W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 19, 2023 – 12:58 PM EST

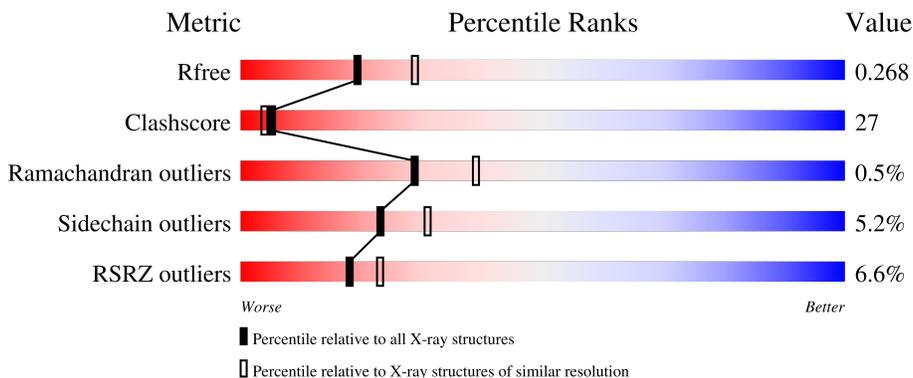|		
PDB ID	:	4EN3
Title	:	Crystal structure of a human Valpha24(-) NKT TCR in complex with CD1d/ alpha-galactosylceramide
Authors	:	Lopez-Sagaseta, J.; Adams, E.J.
Deposited on	:	2012-04-12
Resolution	:	2.57 Å(reported)

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 2.57 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	Percentile Ranks	Value
Rfree		0.268
Clashscore		27
Ramachandran outliers		0.5%
Sidechain outliers		5.2%
RSRZ outliers		6.6%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1279 (2.58-2.54)
Clashscore	141614	1327 (2.58-2.54)
Ramachandran outliers	138981	1312 (2.58-2.54)
Sidechain outliers	138945	1312 (2.58-2.54)
RSRZ outliers	127900	1269 (2.58-2.54)

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	C	284	5% · 56% · 33% · 10%
2	A	220	4% · 47% · 33% · 15%
3	B	259	8% · 58% · 33% · 7%
4	D	100	8% · 52% · 43% · ·

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:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	A	302	-	-	X	-
5	NAG	C	301	-	-	X	-
8	GOL	B	303	-	-	X	-

2 Entry composition ⓘ

There are 9 unique types of molecules in this entry. The entry contains 6389 atoms, of which 8 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 Antigen-presenting glycoprotein CD1d.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	257	Total	C	N	O	S	0	0	0
			2056	1323	351	375	7			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	ALA	-	expression tag	UNP P15813
C	1	ASP	-	expression tag	UNP P15813
C	2	PRO	-	expression tag	UNP P15813
C	278	HIS	-	expression tag	UNP P15813
C	279	HIS	-	expression tag	UNP P15813
C	280	HIS	-	expression tag	UNP P15813
C	281	HIS	-	expression tag	UNP P15813
C	282	HIS	-	expression tag	UNP P15813
C	283	HIS	-	expression tag	UNP P15813

- Molecule 2 is a protein called HUMAN NKT TCR ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	186	Total	C	N	O	S	0	0	0
			1369	858	231	273	7			

- Molecule 3 is a protein called HUMAN NKT TCR BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	240	Total	C	N	O	S	0	0	0
			1892	1193	326	365	8			

- Molecule 4 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	99	Total 811	C 517	N 135	O 157	S 2	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	PRO	-	expression tag	UNP P61769

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
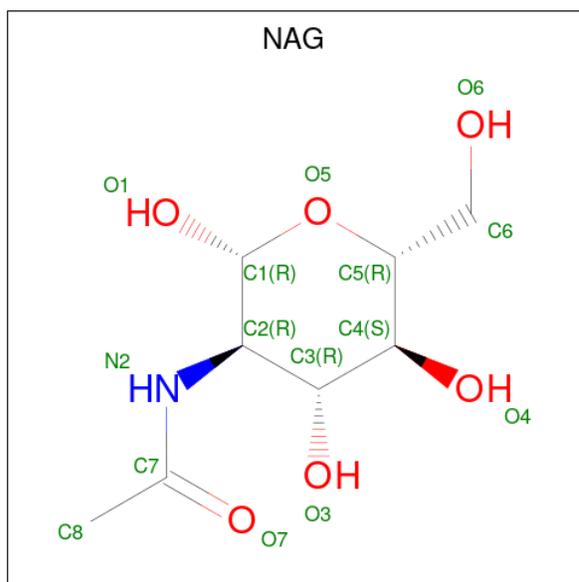
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total 14	C 8	N 1	O 5	0	0
5	C	1	Total 14	C 8	N 1	O 5	0	0
5	C	1	Total 14	C 8	N 1	O 5	0	0
5	A	1	Total 14	C 8	N 1	O 5	0	0
5	B	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 6 is N-{(1S,2R,3S)-1-[(ALPHA-D-GALACTOPYRANOSYLOXY)METHYL]-2,3-DIHYDROXYHEPTADECYL}HEXACOSANAMIDE (three-letter code: AGH) (formula: $C_{50}H_{99}NO_9$).

AGH
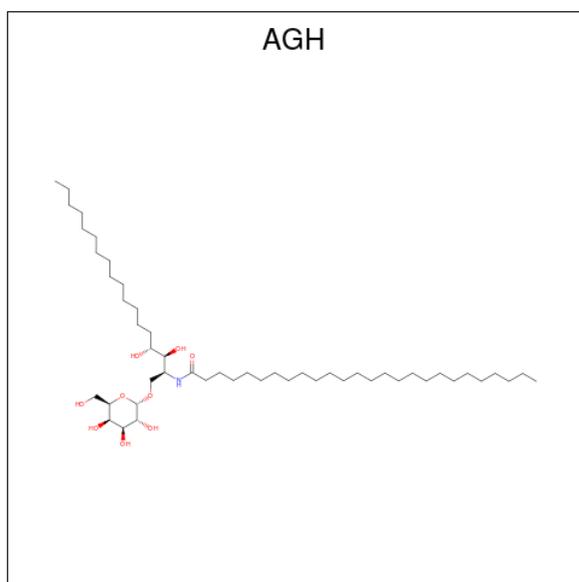
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	C	1	60	50	1	9	0	0

- Molecule 7 is alpha-L-fucopyranose (three-letter code: FUC) (formula: $C_6H_{12}O_5$).

FUC

O5

C6

O1

OH

C5(S)　C1(R)

C4(S)　C2(S)

C3(R)

HO

OH

O4　　　　　　O2

OH

O3
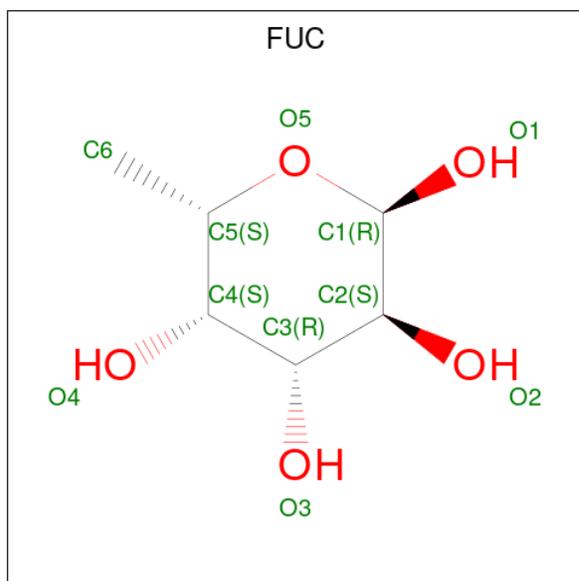
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
7	A	1	10	6	4	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

WORLDWIDE PDB
PROTEIN DATA BANK

GOL

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

• Molecule 9 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	33	Total 35	H 2	O 33	0	0
9	A	20	Total 22	H 2	O 20	0	0
9	B	26	Total 30	H 4	O 26	0	0
9	D	22	Total 22		O 22	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: Antigen-presenting glycoprotein CD1d

Chain C:

• Molecule 2: HUMAN NKT TCR ALPHA CHAIN

Chain A:

• Molecule 3: HUMAN NKT TCR BETA CHAIN

Chain B:

L76 T77 L78 A81 R82 P83 S84 H85 A92 R101 E104 F107 G108 P109 V115 D118 L119 K120 N121 V122 P125 E126 V127 A128 V129 F130 E131 P132 S133 I137 S138 Q141 K142 A143 T144 L145 V146 G151 F152 Y153 P154 V157 E158 L159 S160 W161 W162

V168 H169 S170 G171 V172 C173 P176 L179 K180 E181 R189 Y190 S194 R195 L196 R197 A200 T201 F202 W203 Q204 N208 H209 F210 F216 L219 E224 R229 A230 K231 P232 V233 T234 V237 S238 A239 E240 A241 W242 GLY ARG ALA ASP SER ARG GLY GLY LEU GLU

VAL LEU PHE GLN

● Molecule 4: Beta-2-microglobulin

Chain D:

8%

52% 43%

PRO I1 T4 P5 V9 Y10 S11 R12 H13 P14 A15 E16 K19 S20 N21 F22 L23 H31 P32 S33 E36 V37 L40 R45 I46 E47 K48 V49 E50 H51 S52 D53 L54 F62 Y63 L64 L65 Y66 Y67 T68 E69 F70 T71 P72 T73 E74 N83 H84 V85 T86 L87

S88 Q89 P90 K91 I92 V93 K94 W95 D96 R97 D98 M99

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.11Å　72.57Å　113.75Å 90.00°　103.31°　90.00°	Depositor
Resolution (Å)	44.01　–　2.57 44.01　–　2.57	Depositor EDS
% Data completeness (in resolution range)	99.1 (44.01-2.57) 93.0 (44.01-2.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.20 (at 2.58Å)	Xtriage
Refinement program	PHENIX 1.8_1069	Depositor
R, R_{free}	0.208　,　0.266 0.210　,　0.268	Depositor DCC
R_{free} test set	1471 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	49.6	Xtriage
Anisotropy	0.703	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 60.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6389	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.47% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: AGH, FUC, GOL, NAG

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	C	0.21	0/2116	0.37	0/2882
2	A	0.21	0/1392	0.37	0/1896
3	B	0.21	0/1944	0.39	0/2651
4	D	0.21	0/834	0.37	0/1134
All	All	0.21	0/6286	0.38	0/8563

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	C	2056	0	1972	104	0
2	A	1369	0	1218	100	0
3	B	1892	0	1759	104	0
4	D	811	0	754	41	0
5	A	14	0	13	9	0
5	B	14	0	13	3	0
5	C	42	0	39	12	0
6	C	60	0	99	8	0
7	A	10	0	10	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	12	0	16	6	0
9	A	20	2	0	1	0
9	B	26	4	0	1	0
9	C	33	2	0	4	0
9	D	22	0	0	1	0
All	All	6381	8	5893	325	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 27.

The worst 5 of 325 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)
3:B:137:ILE:HG23	3:B:200:ALA:HB1	1.41	1.03
1:C:42:ASN:HD21	5:C:301:NAG:C1	1.77	0.97
3:B:41:MET:HA	3:B:41:MET:HE2	1.51	0.93
1:C:154:THR:HG23	6:C:304:AGH:HAB2	1.50	0.92
2:A:20:MET:HE1	2:A:108:LEU:HB2	1.48	0.91

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	C	247/284 (87%)	240 (97%)	6 (2%)	1 (0%)	34	45
2	A	176/220 (80%)	165 (94%)	10 (6%)	1 (1%)	25	33
3	B	236/259 (91%)	220 (93%)	16 (7%)	0	100	100
4	D	97/100 (97%)	91 (94%)	4 (4%)	2 (2%)	7	8
All	All	756/863 (88%)	716 (95%)	36 (5%)	4 (0%)	29	39

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	48	LYS
1	C	269	GLU
2	A	164	MET
4	D	47	GLU

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	224/249 (90%)	215 (96%)	9 (4%)	31	42
2	A	137/195 (70%)	128 (93%)	9 (7%)	16	21
3	B	203/227 (89%)	193 (95%)	10 (5%)	25	34
4	D	90/95 (95%)	84 (93%)	6 (7%)	16	21
All	All	654/766 (85%)	620 (95%)	34 (5%)	23	31

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	64	LEU
4	D	69	GLU
4	D	94	LYS
2	A	88	CYS
2	A	60	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	20	ASN
1	C	42	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
5	NAG	C	301	-	14,14,15	0.58	0	17,19,21	1.25	1 (5%)
6	AGH	C	304	-	60,60,60	0.33	0	65,69,69	0.79	1 (1%)
5	NAG	B	302	-	14,14,15	0.60	0	17,19,21	0.72	0
5	NAG	C	302	-	14,14,15	0.45	0	17,19,21	0.77	0
5	NAG	C	303	-	14,14,15	0.60	0	17,19,21	0.92	1 (5%)
5	NAG	A	302	-	14,14,15	0.60	0	17,19,21	1.65	3 (17%)
8	GOL	B	301	-	5,5,5	0.36	0	5,5,5	0.33	0
8	GOL	B	303	-	5,5,5	0.36	0	5,5,5	0.33	0
7	FUC	A	301	-	10,10,11	1.48	1 (10%)	14,14,16	1.10	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	301	-	-	1/6/23/26	0/1/1/1
6	AGH	C	304	-	-	14/58/78/78	0/1/1/1
5	NAG	B	302	-	-	2/6/23/26	0/1/1/1
5	NAG	C	302	-	-	5/6/23/26	0/1/1/1
5	NAG	C	303	-	-	5/6/23/26	0/1/1/1
5	NAG	A	302	-	-	2/6/23/26	0/1/1/1
8	GOL	B	301	-	-	2/4/4/4	-
8	GOL	B	303	-	-	2/4/4/4	-
7	FUC	A	301	-	-	-	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	301	FUC	O5-C1	-3.65	1.37	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	302	NAG	C4-C3-C2	3.46	116.09	111.02
5	C	301	NAG	O5-C1-C2	-3.46	105.83	111.29
5	A	302	NAG	O5-C1-C2	-3.34	106.01	111.29
5	A	302	NAG	C3-C4-C5	3.24	116.01	110.24
5	C	303	NAG	O5-C1-C2	-2.97	106.59	111.29

There are no chirality outliers.

5 of 33 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	301	NAG	C1-C2-N2-C7
5	C	302	NAG	C3-C2-N2-C7
5	C	302	NAG	C8-C7-N2-C2
5	C	302	NAG	O7-C7-N2-C2
5	C	303	NAG	C3-C2-N2-C7

There are no ring outliers.

9 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	301	NAG	7	0
6	C	304	AGH	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	302	NAG	3	0
5	C	302	NAG	3	0
5	C	303	NAG	5	0
5	A	302	NAG	9	0
8	B	301	GOL	2	0
8	B	303	GOL	4	0
7	A	301	FUC	4	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 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.

Ligand AGH C 304

Bond lengths

Bond angles

Torsions

Rings
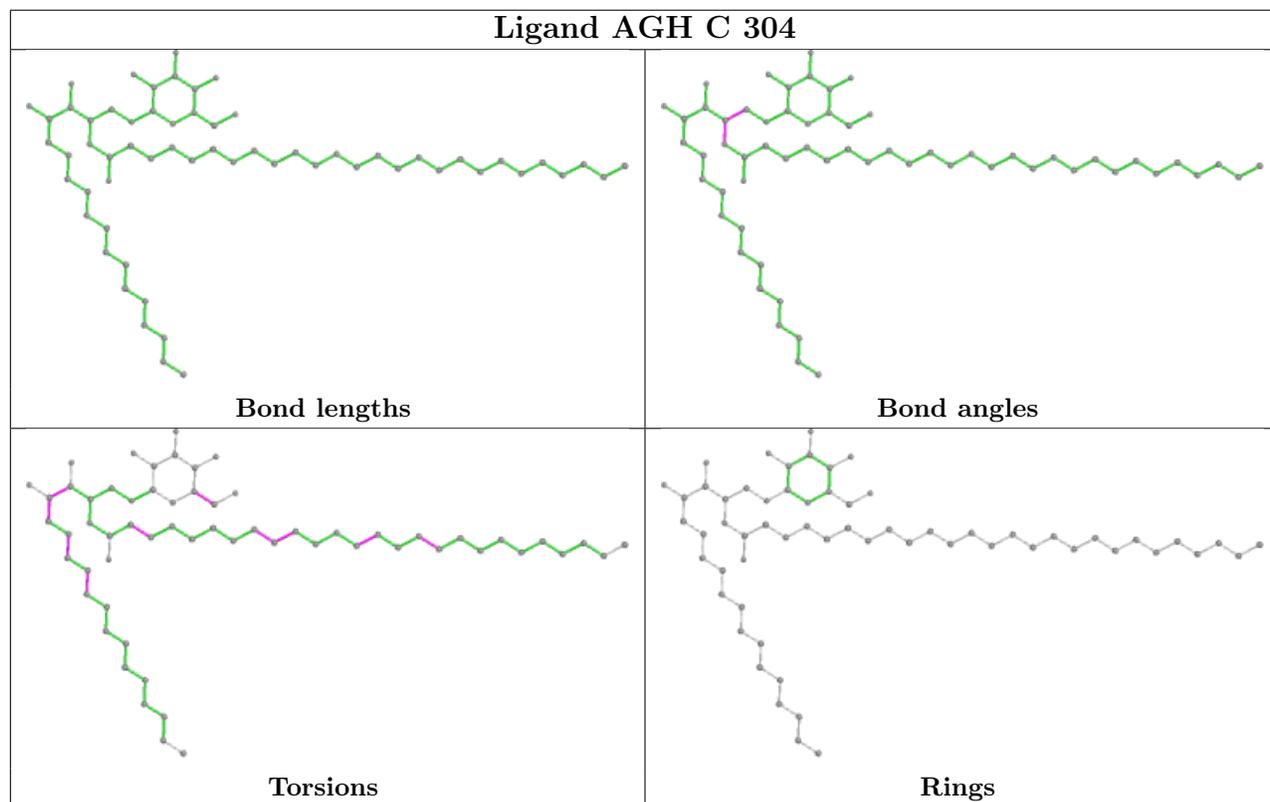
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	C	257/284 (90%)	0.47	15 (5%)	23	29	35, 55, 89, 97	0
2	A	186/220 (84%)	0.52	8 (4%)	35	44	40, 67, 108, 124	0
3	B	240/259 (92%)	0.62	21 (8%)	10	13	35, 61, 106, 118	0
4	D	99/100 (99%)	0.54	8 (8%)	12	16	33, 59, 92, 113	0
All	All	782/863 (90%)	0.53	52 (6%)	18	23	33, 59, 102, 124	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	1	ILE	8.4
1	C	202	LEU	4.6
1	C	252	VAL	4.5
3	B	242	TRP	4.5
1	C	195	PRO	4.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

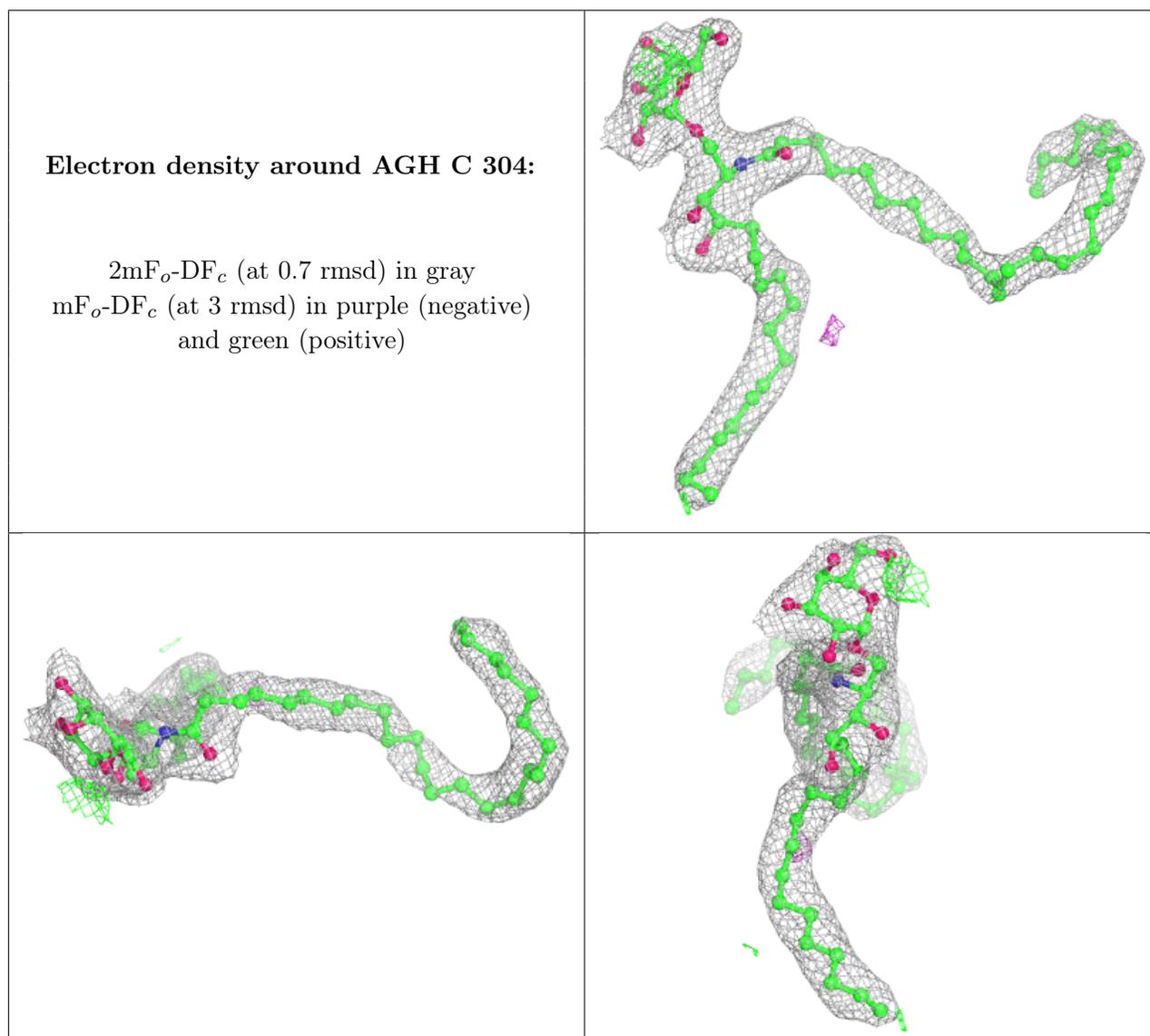There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
5	NAG	A	302	14/15	0.70	0.26	98,105,112,114	0
5	NAG	C	303	14/15	0.71	0.23	75,92,95,96	0
5	NAG	C	302	14/15	0.80	0.19	102,107,111,112	0
5	NAG	B	302	14/15	0.86	0.22	104,105,108,109	0
7	FUC	A	301	10/11	0.87	0.21	112,116,121,124	0
5	NAG	C	301	14/15	0.90	0.19	54,70,82,83	0
8	GOL	B	301	6/6	0.92	0.22	55,60,62,62	0
6	AGH	C	304	60/60	0.94	0.23	32,48,61,63	0
8	GOL	B	303	6/6	0.94	0.27	41,57,59,61	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 AGH C 304:

$2mF_o$-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 ⓘ

There are no such residues in this entry.