wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2022 – 01:17 pm BST

PDB ID	:	7OHO
Title	:	Crystal structure of AP2 FCHO2 chimera
Authors	:	Zaccai, N.R.; Kelly, B.T.; Evans, P.R.; Owen, D.J.
Deposited on	:	2021-05-11
Resolution	:	2.88 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 following versions of software and data (see references (i)) were used in the production of this report:

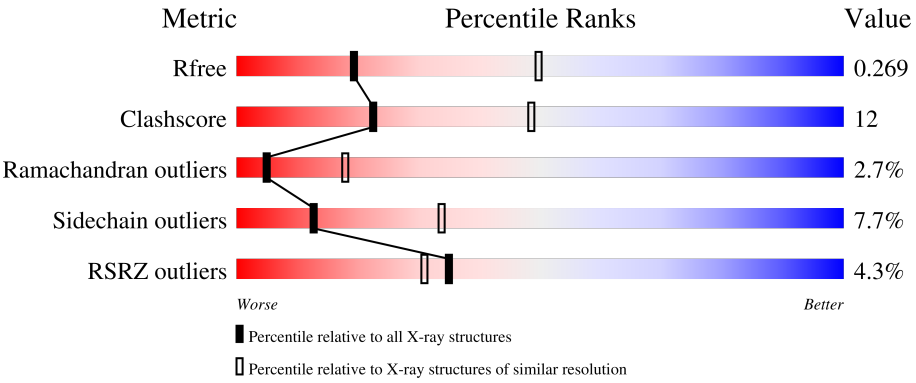
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.28.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.28.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

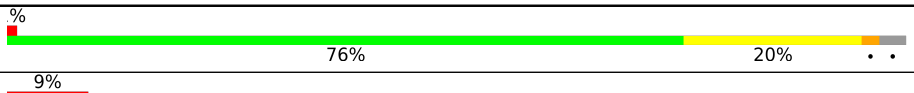
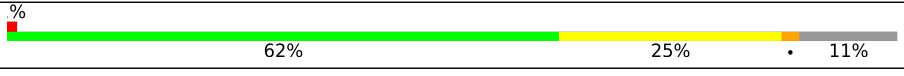
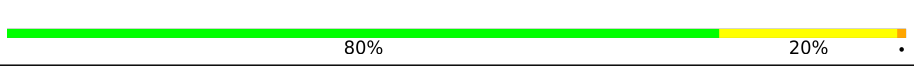
The reported resolution of this entry is 2.88 Å.

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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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	AAA	621	
2	BBB	636	
3	MMM	446	
4	SSS	142	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 13681 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 AP-2 complex subunit alpha-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	605	Total	C	N	O	S	0	0	0
			4774	3040	820	893	21			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	272	GLU	-	insertion	UNP P18484

- Molecule 2 is a protein called AP-2 complex subunit beta,F-BAR domain only protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	561	Total	C	N	O	S	0	0	0
			4446	2830	741	851	24			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	626	PHE	SER	conflict	UNP Q0JRZ9
BBB	630	GLY	-	expression tag	UNP Q0JRZ9
BBB	631	HIS	-	expression tag	UNP Q0JRZ9
BBB	632	HIS	-	expression tag	UNP Q0JRZ9
BBB	633	HIS	-	expression tag	UNP Q0JRZ9
BBB	634	HIS	-	expression tag	UNP Q0JRZ9
BBB	635	HIS	-	expression tag	UNP Q0JRZ9
BBB	636	HIS	-	expression tag	UNP Q0JRZ9

- Molecule 3 is a protein called AP-2 complex subunit mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	MMM	397	Total	C	N	O	S	0	0	0
			3201	2058	561	563	19			

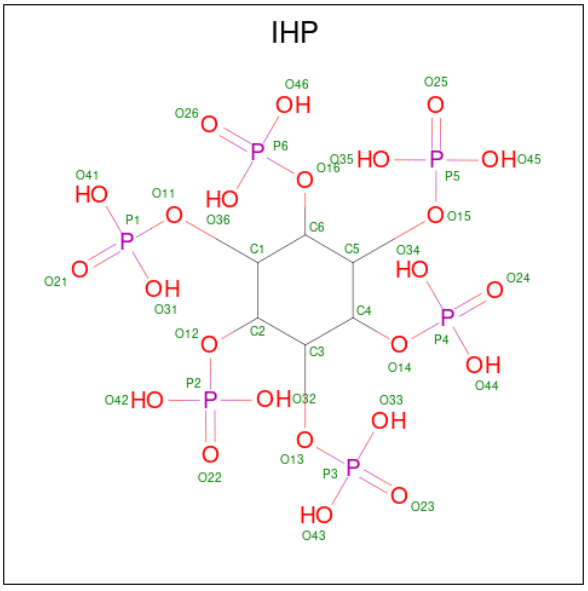
There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
MMM	222O	MET	-	insertion	UNP P84092
MMM	222P	GLU	-	insertion	UNP P84092
MMM	222Q	GLN	-	insertion	UNP P84092
MMM	222R	LYS	-	insertion	UNP P84092
MMM	222S	LEU	-	insertion	UNP P84092
MMM	222T	ILE	-	insertion	UNP P84092
MMM	222U	SER	-	insertion	UNP P84092
MMM	222V	GLU	-	insertion	UNP P84092
MMM	222W	GLU	-	insertion	UNP P84092
MMM	222X	ASP	-	insertion	UNP P84092
MMM	222Y	LEU	-	insertion	UNP P84092

- Molecule 4 is a protein called AP-2 complex subunit sigma.

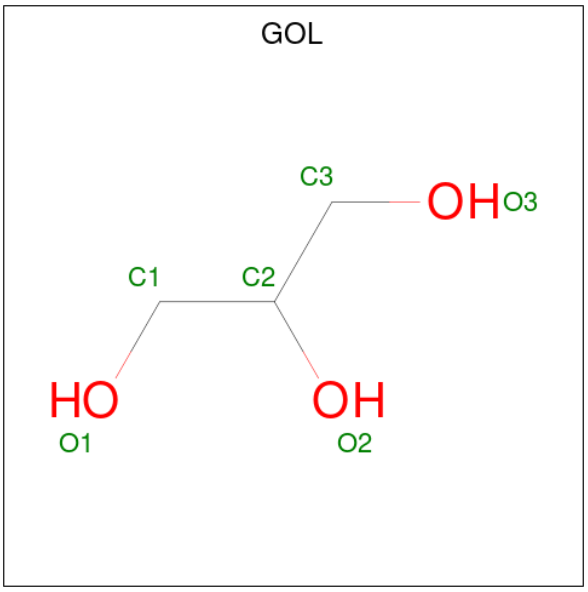
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	SSS	142	Total	C	N	O	S	0	0	0
			1200	778	200	215	7			

- Molecule 5 is INOSITOL HEXAKISPHOSPHATE (three-letter code: IHP) (formula: $C_6H_{18}O_{24}P_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	AAA	1	Total	C	O	P	0	0
			36	6	24	6		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



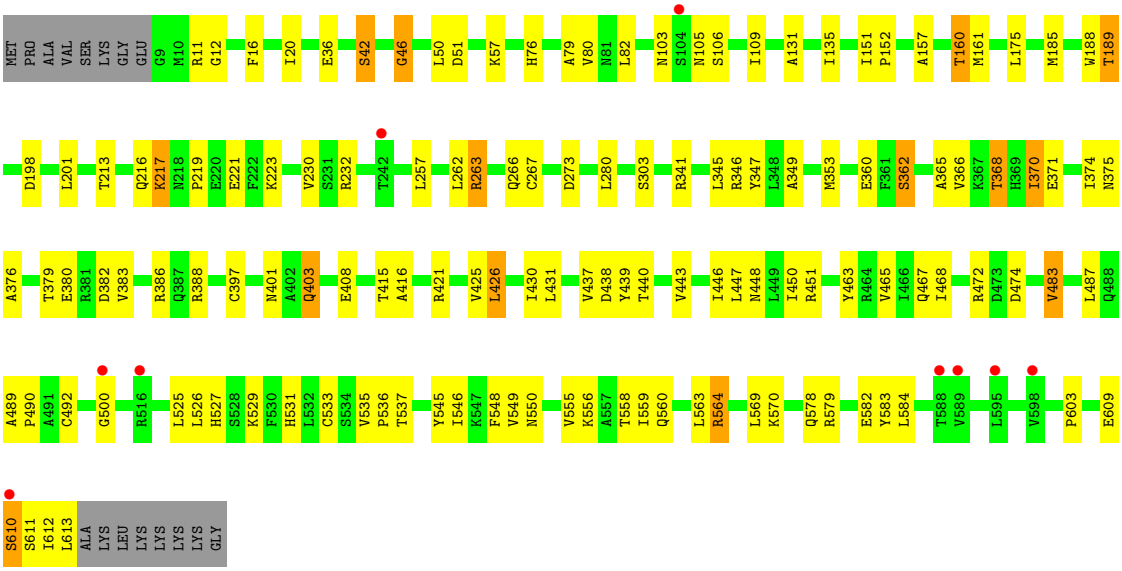
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	AAA	1	6	3	3	0	0
			Total	C	O		
6	AAA	1	6	3	3	0	0
			Total	C	O		
6	AAA	1	6	3	3	0	0
			Total	C	O		
6	MMM	1	6	3	3	0	0

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.

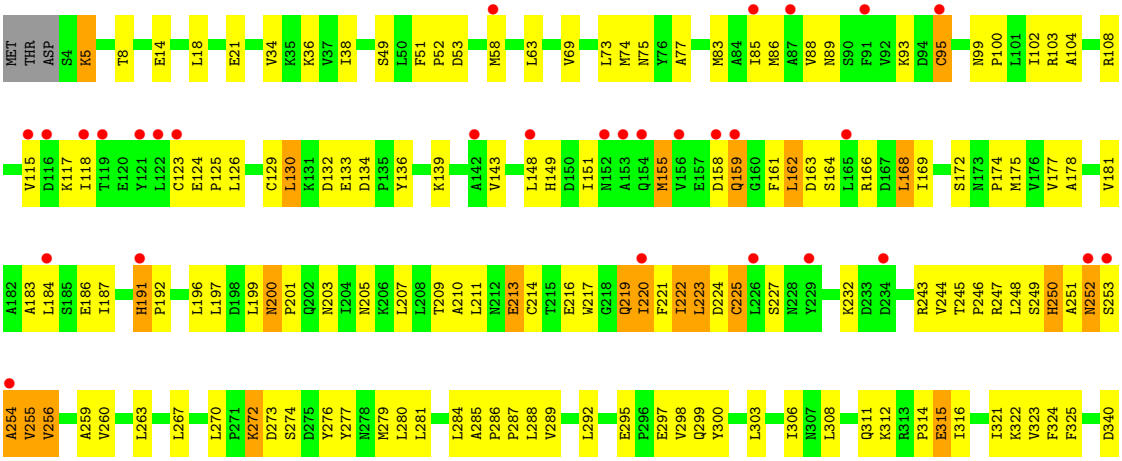
• Molecule 1: AP-2 complex subunit alpha-2

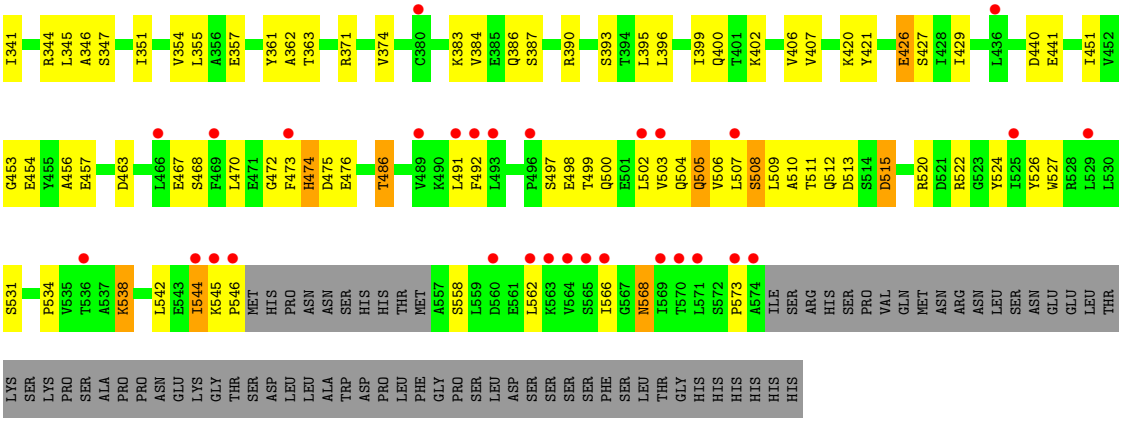
Chain AAA:



• Molecule 2: AP-2 complex subunit beta,F-BAR domain only protein 2

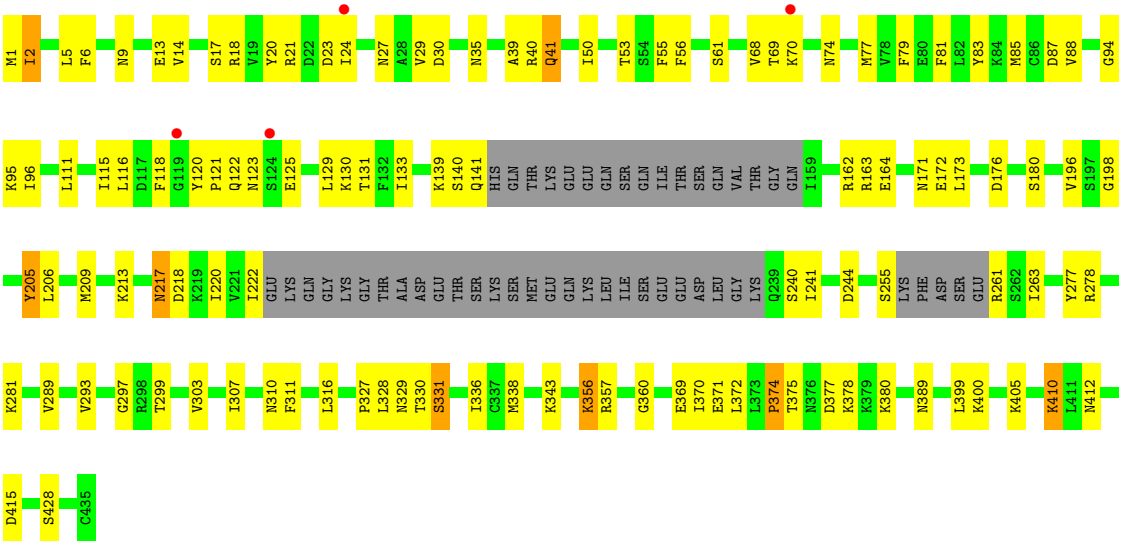
Chain BBB:





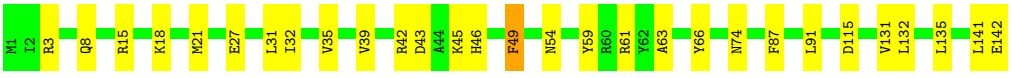
• Molecule 3: AP-2 complex subunit mu

Chain MMM:



• Molecule 4: AP-2 complex subunit sigma

Chain SSS:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	122.04Å 122.04Å 257.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	61.09 – 2.88 61.02 – 2.88	Depositor EDS
% Data completeness (in resolution range)	100.0 (61.09-2.88) 100.0 (61.02-2.88)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 2.86Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.182 , 0.273 0.186 , 0.269	Depositor DCC
R_{free} test set	2547 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	91.5	Xtriage
Anisotropy	0.161	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.027 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	13681	wwPDB-VP
Average B, all atoms (Å^2)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.34% 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: GOL, IHP

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	AAA	0.74	0/4859	0.86	0/6591
2	BBB	0.72	0/4511	0.85	0/6113
3	MMM	0.71	0/3264	0.92	1/4394 (0.0%)
4	SSS	0.68	0/1224	0.87	0/1650
All	All	0.72	0/13858	0.87	1/18748 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	MMM	205	TYR	CB-CA-C	5.11	120.62	110.40

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	AAA	4774	0	4868	73	0
2	BBB	4446	0	4576	178	1
3	MMM	3201	0	3303	69	0
4	SSS	1200	0	1195	17	0
5	AAA	36	0	6	1	0
6	AAA	18	0	24	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	MMM	6	0	8	0	0
All	All	13681	0	13980	322	1

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 12.

The worst 5 of 322 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)
2:BBB:254:ALA:O	2:BBB:255:VAL:HG23	1.63	0.97
1:AAA:198:ASP:O	1:AAA:232:ARG:NH2	1.99	0.94
2:BBB:203:ASN:O	2:BBB:207:LEU:HG	1.73	0.88
3:MMM:293:VAL:HG22	3:MMM:303:VAL:HG22	1.63	0.81
2:BBB:267:LEU:HD21	2:BBB:281:LEU:HD11	1.64	0.80

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:BBB:213:GLU:O	2:BBB:213:GLU:O[4_555]	2.17	0.03

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	AAA	603/621 (97%)	555 (92%)	41 (7%)	7 (1%)	13 38
2	BBB	557/636 (88%)	446 (80%)	84 (15%)	27 (5%)	2 7
3	MMM	389/446 (87%)	336 (86%)	42 (11%)	11 (3%)	5 17
4	SSS	140/142 (99%)	129 (92%)	10 (7%)	1 (1%)	22 52
All	All	1689/1845 (92%)	1466 (87%)	177 (10%)	46 (3%)	5 18

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	189	THR
2	BBB	200	ASN
2	BBB	252	ASN
2	BBB	255	VAL
2	BBB	362	ALA

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	AAA	531/543 (98%)	504 (95%)	27 (5%)	24 53
2	BBB	504/574 (88%)	450 (89%)	54 (11%)	6 18
3	MMM	354/398 (89%)	323 (91%)	31 (9%)	10 28
4	SSS	131/131 (100%)	126 (96%)	5 (4%)	33 65
All	All	1520/1646 (92%)	1403 (92%)	117 (8%)	13 34

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	BBB	274	SER
3	MMM	380	LYS
2	BBB	468	SER
3	MMM	378	LYS
3	MMM	281	LYS

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)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

5 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	IHP	AAA	701	-	36,36,36	0.98	3 (8%)	54,60,60	1.27	6 (11%)
6	GOL	AAA	702	-	5,5,5	0.22	0	5,5,5	0.50	0
6	GOL	AAA	704	-	5,5,5	0.12	0	5,5,5	0.32	0
6	GOL	AAA	703	-	5,5,5	0.09	0	5,5,5	0.31	0
6	GOL	MMM	501	-	5,5,5	0.14	0	5,5,5	0.39	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
5	IHP	AAA	701	-	-	4/30/54/54	0/1/1/1
6	GOL	AAA	702	-	-	1/4/4/4	-
6	GOL	AAA	704	-	-	1/4/4/4	-
6	GOL	AAA	703	-	-	3/4/4/4	-
6	GOL	MMM	501	-	-	4/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	AAA	701	IHP	P3-O13	2.66	1.64	1.59
5	AAA	701	IHP	P2-O12	2.15	1.63	1.59
5	AAA	701	IHP	P1-O11	-2.10	1.55	1.59

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	AAA	701	IHP	C6-C1-C2	4.34	119.91	110.41
5	AAA	701	IHP	O11-C1-C6	-3.35	100.78	108.69
5	AAA	701	IHP	C6-C5-C4	2.49	115.87	110.41
5	AAA	701	IHP	C3-C2-C1	2.32	115.49	110.41
5	AAA	701	IHP	C5-C6-C1	2.29	115.42	110.41

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	AAA	701	IHP	C5-O15-P5-O25
5	AAA	701	IHP	C6-O16-P6-O26
6	AAA	703	GOL	C1-C2-C3-O3
6	MMM	501	GOL	O1-C1-C2-O2
6	MMM	501	GOL	O1-C1-C2-C3

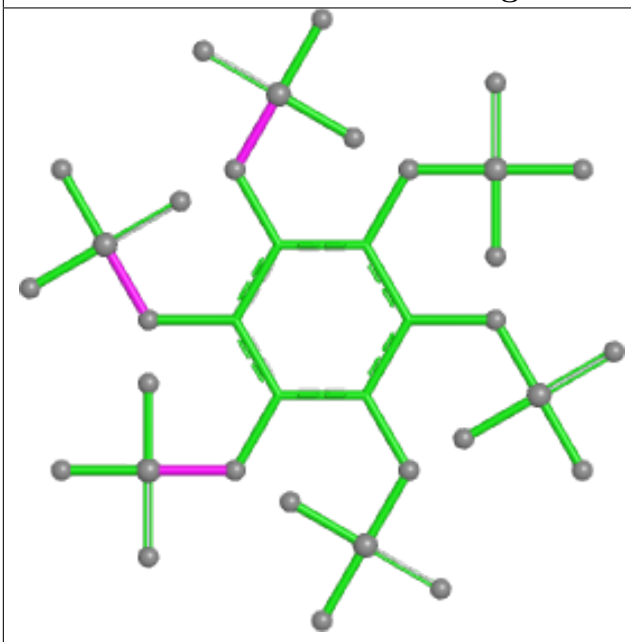
There are no ring outliers.

2 monomers are involved in 2 short contacts:

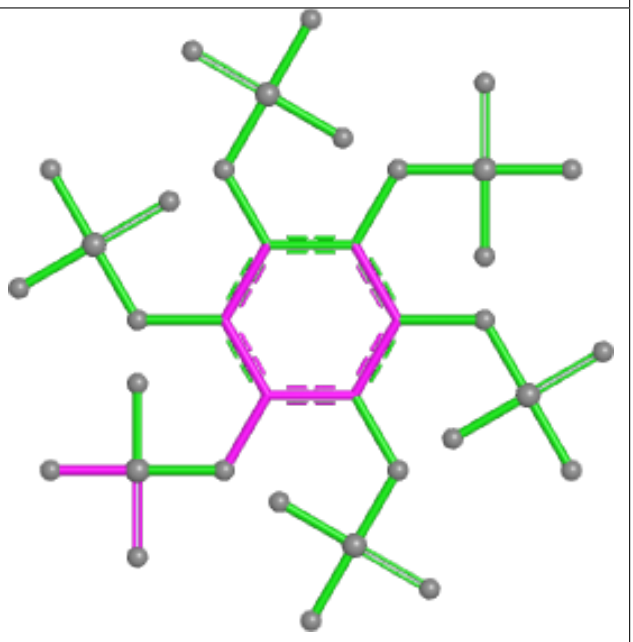
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	AAA	701	IHP	1	0
6	AAA	704	GOL	1	0

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.

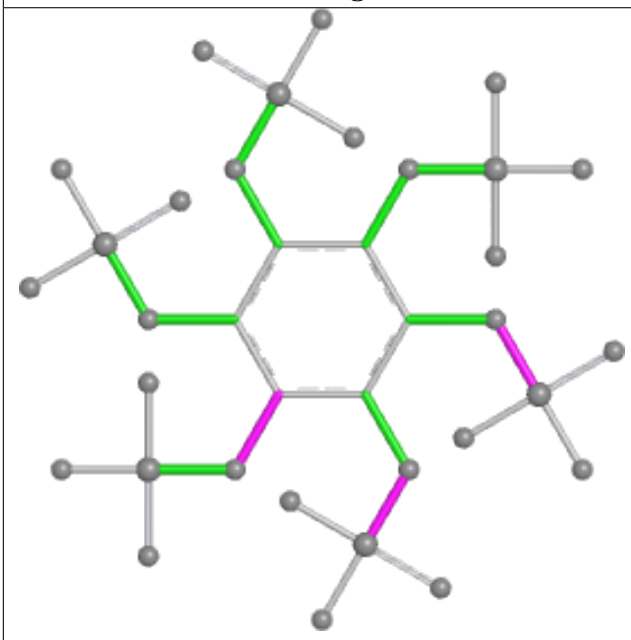
Ligand IHP AAA 701



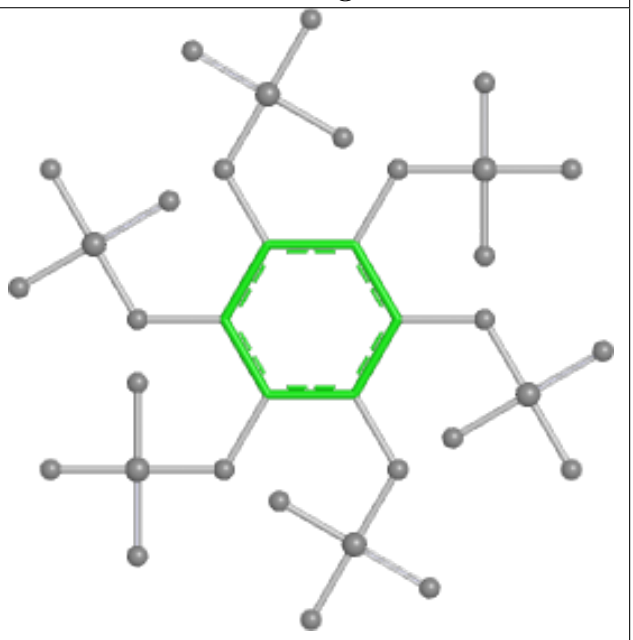
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

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^{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	AAA	605/621 (97%)	0.03	9 (1%) 73 73	57, 88, 126, 172	0
2	BBB	561/636 (88%)	0.52	60 (10%) 6 4	63, 126, 171, 202	0
3	MMM	397/446 (89%)	0.15	4 (1%) 82 82	59, 96, 147, 190	0
4	SSS	142/142 (100%)	0.07	0 100 100	52, 70, 95, 118	0
All	All	1705/1845 (92%)	0.22	73 (4%) 35 31	52, 99, 157, 202	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	BBB	546	PRO	8.2
2	BBB	152	ASN	4.9
2	BBB	560	ASP	4.5
2	BBB	570	THR	4.5
2	BBB	226	LEU	4.3

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

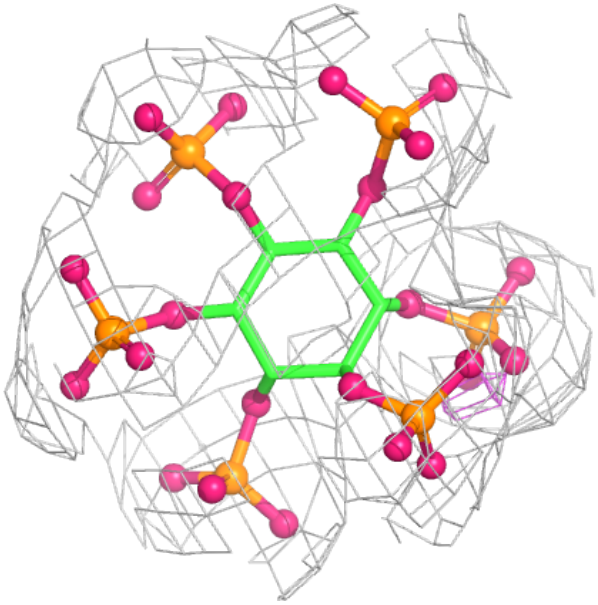
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

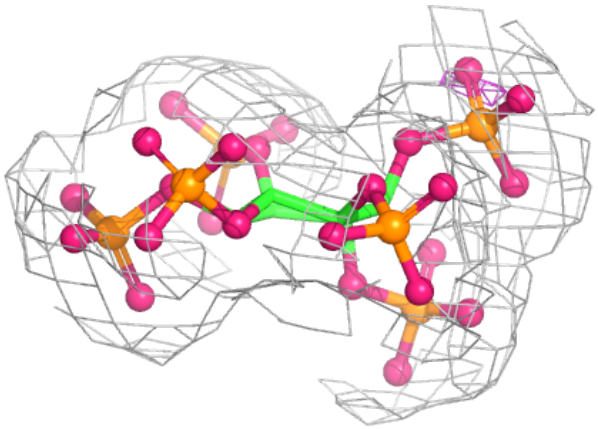
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	GOL	AAA	704	6/6	0.42	0.30	131,142,147,149	0
6	GOL	AAA	702	6/6	0.60	0.38	104,125,140,143	0
6	GOL	AAA	703	6/6	0.64	0.27	106,128,135,137	0
6	GOL	MMM	501	6/6	0.85	0.27	96,110,119,125	0
5	IHP	AAA	701	36/36	0.95	0.13	66,98,119,123	0

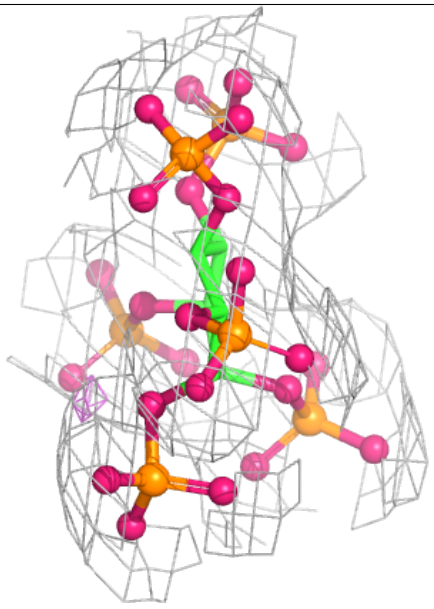
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around IHP AAA 701:

$2mF_o$-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.