wwPDB X-ray Structure Validation Summary Report (i)

Jul 3, 2024 – 02:05 pm BST

PDB ID	:	9FWG
Title	:	LSD1/CoREST bound to bomedemstat
Authors	:	Speranzini, V.; Mattevi, A.
Deposited on	:	2024-06-30
Resolution	:	3.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

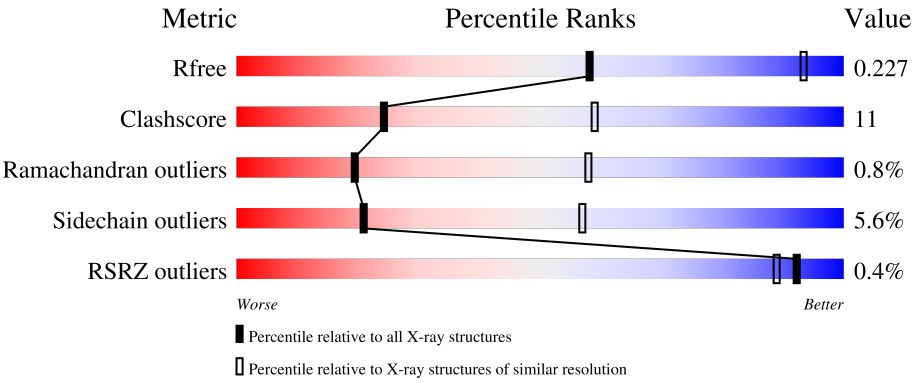
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	852	56% 20% · 22%
2	B	485	18% 8% · 73%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6357 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.

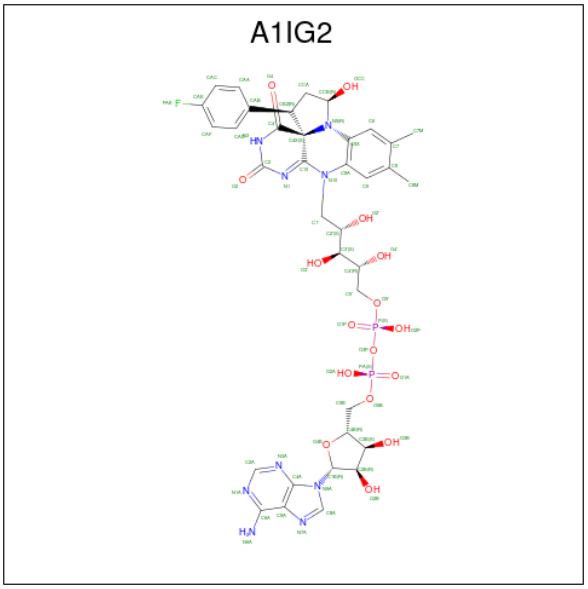
- Molecule 1 is a protein called Lysine-specific histone demethylase 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	666	Total	C	N	O	S	0	0	0
			5217	3324	906	967	20			

- Molecule 2 is a protein called REST corepressor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	133	Total	C	N	O	S	0	0	0
			1076	676	194	203	3			

- Molecule 3 is Bomedemstat FAD adduct (three-letter code: A1IG2) (formula: $C_{36}H_{42}FN_9O_{16}P_2$) (labeled as "Ligand of Interest" by depositor).



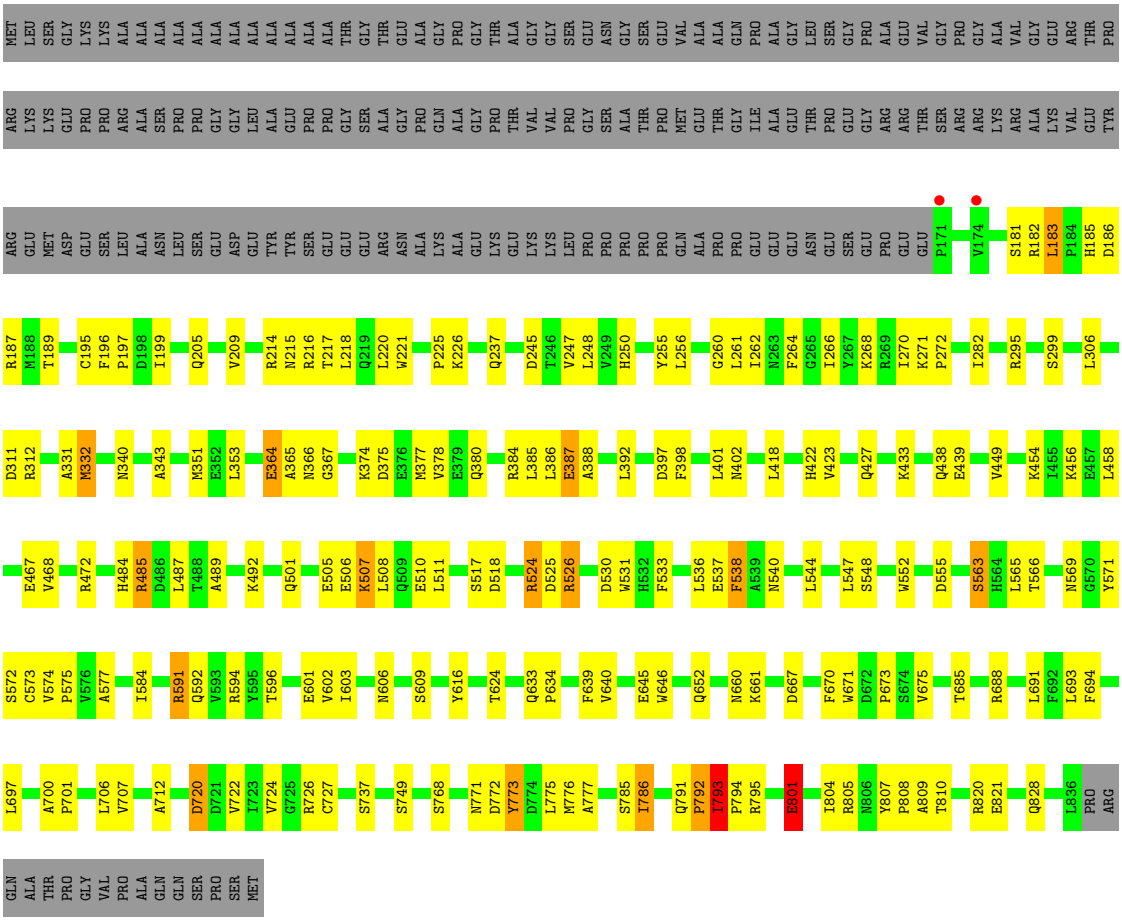
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	C	F	N	O	P	0	0
			64	36	1	9	16	2		

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.

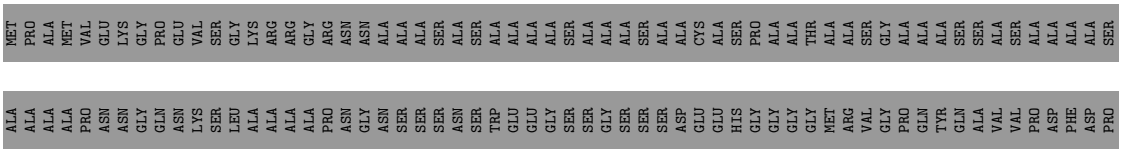
• Molecule 1: Lysine-specific histone demethylase 1A

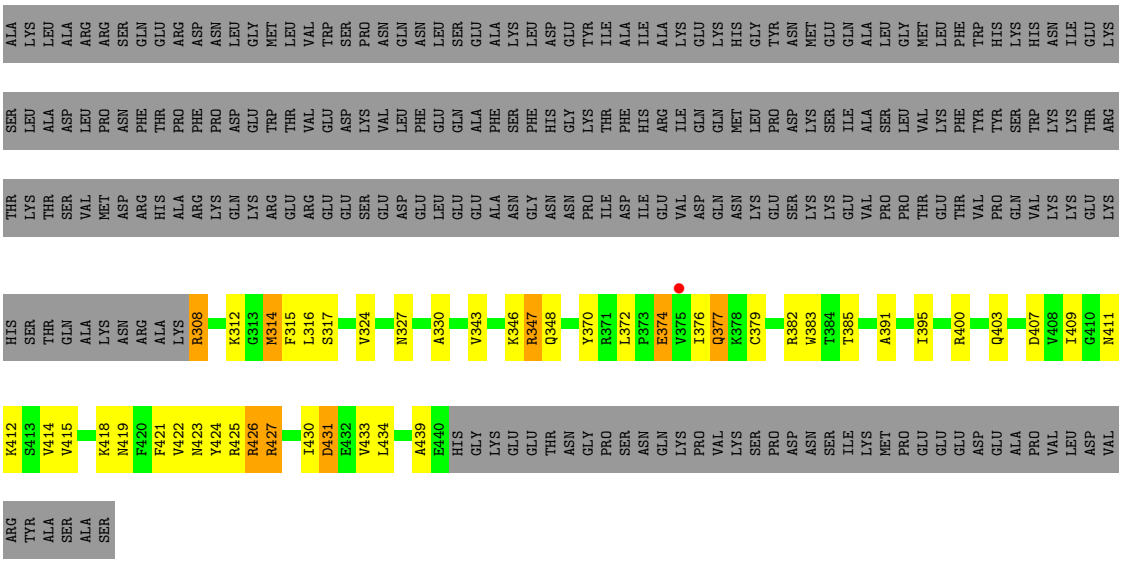
Chain A:



• Molecule 2: REST corepressor 1

Chain B:




ALA LYS LEU ALA ARG ARG SER GLN GLU ARG ASP ASN LEU GLY MET LEU VAL TRP SER PRO ASN GLN ASN LEU SER GLU ALA LYS LEU ASP GLU TYR ILE ALA ILE ALA LYS GLU LYS HIS GLY TYR ASN MET GLU GLN ALA LEU GLY MET LEU PHE TRP HIS LYS HIS ASN ILE GLU LYS
SER LEU ALA ASP LEU PRO ASN PHE THR PRO PHE PRO ASP GLU TRP THR VAL GLU ASP LYS VAL LEU PHE GLU GLN ALA PHE SER PHE HIS GLY LYS THR PHE HIS ARG ILE GLN GLN MET LEU PRO ASP LYS SER LYS LYS GLU SER LEU ALA SER LEU VAL LYS PHE TYR TYR SER TRP LYS LYS THR ARG
THR LYS THR SER VAL MET ASP ARG HIS ALA ARG LYS GLN LYS ARG GLU ARG GLU GLU SER GLU ASP GLU LEU GLU GLU ALA ASN GLY ASN ASN PRO ILE ASP ILE GLU VAL ASP GLN GLN ASN LYS GLU SER LYS LYS GLU VAL PRO PRO THR GLU THR VAL PRO GLN VAL LYS LYS GLU LYS
HIS SER THR GLN ALA LYS ASN ARG ALA LYS R308 K312 G313 M314 F315 L316 S317 V324 N327 A330 V343 K346 R347 Q348 Y370 R371 L372 P373 E374 V375 I376 Q377 K378 C379 R382 W383 T384 T385 A391 I395 R400 Q403 D407 V408 I409 G410 N411
K412 S413 V414 V415 K418 N419 F420 F421 V422 N423 Y424 R425 R426 R427 I430 D431 E432 V433 L434 A439 E440 HIS GLY LYS GLU GLU THR ASN GLY PRO SER ASN GLN LYS PRO VAL LYS SER PRO ASP ASN SER ILE LYS MET PRO GLU GLU GLU ASP GLU ALA PRO VAL LEU ASP VAL
ARG TYR ALA SER ALA SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	119.66Å 179.02Å 234.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	98.00 – 3.20 99.48 – 3.20	Depositor EDS
% Data completeness (in resolution range)	97.7 (98.00-3.20) 97.7 (99.48-3.20)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.191 , 0.227 0.194 , 0.227	Depositor DCC
R_{free} test set	802 reflections (1.96%)	wwPDB-VP
Wilson B-factor (Å^2)	77.8	Xtriage
Anisotropy	0.525	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 57.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6357	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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: A1IG2

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	1.13	11/5331 (0.2%)	1.17	26/7232 (0.4%)
2	B	0.86	0/1091	1.08	7/1471 (0.5%)
All	All	1.09	11/6422 (0.2%)	1.16	33/8703 (0.4%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
2	B	0	1
All	All	0	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	646	TRP	CB-CG	-10.97	1.30	1.50
1	A	821	GLU	CD-OE1	6.43	1.32	1.25
1	A	387	GLU	CG-CD	6.06	1.61	1.51
1	A	591	ARG	CZ-NH1	5.92	1.40	1.33
1	A	596	THR	CB-CG2	5.82	1.71	1.52

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	795	ARG	NE-CZ-NH1	9.77	125.18	120.30
1	A	397	ASP	CB-CG-OD1	9.03	126.42	118.30
1	A	555	ASP	CB-CG-OD1	8.68	126.11	118.30
1	A	795	ARG	NE-CZ-NH2	-8.60	116.00	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	295	ARG	NE-CZ-NH2	-8.30	116.15	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	792	PRO	Peptide
2	B	377	GLN	Peptide

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	5217	0	5252	112	0
2	B	1076	0	1090	36	0
3	A	64	0	0	2	0
All	All	6357	0	6342	139	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 11.

The worst 5 of 139 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:B:425:ARG:CZ	2:B:431:ASP:OD1	2.03	1.06
1:A:566:THR:HG21	1:A:697:LEU:HD22	1.55	0.89
1:A:801:GLU:HG2	1:A:809:ALA:HA	1.59	0.85
2:B:425:ARG:NH2	2:B:431:ASP:OD1	2.12	0.81
1:A:606:ASN:O	1:A:609:SER:O	2.03	0.77

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	664/852 (78%)	602 (91%)	57 (9%)	5 (1%)	19 58
2	B	131/485 (27%)	117 (89%)	13 (10%)	1 (1%)	19 58
All	All	795/1337 (60%)	719 (90%)	70 (9%)	6 (1%)	19 58

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	793	ILE
1	A	801	GLU
1	A	507	LYS
1	A	271	LYS
2	B	439	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	A	566/699 (81%)	539 (95%)	27 (5%)	25 61
2	B	117/397 (30%)	106 (91%)	11 (9%)	8 33
All	All	683/1096 (62%)	645 (94%)	38 (6%)	21 57

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	317	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	385	THR
2	B	343	VAL
2	B	374	GLU
2	B	426	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	419	ASN
2	B	423	ASN
1	A	484	HIS
1	A	554	GLN
1	A	632	GLN

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)

1 ligand is 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	$\#\|Z\|>2$
3	A1IG2	A	901	-	62,71,71	2.05	18 (29%)	72,111,111	1.69	14 (19%)

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
3	A1IG2	A	901	-	-	3/34/106/106	0/7/8/8

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	901	A1IG2	CAD-CAB	5.38	1.47	1.39
3	A	901	A1IG2	CAC-CAE	5.01	1.46	1.37
3	A	901	A1IG2	CAF-CAE	4.95	1.46	1.37
3	A	901	A1IG2	C4X-C4	-4.75	1.49	1.54
3	A	901	A1IG2	CAC-CAA	4.41	1.46	1.38

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	901	A1IG2	N3A-C2A-N1A	-5.23	120.50	128.68
3	A	901	A1IG2	C5A-C6A-N6A	-4.23	113.92	120.35
3	A	901	A1IG2	N6A-C6A-N1A	3.80	126.47	118.57
3	A	901	A1IG2	P-O3P-PA	-3.78	119.85	132.83
3	A	901	A1IG2	O5'-P-O1P	3.24	121.74	109.07

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	901	A1IG2	PA-O3P-P-O5'
3	A	901	A1IG2	O4'-C4'-C5'-O5'
3	A	901	A1IG2	O4B-C4B-C5B-O5B

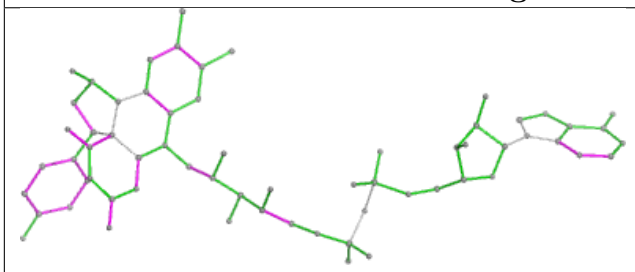
There are no ring outliers.

1 monomer is involved in 2 short contacts:

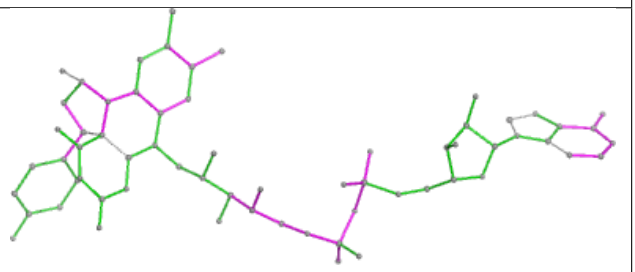
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	901	A1IG2	2	0

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.

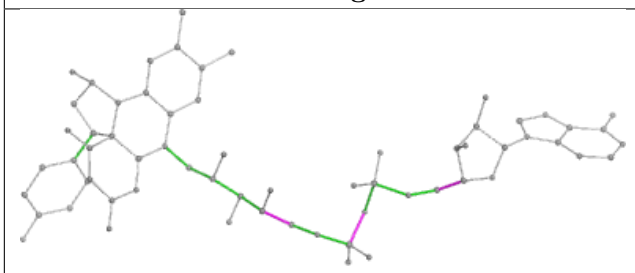
Ligand A1IG2 A 901



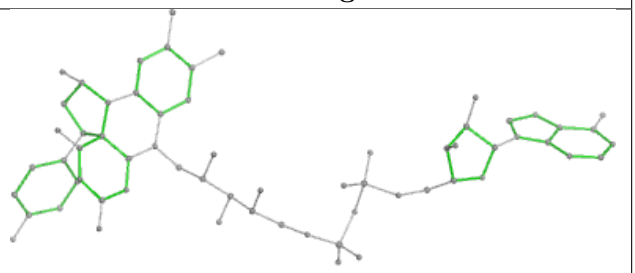
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

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	A	666/852 (78%)	0.20	2 (0%)	94	92	45, 75, 112, 139	0
2	B	133/485 (27%)	0.31	1 (0%)	86	78	76, 106, 134, 141	0
All	All	799/1337 (59%)	0.22	3 (0%)	92	89	45, 81, 120, 141	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	171	PRO	2.7
1	A	174	VAL	2.1
2	B	375	VAL	2.0

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

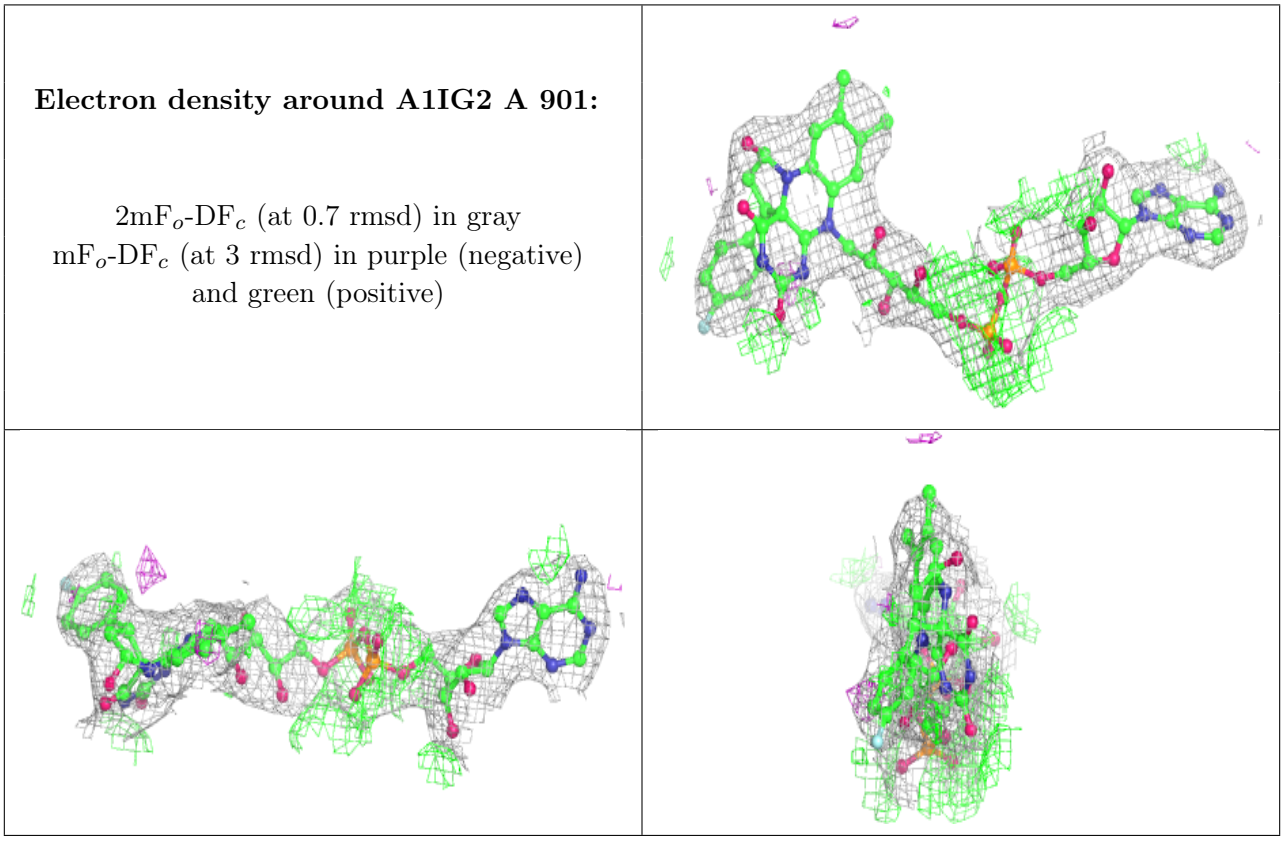
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	A1IG2	A	901	64/64	0.98	0.24	41,58,70,93	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 A1IG2 A 901:

$2mF_o$-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.