

Full wwPDB X-ray Structure Validation Report (i)

May 21, 2020 – 09:55 pm BST

PDB ID	:	2RH9
Title	:	Tryptophan synthase complexed with IGP, internal aldimine, pH 9.0
Authors	:	Kulik, V.; Barends, T.R.M.; Schlichting, I.
Deposited on		
Resolution	:	1.70 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 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 (1)) were used in the production of this report:

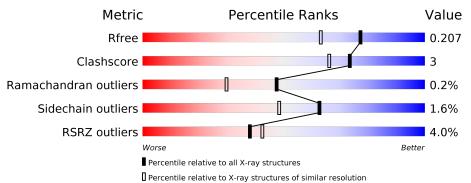
MolProbity	:	4.02b-467
e e e e e e e e e e e e e e e e e e e	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
$\operatorname{CCP4}$:	$7.0.044 (\mathrm{Gargrove})$
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.70 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries},{ m resolution\ range}({ m \AA}))$
R _{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695(1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 on 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	А	268	86%	8%	5%
2	В	397	94%		5%•



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5463 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.

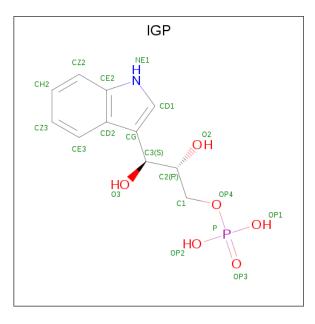
• Molecule 1 is a protein called Tryptophan synthase alpha chain.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Δ	254	Total	С	Ν	Ο	\mathbf{S}	0	6	0
	А	204	1953	1240	339	366	8	0	0	0

• Molecule 2 is a protein called Tryptophan synthase beta chain.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
2	В	394	Total 2992	C 1882	N 526	O 564	S 20	0	3	0

• Molecule 3 is INDOLE-3-GLYCEROL PHOSPHATE (three-letter code: IGP) (formula: $C_{11}H_{14}NO_6P$).



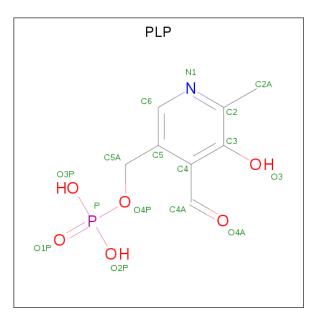
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf		
2	Δ	1	Total	С	Ν	Ο	Р	0	0
J	A	T	19	11	1	6	1	0	0

• Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	Total Na 1 1	0	0

• Molecule 5 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf		
Б	р	1	Total	С	Ν	0	Р	0	0
0	D		15	8	1	5	1	0	0

• Molecule 6 is water.

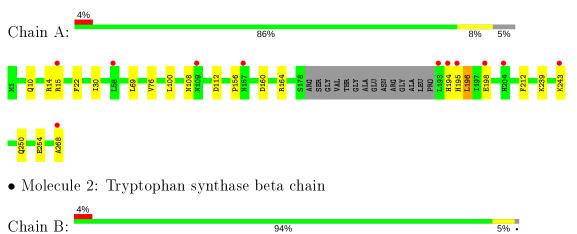
[Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
	6	А	173	Total O 173 173	0	0
	6	В	310	Total O 310 310	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Tryptophan synthase alpha chain





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	183.07Å 59.28Å 67.34Å	Depositor
a, b, c, α , β , γ	90.00° 94.67° 90.00°	Depositor
Resolution (Å)	43.73 - 1.70	Depositor
Resolution (A)	43.72 - 1.70	EDS
% Data completeness	$100.0 \ (43.73 - 1.70)$	Depositor
(in resolution range)	97.3(43.72-1.70)	EDS
R _{merge}	0.06	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.25 (at 1.70 \text{\AA})$	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
D D.	0.172 , 0.200	Depositor
R, R_{free}	0.180 , 0.207	DCC
R_{free} test set	3922 reflections $(5.08%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	18.2	Xtriage
Anisotropy	0.681	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.36 , 49.5	EDS
L-test for twinning ²	$ \langle L \rangle = 0.48, \langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5463	wwPDB-VP
Average B, all atoms $(Å^2)$	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 5.34% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: NA, IGP, PLP

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).

Mal	Mol Chain		lengths	Bond angles		
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.51	0/2003	0.66	0/2720	
2	В	0.58	0/3068	0.68	1/4144~(0.0%)	
All	All	0.55	0/5071	0.67	1/6864~(0.0%)	

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	В	206	ARG	NE-CZ-NH1	5.20	122.90	120.30

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	А	1953	0	1955	18	0
2	В	2992	0	2964	15	0
3	А	19	0	12	0	0
4	В	1	0	0	0	0
5	В	15	0	6	0	0
6	А	173	0	0	1	0
6	В	310	0	0	3	0
All	All	5463	0	4937	33	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 3.

All (33) close contacts	within th	e same	a symmetric	unit	are	listed	below,	sorted	by	their	clash
magnitude.											

Atom-1	Atom-2	Interatomic $distance (\lambda)$	$\frac{\text{Clash}}{(\hat{\lambda})}$
1.4.14[D]. ADC.IIII11		distance (Å) 1.58	$\frac{\text{overlap (Å)}}{0.68}$
1:A:14[B]:ARG:HH11 1:A:239:LYS:O	1:A:14[B]:ARG:HG2 1:A:243:LYS:HD3	1.94	0.67
2:B:379:ARG:NH1			
	6:B:561:HOH:O	2.34	0.61
1:A:194:HIS:O	1:A:198:GLU:HG2	2.01	0.60
2:B:222:ARG:NH2	6:B:556:HOH:O	2.36	0.57
2:B:337:LYS:HE2	2:B:391:LEU:HD11	1.87	0.57
1:A:30:ILE:HD11	1:A:76:VAL:HG22	1.88	0.55
2:B:337:LYS:CE	2:B:391:LEU:HD21	2.36	0.54
1:A:156[B]:PRO:HA	1:A:196:LEU:CD2	2.37	0.54
1:A:156[B]:PRO:HA	1:A:196:LEU:HD21	1.90	0.54
2:B:115:HIS:CE1	2:B:189:GLY:HA2	2.46	0.51
2:B:30:GLU:OE2	6:B:691:HOH:O	2.19	0.50
2:B:206:ARG:HD3	2:B:210:GLU:OE2	2.13	0.49
2:B:2:THR:HG22	2:B:3:THR:N	2.27	0.49
2:B:206:ARG:CD	2:B:210:GLU:OE2	2.61	0.48
2:B:337:LYS:HE2	2:B:391:LEU:HD21	1.96	0.48
1:A:160:ASP:O	1:A:164:ARG:HG3	2.15	0.47
1:A:15:ARG:O	1:A:268:ALA:HB2	2.15	0.46
1:A:108:ASN:HA	6:A:338:HOH:O	2.16	0.46
1:A:10:GLN:O	1:A:14[A]:ARG:HG3	2.17	0.45
1:A:195:HIS:HA	1:A:198:GLU:HG2	1.99	0.44
2:B:2:THR:HG22	2:B:3:THR:H	1.83	0.43
1:A:10:GLN:O	1:A:14[B]:ARG:HG3	2.18	0.43
2:B:279[B]:TYR:CD2	2:B:286:MET:HE1	2.54	0.42
2:B:286:MET:HE3	2:B:286:MET:HB3	1.31	0.42
2:B:86:HIS:CE1	2:B:236:ASN:HB3	2.55	0.41
1:A:100:LEU:HD13	1:A:100:LEU:C	2.41	0.41
1:A:22:PHE:C	1:A:22:PHE:CD1	2.94	0.41
2:B:216:ILE:HG21	2:B:224:PRO:HD3	2.03	0.41
1:A:250:GLN:O	1:A:254:GLU:HG3	2.21	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	А	256/268~(96%)	252~(98%)	3~(1%)	1 (0%)	34	18
2	В	395/397~(100%)	388~(98%)	7(2%)	0	100	100
All	All	651/665~(98%)	640 (98%)	10 (2%)	1 (0%)	47	30

All (1) Ramachandran outliers are listed below:

Mol	Chain	\mathbf{Res}	Type
1	А	212	PHE

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	А	203/208~(98%)	200~(98%)	3(2%)	65 51	
2	В	310/311~(100%)	305~(98%)	5(2%)	62 48	
All	All	513/519~(99%)	505~(98%)	8 (2%)	62 48	

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	А	69	LEU
1	А	112	ASP
1	А	196	LEU
2	В	65	ILE
2	В	206	ARG

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
2	В	286	MET
2	В	291	ASP
2	В	392	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	\mathbf{Res}	Type
1	А	80	GLN
2	В	44	GLN
2	В	246	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	Turne	Chain	Chain	Chain	Chain	Chain	Chain	Res	Link	Bo	ond leng	\mathbf{ths}	B	ond ang	les
	Type	Cham	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2					
3	IGP	А	269	-	20,20,20	1.69	2 (10%)	24,29,29	1.81	4 (16%)					
5	PLP	В	399	2	15, 15, 16	1.63	1 (6%)	20,22,23	1.67	3 (15%)					



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	IGP	А	269	-	-	1/10/14/14	0/2/2/2
5	PLP	В	399	2	-	0/6/6/8	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\operatorname{\AA})$	Ideal(Å)
3	А	269	IGP	CG-CD2	5.51	1.47	1.40
5	В	399	PLP	O3-C3	-4.95	1.25	1.37
3	А	269	IGP	CG-C3	-2.98	1.48	1.52

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
3	А	269	IGP	CG-C3-C2	-5.46	103.61	112.14
5	В	399	PLP	O4P-C5A-C5	4.70	118.31	109.35
3	А	269	IGP	CG-CD2-CE2	-4.45	102.20	106.83
3	А	269	IGP	CE3-CD2-CE2	2.98	122.12	118.17
5	В	399	PLP	C6-C5-C4	2.74	120.31	118.16
5	В	399	PLP	C5A-C5-C6	-2.28	115.62	119.37
3	А	269	IGP	O3-C3-CG	2.02	114.13	111.03

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	\mathbf{Res}	Type	Atoms
3	А	269	IGP	C2-C1-OP4-P

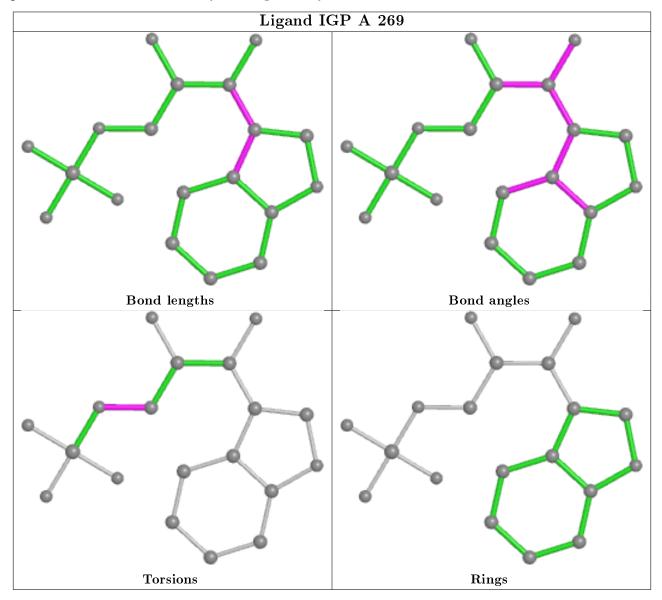
There are no ring outliers.

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 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.



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 (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	А	254/268~(94%)	-0.00	11 (4%) 35 39	15, 25, 37, 56	0
2	В	394/397~(99%)	0.08	15 (3%) 40 45	12, 18, 36, 63	0
All	All	648/665~(97%)	0.05	26 (4%) 38 42	12, 21, 37, 63	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	В	393	ALA	6.1
1	А	268	ALA	5.8
2	В	392	LYS	4.8
2	В	385	PHE	4.3
1	А	157[A]	ASN	4.0
1	А	15	ARG	3.7
1	А	194	HIS	3.7
2	В	394	ARG	3.6
1	А	195	HIS	3.4
2	В	395	GLY	3.1
1	А	193	LEU	3.0
2	В	162	GLY	2.9
2	В	159	VAL	2.9
2	В	388	HIS	2.7
2	В	390	ILE	2.6
2	В	2	THR	2.6
1	А	204	HIS	2.5
2	В	141	ARG	2.5
2	В	160	HIS	2.5
2	В	391	LEU	2.4
1	А	58	LEU	2.3
1	А	109	ASN	2.2
1	А	198	GLU	2.1
2	В	147 Continue	PHE	2.1

Continued on next page...



Continued from previous page...

Mol	Chain	\mathbf{Res}	Type	RSRZ
1	А	243	LYS	2.1
2	В	40	GLU	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)

There are no carbohydrates in this entry.

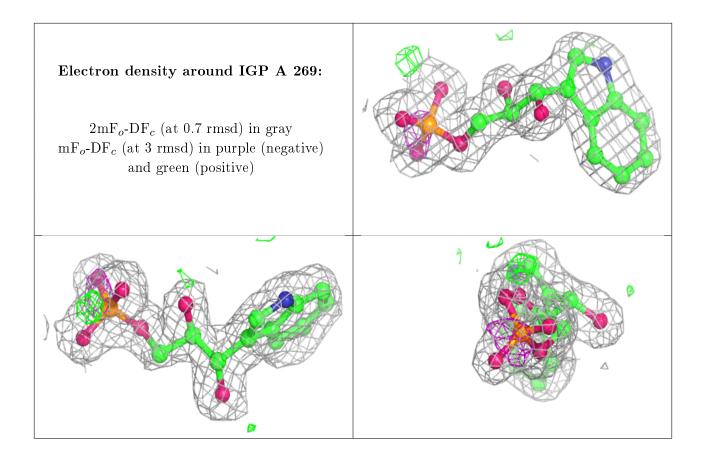
6.4 Ligands (i)

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, 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	$\mathbf{B} ext{-factors}(\mathbf{\mathring{A}}^2)$	$\mathbf{Q}{<}0.9$
3	IGP	А	269	19/19	0.95	0.11	$27,\!29,\!31,\!32$	0
5	PLP	В	399	15/16	0.98	0.13	$13,\!16,\!20,\!22$	0
4	NA	В	398	1/1	0.99	0.06	$18,\!18,\!18,\!18$	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.





6.5 Other polymers (i)

There are no such residues in this entry.

