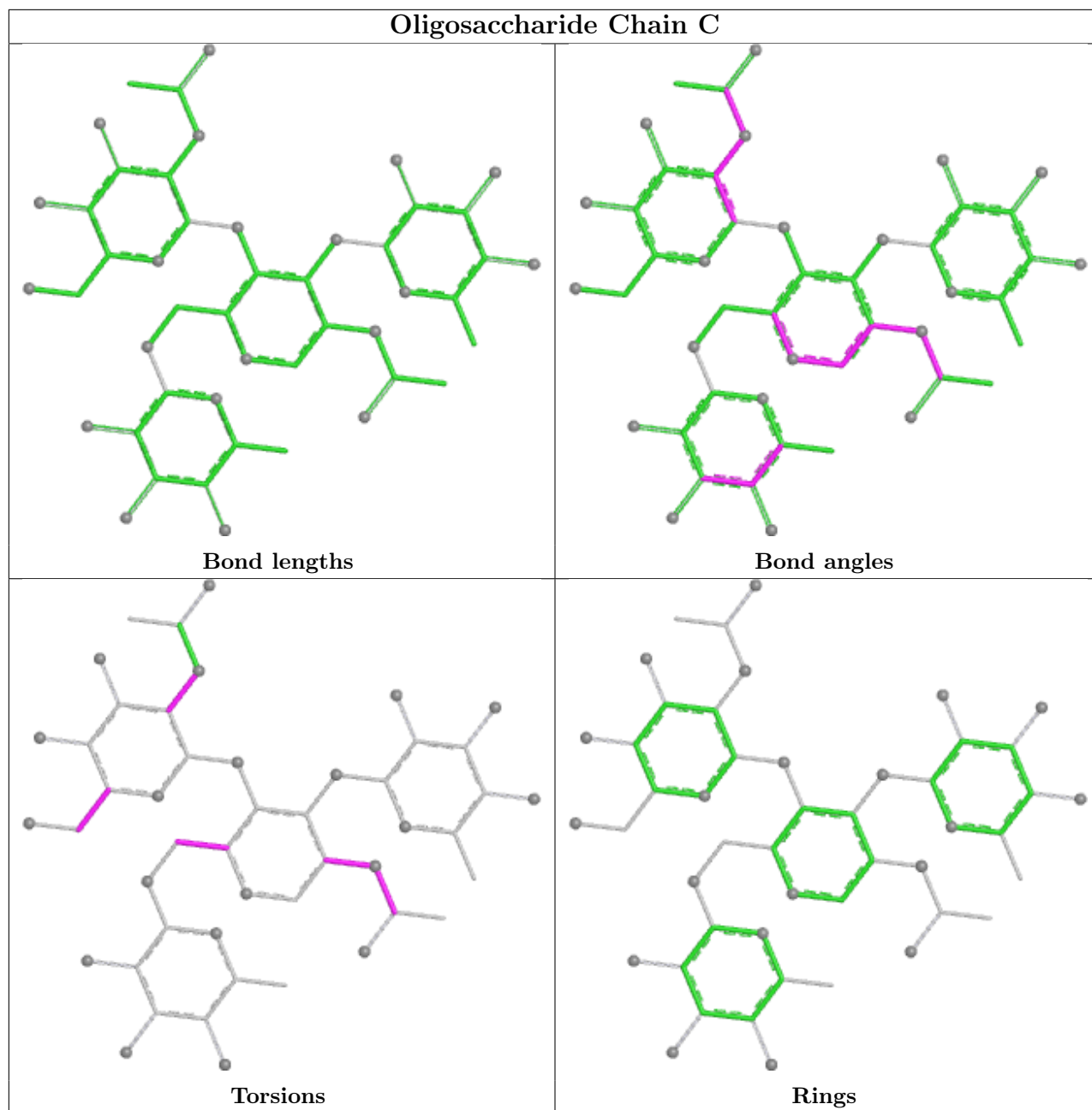
All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1	NAG	C1-C2-N2-C7
5	C	3	NAG	C1-C2-N2-C7
5	C	1	NAG	O5-C5-C6-O6
5	C	1	NAG	C4-C5-C6-O6
5	C	1	NAG	C8-C7-N2-C2
5	C	3	NAG	O5-C5-C6-O6
5	C	1	NAG	O7-C7-N2-C2
5	C	3	NAG	C3-C2-N2-C7

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 oligosaccharide.



## 5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 3 are monoatomic - leaving 17 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
8	EDO	D	303	-	3,3,3	0.55	0	2,2,2	0.19	0
8	EDO	B	201	-	3,3,3	0.55	0	2,2,2	0.24	0
8	EDO	A	408	-	3,3,3	0.49	0	2,2,2	0.38	0
8	EDO	A	407	-	3,3,3	0.52	0	2,2,2	0.46	0
8	EDO	A	410	-	3,3,3	0.61	0	2,2,2	0.12	0
8	EDO	A	411	-	3,3,3	0.57	0	2,2,2	0.17	0
11	ACT	E	301	-	3,3,3	1.10	0	3,3,3	0.92	0
8	EDO	A	409	-	3,3,3	0.56	0	2,2,2	0.18	0
7	NAG	A	402	1	14,14,15	0.41	0	17,19,21	1.08	1 (5%)
8	EDO	D	304	-	3,3,3	0.55	0	2,2,2	0.28	0
8	EDO	D	305	-	3,3,3	0.49	0	2,2,2	0.40	0
8	EDO	D	306	-	3,3,3	0.58	0	2,2,2	0.23	0
8	EDO	D	302	-	3,3,3	0.53	0	2,2,2	0.33	0
6	CUY	A	401	-	22,22,43	0.83	1 (4%)	22,22,43	0.99	1 (4%)
8	EDO	D	301	-	3,3,3	0.54	0	2,2,2	0.28	0
11	ACT	E	302	-	3,3,3	1.11	0	3,3,3	1.16	0
10	D3D	D	308	-	48,48,50	0.95	3 (6%)	51,54,56	1.12	2 (3%)

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
8	EDO	D	303	-	-	0/1/1/1	-
8	EDO	B	201	-	-	0/1/1/1	-
8	EDO	A	408	-	-	0/1/1/1	-
8	EDO	A	407	-	-	1/1/1/1	-
8	EDO	A	410	-	-	1/1/1/1	-
8	EDO	A	411	-	-	1/1/1/1	-
8	EDO	A	409	-	-	1/1/1/1	-
7	NAG	A	402	1	-	1/6/23/26	0/1/1/1
8	EDO	D	304	-	-	1/1/1/1	-
8	EDO	D	305	-	-	0/1/1/1	-
8	EDO	D	306	-	-	1/1/1/1	-
8	EDO	D	302	-	-	0/1/1/1	-
6	CUY	A	401	-	-	10/21/21/42	-
8	EDO	D	301	-	-	1/1/1/1	-
10	D3D	D	308	-	-	21/53/53/55	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
10	D	308	D3D	O1-C15	2.98	1.42	1.33
6	A	401	CUY	OAQ-CAP	2.87	1.41	1.33
10	D	308	D3D	O8-C22	2.49	1.41	1.34
10	D	308	D3D	O8-C17	-2.20	1.41	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	D	308	D3D	O8-C22-C23	3.84	119.78	111.48
7	A	402	NAG	C1-O5-C5	3.53	116.92	112.19
10	D	308	D3D	P-O5-C19	-2.77	105.45	121.35
6	A	401	CUY	OAQ-CAP-CAO	2.64	119.89	111.83

There are no chirality outliers.

All (39) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
10	D	308	D3D	C23-C22-O8-C17
10	D	308	D3D	O9-C22-O8-C17
10	D	308	D3D	O5-C19-C20-C21
10	D	308	D3D	O5-C19-C20-O6
10	D	308	D3D	C2-C3-C4-C5
10	D	308	D3D	C24-C25-C26-C27
6	A	401	CUY	OAQ-CAR-CAS-CAT
10	D	308	D3D	C25-C26-C27-C28
6	A	401	CUY	CBK-CBL-CBM-CAW
10	D	308	D3D	C11-C10-C9-C8
6	A	401	CUY	CAV-CAW-CBM-CBL
7	A	402	NAG	O5-C5-C6-O6
8	A	407	EDO	O1-C1-C2-O2
8	D	301	EDO	O1-C1-C2-O2
6	A	401	CUY	CAT-CAU-CAV-CAW
6	A	401	CUY	CBF-CBG-CBH-CBI
10	D	308	D3D	C9-C10-C11-C12
10	D	308	D3D	C7-C8-C9-C10
8	D	304	EDO	O1-C1-C2-O2
10	D	308	D3D	C4-C5-C6-C7
10	D	308	D3D	C10-C11-C12-C13
10	D	308	D3D	C34-C35-C36-C37
6	A	401	CUY	CBG-CBH-CBI-CBJ
6	A	401	CUY	CAL-CAM-CAN-CAO
6	A	401	CUY	CAK-CAL-CAM-CAN

*Continued on next page...*

Continued from previous page...

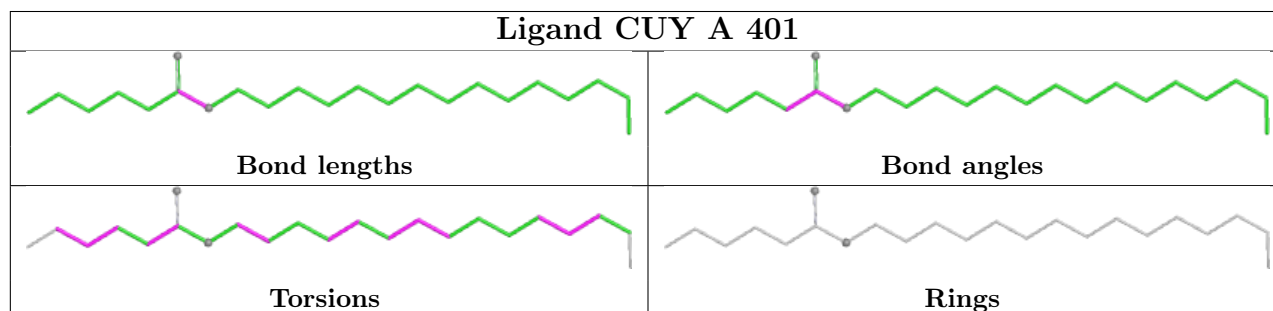
Mol	Chain	Res	Type	Atoms
10	D	308	D3D	C16-C17-O8-C22
8	A	410	EDO	O1-C1-C2-O2
10	D	308	D3D	C6-C7-C8-C9
8	A	409	EDO	O1-C1-C2-O2
8	A	411	EDO	O1-C1-C2-O2
10	D	308	D3D	O1-C16-C17-O8
10	D	308	D3D	C26-C27-C28-C29
10	D	308	D3D	C28-C29-C30-C31
10	D	308	D3D	O1-C16-C17-C18
8	D	306	EDO	O1-C1-C2-O2
10	D	308	D3D	C11-C12-C13-C14
6	A	401	CUY	CAN-CAO-CAP-OAQ
10	D	308	D3D	C18-C17-O8-C22
6	A	401	CUY	CAN-CAO-CAP-OBO

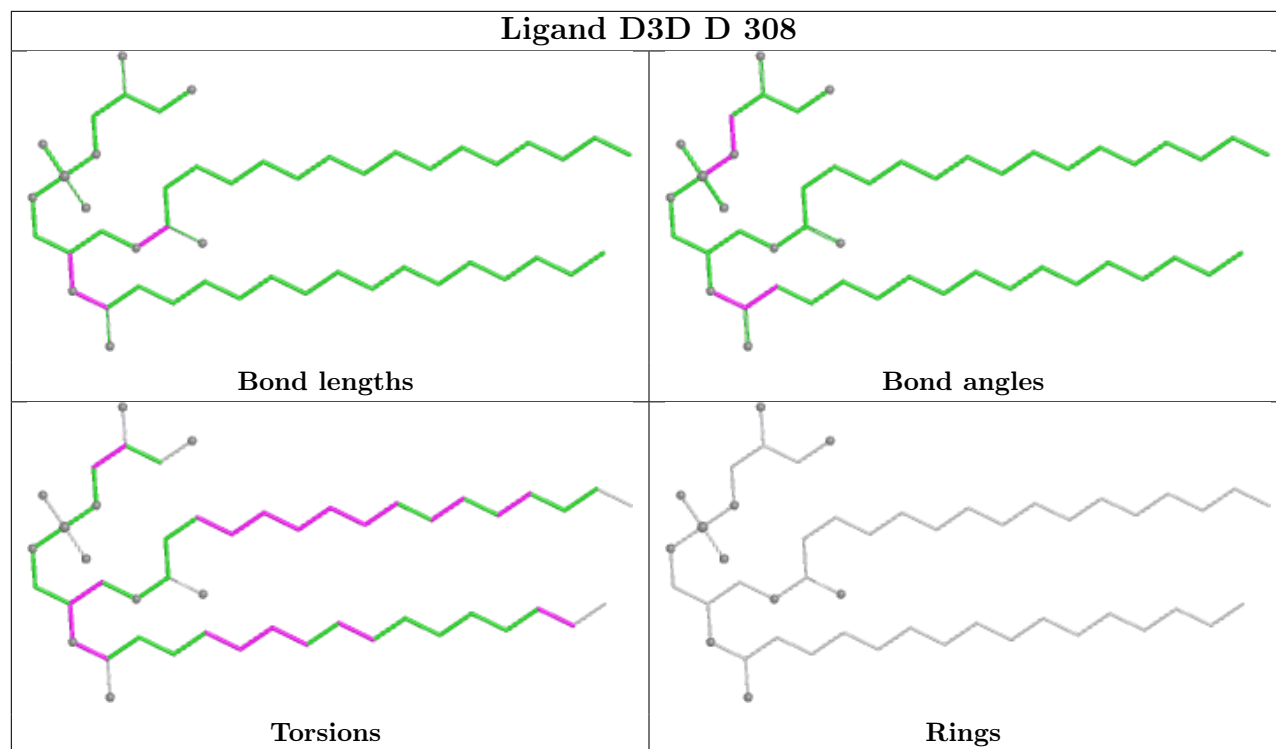
There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	A	407	EDO	1	0
11	E	302	ACT	3	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	267/300 (89%)	0.99	46 (17%) <b>5</b> <b>6</b>	31, 58, 100, 204	0
2	B	97/99 (97%)	0.99	9 (9%) <b>16</b> <b>18</b>	35, 62, 86, 104	0
3	D	198/204 (97%)	0.50	21 (10%) <b>13</b> <b>15</b>	28, 50, 86, 102	0
4	E	248/249 (99%)	0.03	10 (4%) <b>43</b> <b>44</b>	19, 39, 78, 94	0
All	All	810/852 (95%)	0.58	86 (10%) <b>13</b> <b>15</b>	19, 50, 88, 204	0

All (86) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	106	SER	5.8
1	A	239	ALA	5.6
1	A	252	ALA	4.2
1	A	104	LEU	3.6
1	A	277	TRP	3.6
4	E	187	ALA	3.5
2	B	3	ILE	3.5
1	A	8	THR	3.5
2	B	99	ARG	3.5
1	A	129	ALA	3.4
1	A	204	LEU	3.3
1	A	196	SER	3.3
3	D	183	CYS	3.3
1	A	251	VAL	3.3
3	D	116	ASP	3.2
1	A	91	LYS	3.1
1	A	240	ASN	3.1
1	A	88	PHE	3.0
1	A	194	GLY	3.0
1	A	253	ASP	3.0
4	E	24	ASP	2.9

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
3	D	52	LYS	2.8
3	D	200	SER	2.8
1	A	202	LEU	2.8
4	E	63	ASP	2.7
1	A	119	GLY	2.7
3	D	148	LYS	2.6
1	A	180	ASP	2.6
3	D	54	ASN	2.6
2	B	94	ILE	2.6
1	A	249	LEU	2.6
1	A	105	HIS	2.6
1	A	121	LEU	2.6
1	A	24	ALA	2.6
1	A	110	ILE	2.6
4	E	249	ASP	2.5
1	A	109	ALA	2.5
1	A	6	GLY	2.5
1	A	86	GLY	2.5
2	B	48	ILE	2.5
1	A	173	VAL	2.5
1	A	144	PHE	2.5
1	A	270	GLY	2.4
1	A	178	LYS	2.4
1	A	195	PRO	2.4
3	D	24	HIS	2.4
1	A	131	CYS	2.3
3	D	151	ASP	2.3
1	A	89	GLN	2.3
1	A	149	ILE	2.3
1	A	256	ALA	2.3
2	B	84	VAL	2.3
1	A	257	ALA	2.3
1	A	85	ALA	2.3
4	E	248	ALA	2.3
3	D	126	LYS	2.2
3	D	150	SER	2.3
4	E	225	ASN	2.2
2	B	34	PRO	2.2
3	D	191	ILE	2.2
1	A	201	ARG	2.2
2	B	47	ARG	2.2
1	A	93	PRO	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	D	56	THR	2.2
2	B	95	VAL	2.2
3	D	59	MET	2.2
1	A	79	ARG	2.2
3	D	145	SER	2.2
1	A	248	THR	2.2
1	A	227	GLN	2.2
3	D	144	VAL	2.2
3	D	129	ASP	2.1
3	D	152	VAL	2.1
4	E	180	GLN	2.1
3	D	51	LEU	2.1
4	E	188	LEU	2.1
2	B	92	PRO	2.1
3	D	180	ASP	2.1
4	E	247	ARG	2.1
1	A	126	VAL	2.0
1	A	177	GLY	2.0
1	A	255	GLU	2.0
3	D	3	LYS	2.0
1	A	81	VAL	2.0
3	D	57	ASN	2.0
4	E	189	ASN	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](#)

SUGAR-RSR INFOmissingINFO

## 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.

*Continued on next page...*

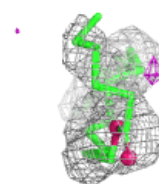
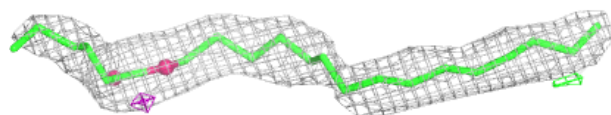
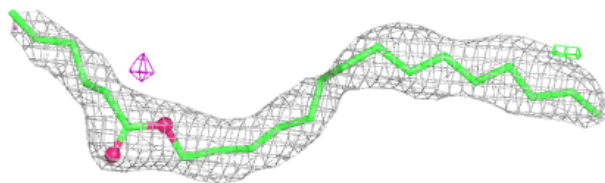
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
8	EDO	D	301	4/4	0.33	0.28	76,76,77,77	0
8	EDO	D	306	4/4	0.60	0.27	67,67,67,67	0
8	EDO	B	201	4/4	0.62	0.27	59,60,60,60	0
7	NAG	A	402	14/15	0.65	0.20	97,98,99,99	0
8	EDO	A	411	4/4	0.69	0.17	75,75,76,76	0
8	EDO	A	410	4/4	0.70	0.18	74,74,75,75	0
8	EDO	D	302	4/4	0.72	0.17	62,63,64,65	0
8	EDO	A	407	4/4	0.75	0.24	46,46,47,47	0
8	EDO	D	305	4/4	0.78	0.20	71,71,72,72	0
8	EDO	D	303	4/4	0.79	0.20	63,63,64,65	0
11	ACT	E	302	4/4	0.83	0.20	55,56,56,57	0
8	EDO	A	409	4/4	0.84	0.18	56,57,58,59	0
9	NA	D	307	1/1	0.88	0.36	63,63,63,63	0
8	EDO	D	304	4/4	0.88	0.14	61,62,63,64	0
11	ACT	E	301	4/4	0.90	0.17	40,40,40,40	0
8	EDO	A	408	4/4	0.90	0.13	69,70,71,72	0
6	CUY	A	401	23/44	0.91	0.15	43,47,49,50	0
10	D3D	D	308	49/51	0.91	0.13	32,36,41,48	0
12	CL	E	304	1/1	0.93	0.08	57,57,57,57	0
9	NA	E	303	1/1	0.94	0.22	62,62,62,62	0

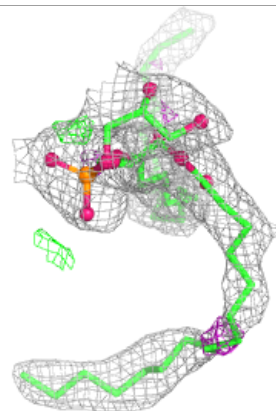
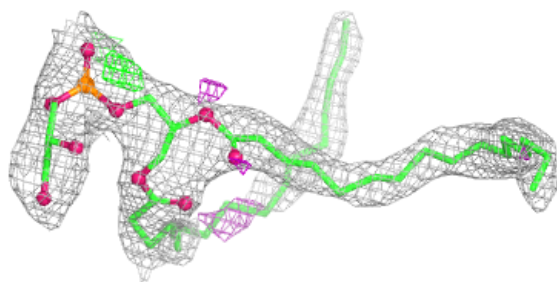
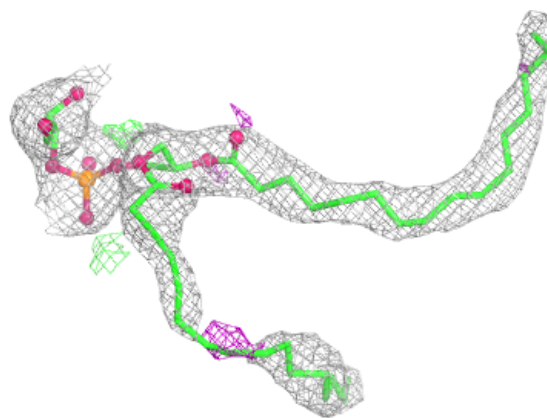
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around CUY A 401:**

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

**Electron density around D3D D 308:**

$2mF_o-DF_c$  (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.