wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 31, 2020 – 11:15 AM GMT

PDB ID	:	7A17
Title	:	Crystal structure of the 5-phosphatase domain of Synaptojanin1 bound to its substrate diC8-PI(3,4,5)P3 in complex with a nanobody
Authors	:	Paesmans, J.; Galicia, C.; Martin, E.; Versees, W.
Deposited on	:	2020-08-12
Resolution	:	2.73 Å(reported)

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	:	2.16
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)	:	2.16

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 2.73 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.
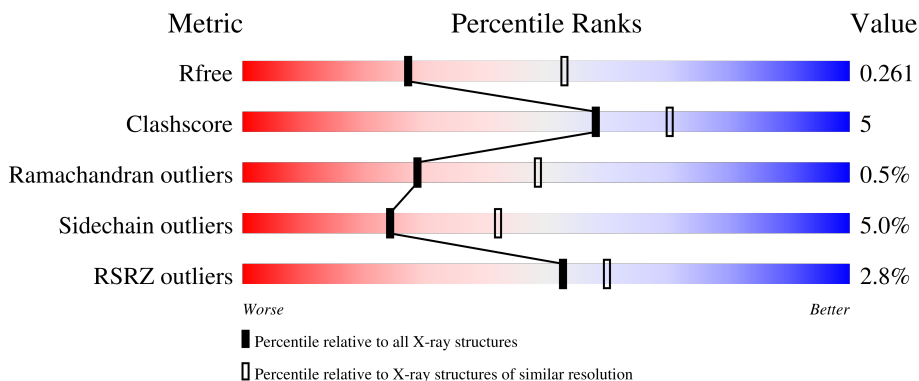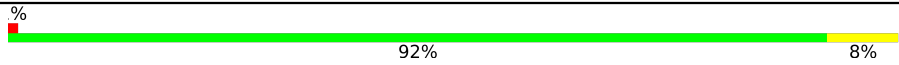

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1271 (2.76-2.72)
Clashscore	141614	1322 (2.76-2.72)
Ramachandran outliers	138981	1297 (2.76-2.72)
Sidechain outliers	138945	1298 (2.76-2.72)
RSRZ outliers	127900	1243 (2.76-2.72)

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	A	346	
2	B	119	
3	C	345	
3	E	345	
4	D	123	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	125	92% 8%

2 Entry composition (i)

There are 10 unique types of molecules in this entry. The entry contains 10780 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 Isoform 2 of Synaptojanin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	336	Total 2708	C 1735	N 471	O 491	S 11	0	1	0

- Molecule 2 is a protein called Nanobody 13015.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	119	Total 896	C 553	N 164	O 175	S 4	0	0	0

- Molecule 3 is a protein called Isoform 2 of Synaptojanin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	319	Total 2581	C 1653	N 446	O 472	S 10	0	0	0
3	E	313	Total 2537	C 1628	N 442	O 456	S 11	0	0	0

- Molecule 4 is a protein called Nanobody 13015.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	123	Total 942	C 582	N 173	O 183	S 4	0	0	0

- Molecule 5 is a protein called Nanobody 13015.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	125	Total 960	C 592	N 177	O 187	S 4	0	0	0

- Molecule 6 is (2R)-3-{[(R)-{[(1S,2S,3R,4S,5S,6S)-2,6-dihydroxy-3,4,5-tris(phosphonooxy) cyclohexyl]oxy}(hydroxy)phosphoryl]oxy}propane -1,2-diyl dioctanoate (three-letter code: IP9) (formula: $C_{25}H_{50}O_{22}P_4$) (labeled as "Ligand of Interest" by depositor).

IP9
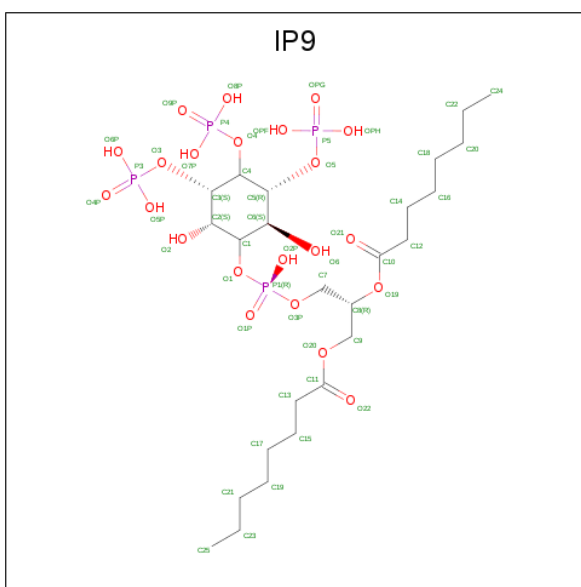
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	O	P	0	0
			51	25	22	4		

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Mg	0	0
			1	1		
7	C	1	Total	Mg	0	0
			1	1		

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

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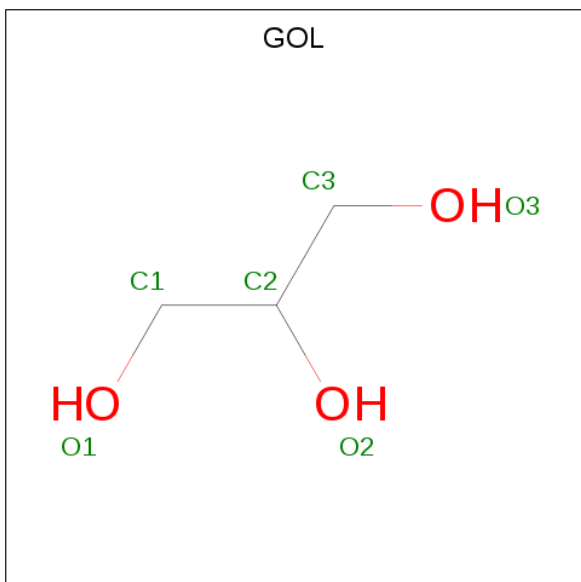
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total 6	C 3	O 3	0	0
8	A	1	Total 6	C 3	O 3	0	0

- Molecule 9 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).
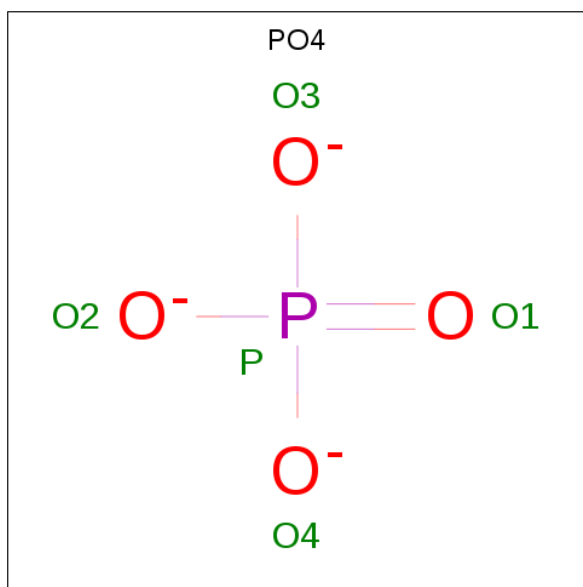
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	C	1	Total 5	O 4	P 1	0	0
9	C	1	Total 5	O 4	P 1	0	0
9	C	1	Total 5	O 4	P 1	0	0
9	E	1	Total 5	O 4	P 1	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	22	Total 22	O 22	0	0
10	B	9	Total 9	O 9	0	0
10	C	12	Total 12	O 12	0	0
10	D	8	Total 8	O 8	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	E	14	Total 14	O 14	0	0
10	F	6	Total 6	O 6	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: Isoform 2 of Synaptojanin-1
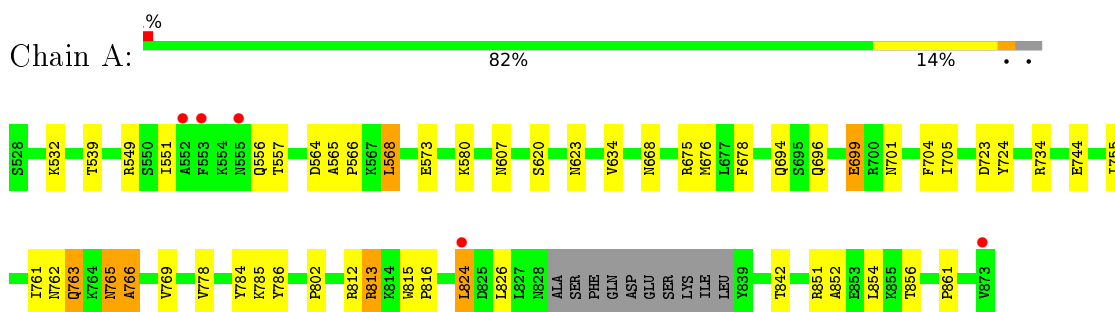
- Molecule 2: Nanobody 13015
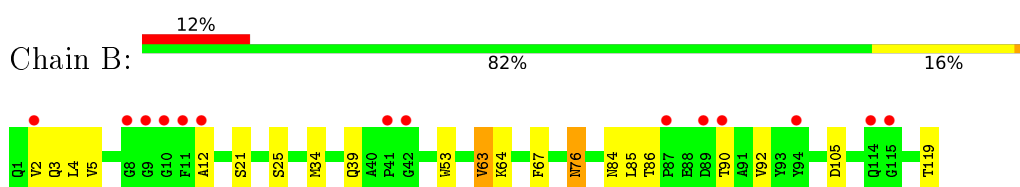
- Molecule 3: Isoform 2 of Synaptojanin-1

- Molecule 3: Isoform 2 of Synaptojanin-1

K529 T539 A552 T557 L558 T559 Q572 E573 R578 M592 L595 N596 A597 GLY ASN ILE VAL SER A603 S620 R621 D622 N623 E631 Q632 L633 V634 V653 A654 V655 G661 M662 GLY GLY ALA THR GLY N668 R675 M676 K698 R709 K710 S721

Y724 R734 D752 Q763 V778 Y784 L788 F789 S790 D791 P802 R811 R812 W815 PRO PHE ASP ARG SER ALA GLU ASP LEU ASP LEU ASN ALA SER PHE GLN ASP GLU SER LYS ILE L838 Y839 T840 W841 T842 T845 L846 L847 A852 E853 L854 K855 T856

P861 V873
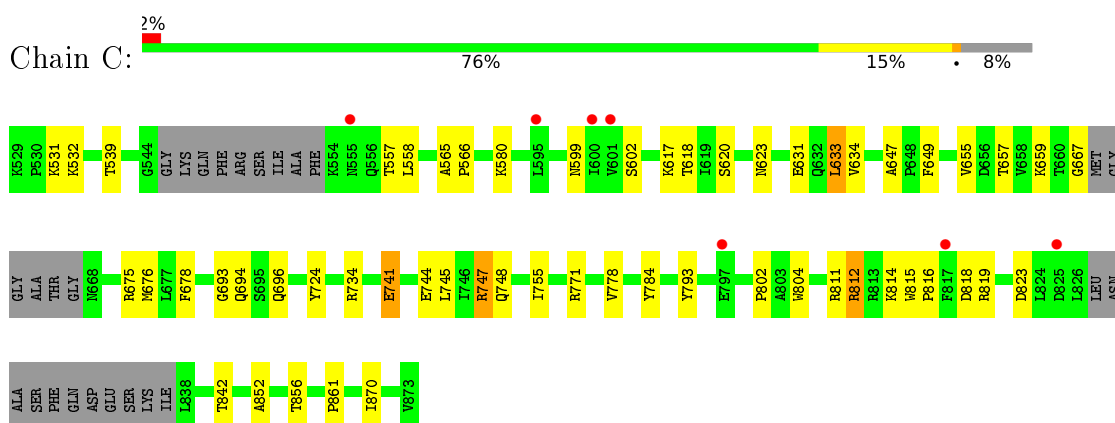
• **Molecule 4: Nanobody 13015**

Chain D: 89% 9%

Q1 Q39 E44 W53 V63 K64 F67 N84 L85 T86 T90 A91 V92 I101 D105 T119 H123
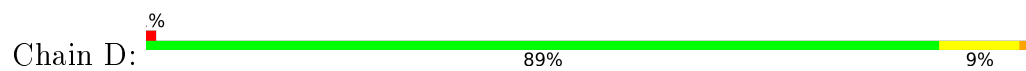
• **Molecule 5: Nanobody 13015**

Chain F: 92% 8%

Q1 F31 W53 V63 F67 N84 L85 T86 P87 T90 T119 S122 H125
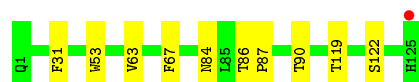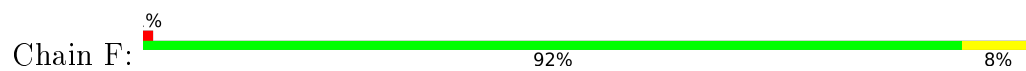
4   Data and refinement statistics $\;(i)$

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	169.32Å 109.20Å 100.90Å 90.00° 120.62° 90.00°	Depositor
Resolution (Å)	87.39 – 2.73 87.39 – 2.73	Depositor EDS
% Data completeness (in resolution range)	76.2 (87.39-2.73) 70.0 (87.39-2.73)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.64 (at 2.73Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.196 , 0.257 0.203 , 0.261	Depositor DCC
R_{free} test set	1629 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	35.2	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 47.1	EDS
L-test for twinning [2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	0.016 for -1/2*h+1/2*k+l,1/2*h-1/2*k+l,1/2*h+1/2*k 0.017 for -1/2*h-1/2*k+l,-1/2*h-1/2*k-l,1/2*h-1/2*k	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10780	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.95% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4, IP9, MG

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	0.35	0/2775	0.64	2/3755 (0.1%)
2	B	0.33	0/915	0.61	0/1237
3	C	0.31	0/2641	0.60	0/3576
3	E	0.32	0/2595	0.61	0/3507
4	D	0.33	0/962	0.57	0/1298
5	F	0.34	0/981	0.60	0/1324
All	All	0.33	0/10869	0.61	2/14697 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	766	ALA	N-CA-CB	5.50	117.80	110.10
1	A	765	ASN	C-N-CA	5.21	134.72	121.70

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	2708	0	2671	28	0
2	B	896	0	821	10	0
3	C	2581	0	2528	30	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	2537	0	2510	26	0
4	D	942	0	889	8	0
5	F	960	0	900	6	0
6	A	51	0	43	1	0
7	A	1	0	0	0	0
7	C	1	0	0	0	0
8	A	12	0	16	0	0
9	C	15	0	0	0	0
9	E	5	0	0	0	0
10	A	22	0	0	0	0
10	B	9	0	0	0	0
10	C	12	0	0	0	0
10	D	8	0	0	0	0
10	E	14	0	0	0	0
10	F	6	0	0	0	0
All	All	10780	0	10378	101	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 5.

The worst 5 of 101 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:762:ASN:O	1:A:765:ASN:O	2.05	0.73
2:B:90:THR:HG22	2:B:119:THR:HA	1.69	0.73
1:A:668:ASN:ND2	6:A:901:IP9:O1P	2.20	0.72
3:E:724:TYR:CZ	3:E:812:ARG:HD2	2.24	0.72
4:D:105:ASP:N	4:D:105:ASP:OD1	2.23	0.70

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	333/346 (96%)	317 (95%)	14 (4%)	2 (1%)	25	44
2	B	117/119 (98%)	114 (97%)	3 (3%)	0	100	100
3	C	311/345 (90%)	298 (96%)	11 (4%)	2 (1%)	25	44
3	E	305/345 (88%)	296 (97%)	7 (2%)	2 (1%)	22	40
4	D	121/123 (98%)	120 (99%)	1 (1%)	0	100	100
5	F	123/125 (98%)	121 (98%)	2 (2%)	0	100	100
All	All	1310/1403 (93%)	1266 (97%)	38 (3%)	6 (0%)	29	48

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	766	ALA
3	C	823	ASP
1	A	634	VAL
3	C	634	VAL
3	E	573	GLU

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	289/301 (96%)	275 (95%)	14 (5%)	25	44
2	B	87/93 (94%)	79 (91%)	8 (9%)	9	17
3	C	277/300 (92%)	265 (96%)	12 (4%)	29	48
3	E	272/300 (91%)	258 (95%)	14 (5%)	24	41
4	D	96/97 (99%)	90 (94%)	6 (6%)	18	31
5	F	98/99 (99%)	96 (98%)	2 (2%)	55	72
All	All	1119/1190 (94%)	1063 (95%)	56 (5%)	24	42

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	633	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	812	ARG
3	E	842	THR
3	C	678	PHE
3	C	747	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	84	ASN
3	C	608	GLN
3	E	632	GLN
2	B	76	ASN
3	C	859	HIS

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 ⓘ

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	GOL	A	904	-	5,5,5	0.09	0	5,5,5	0.28	0
9	PO4	E	901	-	4,4,4	0.64	0	6,6,6	0.45	0
9	PO4	C	902	-	4,4,4	0.63	0	6,6,6	0.49	0
8	GOL	A	903	-	5,5,5	0.09	0	5,5,5	0.27	0
9	PO4	C	901	-	4,4,4	0.72	0	6,6,6	0.40	0
6	IP9	A	901	-	51,51,51	0.74	0	67,72,72	0.69	0
9	PO4	C	903	-	4,4,4	0.59	0	6,6,6	0.45	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
8	GOL	A	903	-	-	2/4/4/4	-
8	GOL	A	904	-	-	2/4/4/4	-
6	IP9	A	901	-	-	19/49/73/73	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	904	GOL	C1-C2-C3-O3
8	A	903	GOL	O1-C1-C2-O2
8	A	903	GOL	O1-C1-C2-C3
6	A	901	IP9	C1-O1-P1-O2P
8	A	904	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	901	IP9	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 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 IP9 A 901

Bond lengths

Bond angles

Torsions

Rings
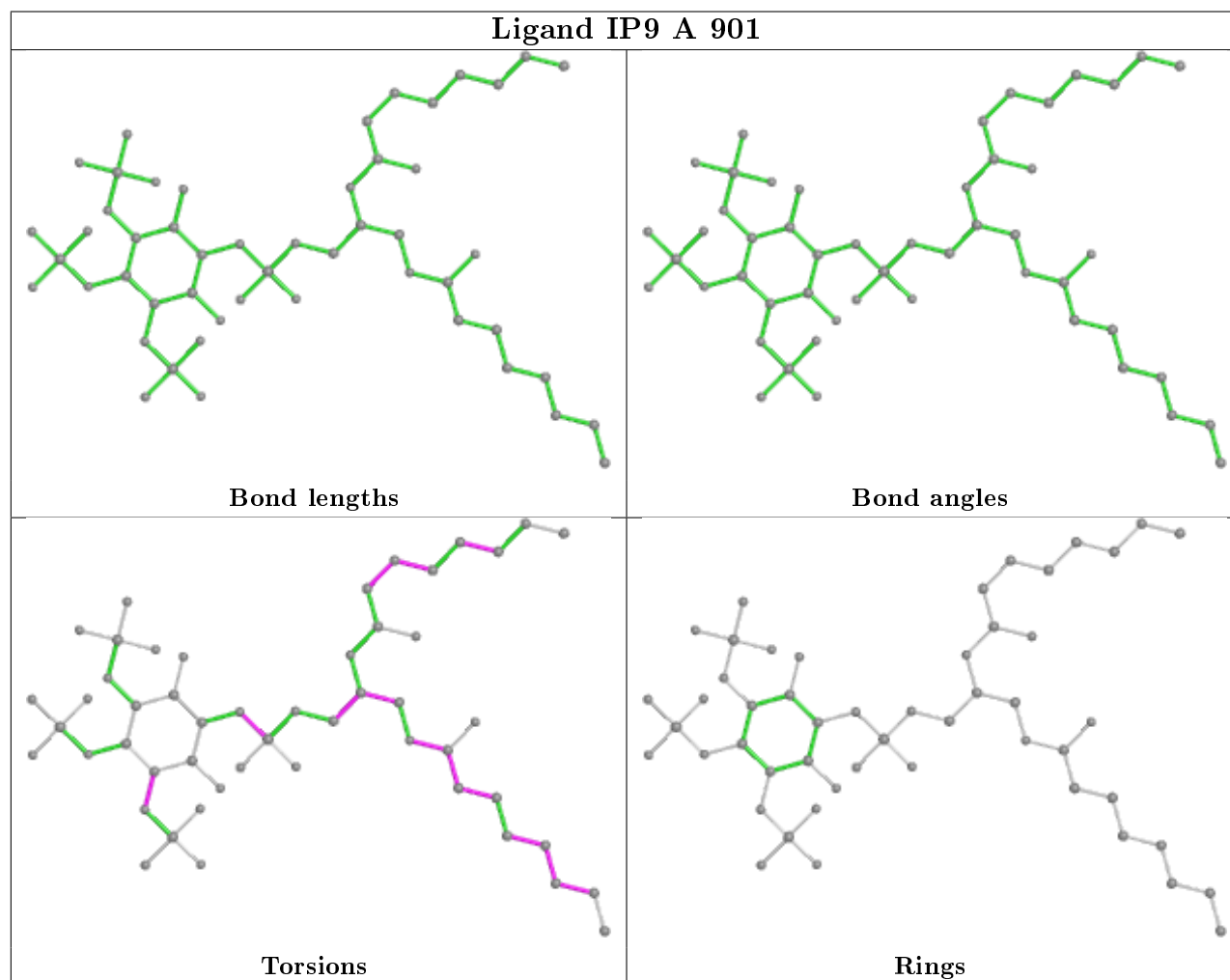
5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	660:THR	C	667:GLY	N	5.44

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, 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	A	336/346 (97%)	-0.10	5 (1%)	73	79	18, 31, 65, 102	0
2	B	119/119 (100%)	0.75	14 (11%)	4	4	28, 68, 99, 106	0
3	C	319/345 (92%)	-0.01	7 (2%)	62	69	26, 44, 80, 113	0
3	E	313/345 (90%)	0.12	9 (2%)	51	58	24, 45, 75, 107	0
4	D	123/123 (100%)	-0.09	1 (0%)	86	89	26, 38, 63, 80	0
5	F	125/125 (100%)	-0.16	1 (0%)	86	89	20, 32, 59, 95	0
All	All	1335/1403 (95%)	0.04	37 (2%)	53	60	18, 40, 83, 113	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	555	ASN	4.7
3	E	839	TYR	4.6
2	B	90	THR	4.5
1	A	873	VAL	3.9
3	E	662	MET	3.7

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

There are no non-standard protein/DNA/RNA residues in this entry.
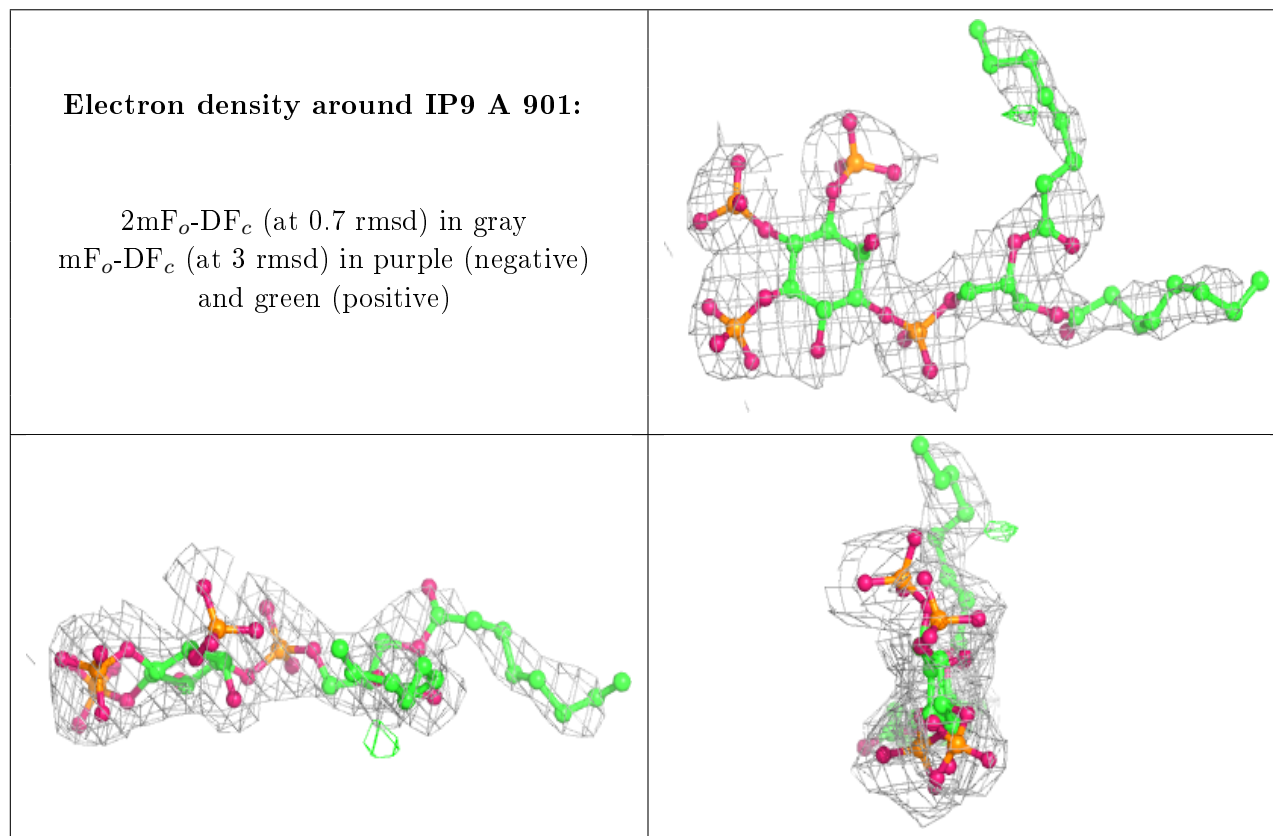
6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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
8	GOL	A	903	6/6	0.77	0.19	39,47,53,55	0
9	PO4	C	901	5/5	0.78	0.18	79,86,97,100	0
9	PO4	C	903	5/5	0.80	0.23	106,114,117,131	0
7	MG	A	902	1/1	0.85	0.17	46,46,46,46	0
9	PO4	E	901	5/5	0.88	0.18	80,82,95,97	0
8	GOL	A	904	6/6	0.91	0.14	35,42,49,54	0
7	MG	C	904	1/1	0.92	0.09	38,38,38,38	0
9	PO4	C	902	5/5	0.92	0.17	78,78,97,98	0
6	IP9	A	901	51/51	0.94	0.17	39,63,74,79	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 IP9 A 901:

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

6.5 Other polymers (i)

There are no such residues in this entry.