Full wwPDB X-ray Structure Validation Report ⓘ

Sep 16, 2021 – 02:02 pm BST

PDB ID	:	7B4I
Title	:	Thermostable omega transaminase PjTA-R6 variant W58G engineered for asymmetric synthesis of enantiopure bulky amines
Authors	:	Capra, N.; Rozeboom, H.J.; Thunnissen, A.M.W.H.; Janssen, D.B.
Deposited on	:	2020-12-02
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 ⓘ symbol.

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

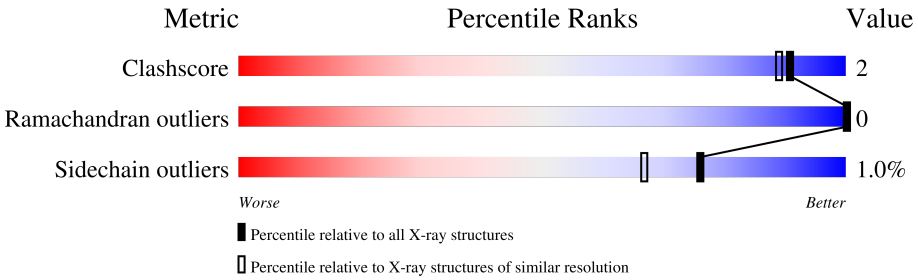
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.1

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 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (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 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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	AAA	464	93%
1	BBB	464	92% 5%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14378 atoms, of which 6959 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 Aspartate aminotransferase family protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	449	Total	C	H	N	O	S	168	0	0
			6889	2201	3444	597	631	16			
1	BBB	452	Total	C	H	N	O	S	171	5	0
			6996	2232	3495	608	643	18			

There are 32 discrepancies between the modelled and reference sequences:

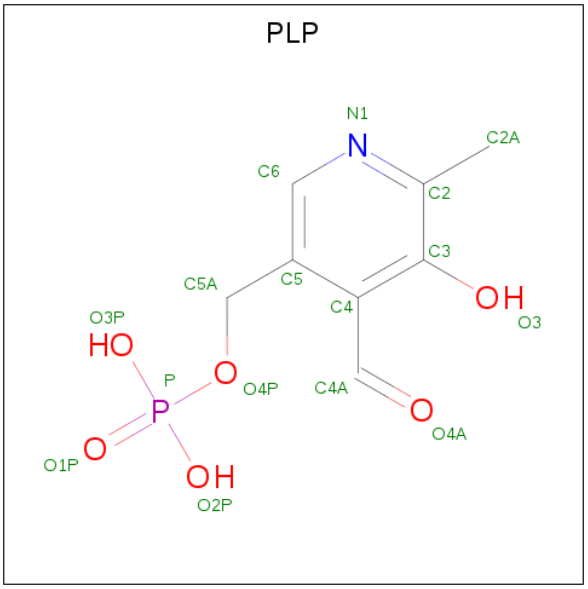
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	9	ALA	PRO	engineered mutation	UNP A0A2D8IND4
AAA	38	GLN	GLU	engineered mutation	UNP A0A2D8IND4
AAA	58	GLY	TRP	engineered mutation	UNP A0A2D8IND4
AAA	60	VAL	ALA	engineered mutation	UNP A0A2D8IND4
AAA	87	ASN	SER	engineered mutation	UNP A0A2D8IND4
AAA	128	PHE	MET	engineered mutation	UNP A0A2D8IND4
AAA	154	VAL	ILE	engineered mutation	UNP A0A2D8IND4
AAA	456	PRO	-	expression tag	UNP A0A2D8IND4
AAA	457	GLY	-	expression tag	UNP A0A2D8IND4
AAA	458	GLY	-	expression tag	UNP A0A2D8IND4
AAA	459	HIS	-	expression tag	UNP A0A2D8IND4
AAA	460	HIS	-	expression tag	UNP A0A2D8IND4
AAA	461	HIS	-	expression tag	UNP A0A2D8IND4
AAA	462	HIS	-	expression tag	UNP A0A2D8IND4
AAA	463	HIS	-	expression tag	UNP A0A2D8IND4
AAA	464	HIS	-	expression tag	UNP A0A2D8IND4
BBB	9	ALA	PRO	engineered mutation	UNP A0A2D8IND4
BBB	38	GLN	GLU	engineered mutation	UNP A0A2D8IND4
BBB	58	GLY	TRP	engineered mutation	UNP A0A2D8IND4
BBB	60	VAL	ALA	engineered mutation	UNP A0A2D8IND4
BBB	87	ASN	SER	engineered mutation	UNP A0A2D8IND4
BBB	128	PHE	MET	engineered mutation	UNP A0A2D8IND4
BBB	154	VAL	ILE	engineered mutation	UNP A0A2D8IND4
BBB	456	PRO	-	expression tag	UNP A0A2D8IND4
BBB	457	GLY	-	expression tag	UNP A0A2D8IND4

Continued on next page...

Continued from previous page...

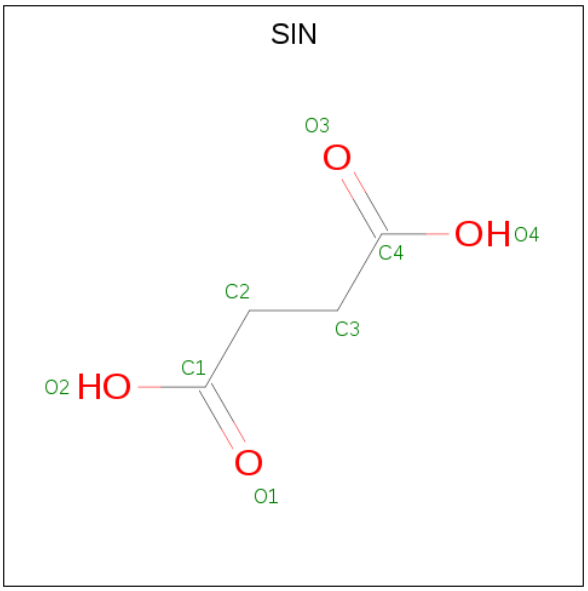
Chain	Residue	Modelled	Actual	Comment	Reference
BBB	458	GLY	-	expression tag	UNP A0A2D8IND4
BBB	459	HIS	-	expression tag	UNP A0A2D8IND4
BBB	460	HIS	-	expression tag	UNP A0A2D8IND4
BBB	461	HIS	-	expression tag	UNP A0A2D8IND4
BBB	462	HIS	-	expression tag	UNP A0A2D8IND4
BBB	463	HIS	-	expression tag	UNP A0A2D8IND4
BBB	464	HIS	-	expression tag	UNP A0A2D8IND4

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	P	1	0
			23	8	8	1	5	1		
2	BBB	1	Total	C	H	N	O	P	1	0
			23	8	8	1	5	1		

- Molecule 3 is SUCCINIC ACID (three-letter code: SIN) (formula: $C_4H_6O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	BBB	1	Total C H O 12 4 4 4	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	AAA	165	Total O 167 167	0	2
4	BBB	263	Total O 268 268	0	5

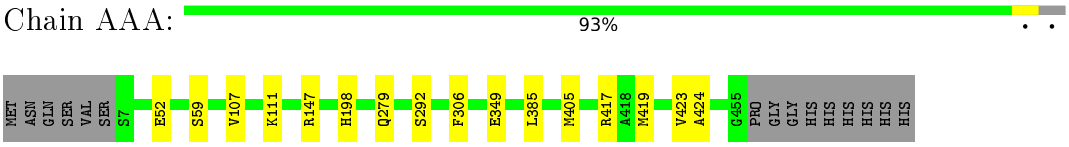
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

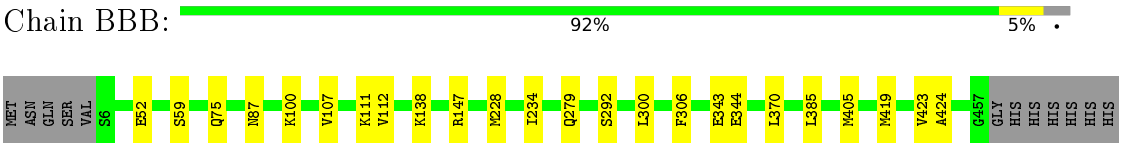
- Molecule 1: Aspartate aminotransferase family protein

Chain AAA:



- Molecule 1: Aspartate aminotransferase family protein

Chain BBB:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α, β, γ	97.78Å 97.78Å 119.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	98.00 – 1.70	Depositor
% Data completeness (in resolution range)	99.9 (98.00-1.70)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.173 , 0.191	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	14378	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SIN, 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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	AAA	0.64	0/3524	0.75	0/4774
1	BBB	0.64	0/3581	0.75	0/4850
All	All	0.64	0/7105	0.75	0/9624

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	AAA	3445	3444	3425	10	0
1	BBB	3501	3495	3471	15	0
2	AAA	15	8	6	0	0
2	BBB	15	8	6	0	0
3	BBB	8	4	4	0	0
4	AAA	167	0	0	1	0
4	BBB	268	0	0	2	0
All	All	7419	6959	6912	24	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 2.

All (24) 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:BBB:419:MET:CE	1:BBB:424:ALA:HB2	2.23	0.68
1:AAA:419:MET:CE	1:AAA:424:ALA:HB2	2.26	0.66
1:AAA:419:MET:HE1	1:AAA:424:ALA:HB2	1.90	0.54
1:BBB:419:MET:HE2	1:BBB:424:ALA:HB2	1.90	0.53
1:BBB:419:MET:HE1	1:BBB:424:ALA:HB2	1.90	0.53
1:AAA:385:LEU:HD11	1:AAA:405:MET:HG3	1.93	0.50
1:BBB:59:SER:HB2	1:BBB:292[A]:SER:OG	2.11	0.50
1:BBB:385:LEU:HD11	1:BBB:405:MET:HG3	1.94	0.50
1:BBB:228[B]:MET:HB2	1:BBB:234:ILE:HB	1.94	0.49
1:BBB:385:LEU:CD1	1:BBB:405:MET:HG3	2.42	0.49
1:AAA:59:SER:HB2	1:AAA:292:SER:OG	2.12	0.49
1:AAA:417:ARG:NH1	1:BBB:87:ASN:OD1	2.45	0.48
1:BBB:344:GLU:OE2	4:BBB:1201:HOH:O	2.20	0.46
1:AAA:107:VAL:HG11	1:AAA:279:GLN:HA	2.00	0.44
1:BBB:107:VAL:HG11	1:BBB:279:GLN:HA	1.99	0.43
1:AAA:111:LYS:HG3	1:AAA:306:PHE:CD1	2.55	0.42
1:AAA:349:GLU:CG	4:AAA:1670:HOH:O	2.68	0.42
1:BBB:138:LYS:CE	4:BBB:1368:HOH:O	2.69	0.41
1:BBB:100:LYS:NZ	1:BBB:343:GLU:OE1	2.53	0.41
1:BBB:112:VAL:HG22	1:BBB:300:LEU:HD22	2.02	0.41
1:BBB:370:LEU:O	1:BBB:385:LEU:HA	2.20	0.41
1:BBB:111:LYS:HG3	1:BBB:306:PHE:CD1	2.56	0.41
1:AAA:405:MET:HB3	1:AAA:423:VAL:HG21	2.04	0.40
1:AAA:385:LEU:CD1	1:AAA:405:MET:HG3	2.51	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	447/464 (96%)	431 (96%)	16 (4%)	0	100	100
1	BBB	455/464 (98%)	436 (96%)	19 (4%)	0	100	100
All	All	902/928 (97%)	867 (96%)	35 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	357/370 (96%)	354 (99%)	3 (1%)	81	74
1	BBB	364/370 (98%)	360 (99%)	4 (1%)	73	63
All	All	721/740 (97%)	714 (99%)	7 (1%)	76	67

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	52	GLU
1	AAA	147	ARG
1	AAA	198	HIS
1	BBB	52	GLU
1	BBB	75	GLN
1	BBB	147	ARG
1	BBB	423	VAL

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)

3 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	$\#\|Z\| > 2$
2	PLP	BBB	1102	1	15,15,16	1.44	2 (13%)	20,22,23	1.01	2 (10%)
3	SIN	BBB	1101	-	1,7,7	0.19	0	2,8,8	1.52	1 (50%)
2	PLP	AAA	1501	1	15,15,16	1.64	5 (33%)	20,22,23	1.02	2 (10%)

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
2	PLP	BBB	1102	1	-	0/6/6/8	0/1/1/1
3	SIN	BBB	1101	-	-	0/1/5/5	-
2	PLP	AAA	1501	1	-	0/6/6/8	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	1501	PLP	C6-N1	3.49	1.41	1.34
2	AAA	1501	PLP	C2-N1	3.14	1.39	1.33
2	BBB	1102	PLP	C6-N1	2.86	1.40	1.34
2	BBB	1102	PLP	C2-N1	2.53	1.38	1.33

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	1501	PLP	C5A-C5	2.16	1.56	1.50
2	AAA	1501	PLP	P-O4P	2.15	1.67	1.60
2	AAA	1501	PLP	C6-C5	2.12	1.42	1.37

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	AAA	1501	PLP	C5-C6-N1	-2.29	120.01	123.82
2	BBB	1102	PLP	C5-C6-N1	-2.19	120.17	123.82
2	BBB	1102	PLP	C6-C5-C4	2.16	119.86	118.16
3	BBB	1101	SIN	C3-C2-C1	2.15	116.28	112.67
2	AAA	1501	PLP	C6-C5-C4	2.11	119.81	118.16

There are no chirality outliers.

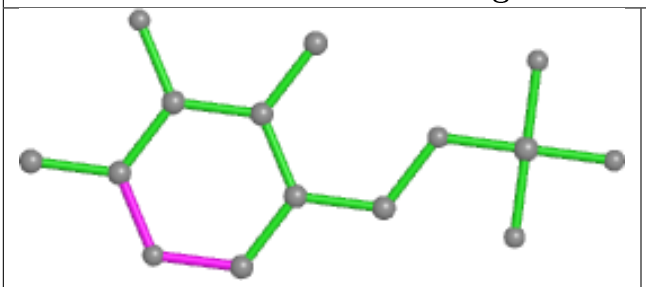
There are no torsion outliers.

There are no ring outliers.

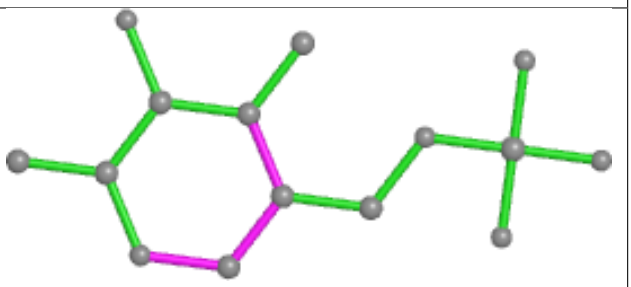
No monomer is involved in short contacts.

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.

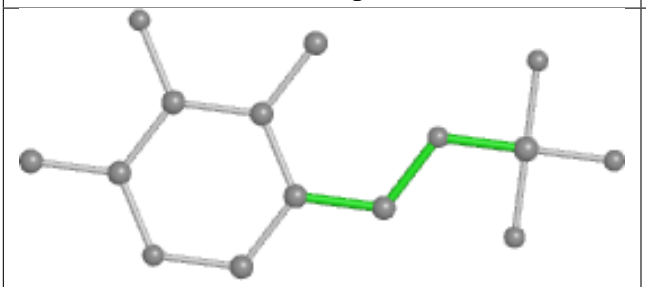
Ligand PLP BBB 1102



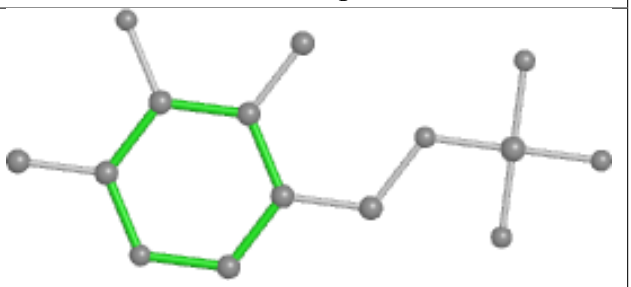
Bond lengths



Bond angles

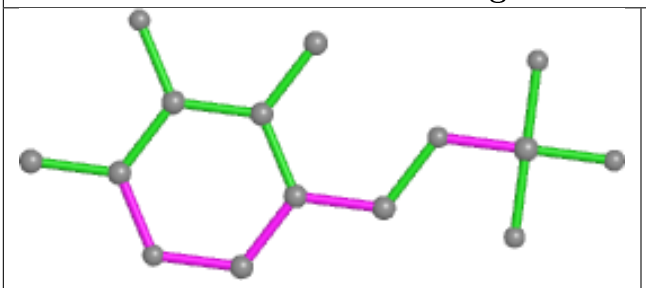


Torsions

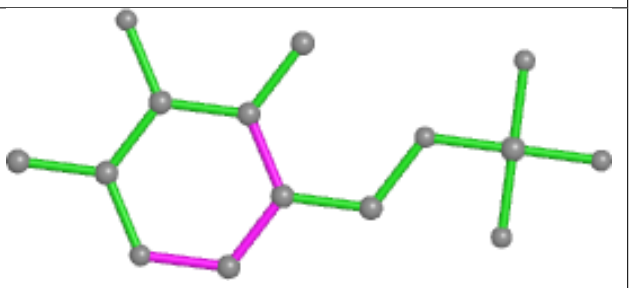


Rings

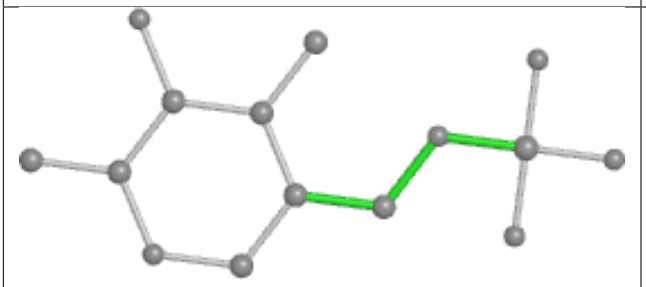
Ligand PLP AAA 1501



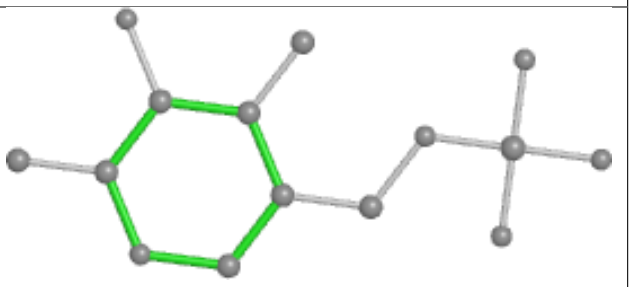
Bond lengths



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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.