



wwPDB X-ray Structure Validation Summary Report

Nov 11, 2023 – 08:15 am GMT

PDB ID : 6RGS
Title : Crystal Structure of Phenylalanine Ammonia Lyase (PAL) from *Petroselinum crispum* bound to cinnamate
Authors : Brem, J.; Lang, P.; Bencze, C.-L.; Schofield, C.
Deposited on : 2019-04-17
Resolution : 2.42 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  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](#) ) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

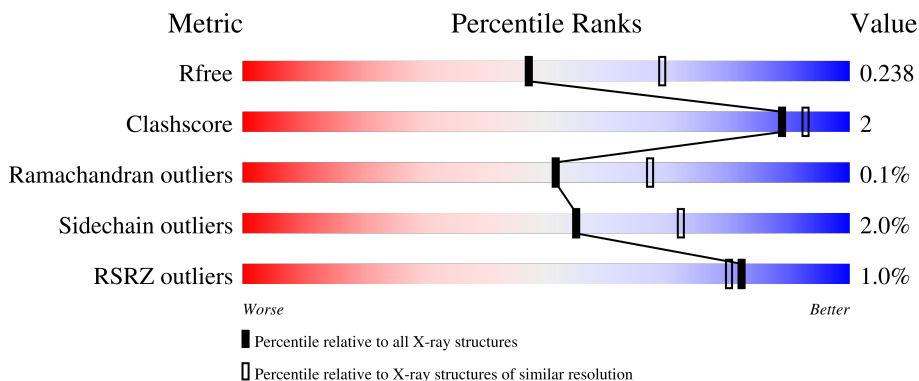
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.42 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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 $\leq 5\%$. 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	740	 81% 5% 15%
1	B	740	 79% 5% 15%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	K3Z	B	801	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 18445 atoms, of which 9030 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 Phenylalanine ammonia-lyase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	628	9259	2983	4569	785	900	22	0	0	0
1	B	626	9082	2941	4452	773	895	21	0	0	0

There are 64 discrepancies between the modelled and reference sequences:

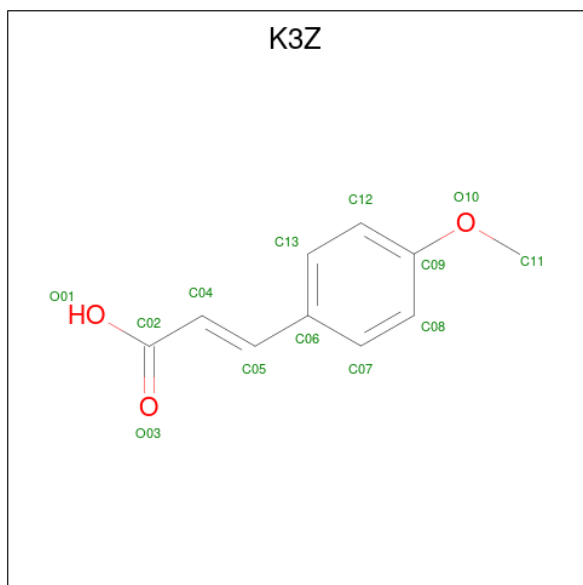
Chain	Residue	Modelled	Actual	Comment	Reference
A	-25	MET	-	initiating methionine	UNP P24481
A	-24	GLY	-	expression tag	UNP P24481
A	-23	HIS	-	expression tag	UNP P24481
A	-22	HIS	-	expression tag	UNP P24481
A	-21	HIS	-	expression tag	UNP P24481
A	-20	HIS	-	expression tag	UNP P24481
A	-19	HIS	-	expression tag	UNP P24481
A	-18	HIS	-	expression tag	UNP P24481
A	-17	HIS	-	expression tag	UNP P24481
A	-16	HIS	-	expression tag	UNP P24481
A	-15	HIS	-	expression tag	UNP P24481
A	-14	HIS	-	expression tag	UNP P24481
A	-13	SER	-	expression tag	UNP P24481
A	-12	SER	-	expression tag	UNP P24481
A	-11	GLY	-	expression tag	UNP P24481
A	-10	HIS	-	expression tag	UNP P24481
A	-9	ILE	-	expression tag	UNP P24481
A	-8	ASP	-	expression tag	UNP P24481
A	-7	ASP	-	expression tag	UNP P24481
A	-6	ASP	-	expression tag	UNP P24481
A	-5	ASP	-	expression tag	UNP P24481
A	-4	LYS	-	expression tag	UNP P24481
A	-3	HIS	-	expression tag	UNP P24481
A	-2	MET	-	expression tag	UNP P24481
A	-1	LEU	-	expression tag	UNP P24481

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLU	-	expression tag	UNP P24481
A	203	MDO	ALA	chromophore	UNP P24481
A	203	MDO	SER	chromophore	UNP P24481
A	203	MDO	GLY	chromophore	UNP P24481
A	460	VAL	ILE	engineered mutation	UNP P24481
A	704	SER	CYS	engineered mutation	UNP P24481
A	716	SER	-	expression tag	UNP P24481
B	-25	MET	-	initiating methionine	UNP P24481
B	-24	GLY	-	expression tag	UNP P24481
B	-23	HIS	-	expression tag	UNP P24481
B	-22	HIS	-	expression tag	UNP P24481
B	-21	HIS	-	expression tag	UNP P24481
B	-20	HIS	-	expression tag	UNP P24481
B	-19	HIS	-	expression tag	UNP P24481
B	-18	HIS	-	expression tag	UNP P24481
B	-17	HIS	-	expression tag	UNP P24481
B	-16	HIS	-	expression tag	UNP P24481
B	-15	HIS	-	expression tag	UNP P24481
B	-14	HIS	-	expression tag	UNP P24481
B	-13	SER	-	expression tag	UNP P24481
B	-12	SER	-	expression tag	UNP P24481
B	-11	GLY	-	expression tag	UNP P24481
B	-10	HIS	-	expression tag	UNP P24481
B	-9	ILE	-	expression tag	UNP P24481
B	-8	ASP	-	expression tag	UNP P24481
B	-7	ASP	-	expression tag	UNP P24481
B	-6	ASP	-	expression tag	UNP P24481
B	-5	ASP	-	expression tag	UNP P24481
B	-4	LYS	-	expression tag	UNP P24481
B	-3	HIS	-	expression tag	UNP P24481
B	-2	MET	-	expression tag	UNP P24481
B	-1	LEU	-	expression tag	UNP P24481
B	0	GLU	-	expression tag	UNP P24481
B	203	MDO	ALA	chromophore	UNP P24481
B	203	MDO	SER	chromophore	UNP P24481
B	203	MDO	GLY	chromophore	UNP P24481
B	460	VAL	ILE	engineered mutation	UNP P24481
B	704	SER	CYS	engineered mutation	UNP P24481
B	716	SER	-	expression tag	UNP P24481

- Molecule 2 is (E)-3-(4-methoxyphenyl)acrylic acid (three-letter code: K3Z) (formula: C₁₀H₁₀O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	B	1	22	10	9	3	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	47	Total	O	0	1
			48	48		
3	B	34	Total	O	0	0
			34	34		

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	119.91Å 159.82Å 142.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	57.31 – 2.42 79.61 – 2.42	Depositor EDS
% Data completeness (in resolution range)	99.7 (57.31-2.42) 89.7 (79.61-2.42)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.60 (at 2.42Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.210 , 0.238 0.210 , 0.238	Depositor DCC
R_{free} test set	2675 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	41.5	Xtrriage
Anisotropy	0.632	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	18445	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

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: K3Z, MDO

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	0.25	0/4759	0.41	1/6463 (0.0%)
1	B	0.25	0/4696	0.42	0/6381
All	All	0.25	0/9455	0.41	1/12844 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	205	ASP	N-CA-CB	-5.32	101.03	110.60

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	4690	4569	4561	16	0
1	B	4630	4452	4444	23	0
2	B	13	9	0	1	0
3	A	48	0	0	2	0
3	B	34	0	0	1	0
All	All	9415	9030	9005	34	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.

The worst 5 of 34 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:253:LYS:NZ	1:A:385:PRO:O	2.01	0.94
1:B:205:ASP:O	1:B:209:LEU:HB2	1.92	0.70
1:A:180:GLU:OE1	3:A:801:HOH:O	2.11	0.68
1:B:200:ILE:HD11	1:B:488:GLN:OE1	1.93	0.66
1:A:203:MDO:HB21	1:B:351:TYR:OH	2.02	0.60

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	615/740 (83%)	605 (98%)	10 (2%)	0	100	100
1	B	613/740 (83%)	596 (97%)	16 (3%)	1 (0%)	47	61
All	All	1228/1480 (83%)	1201 (98%)	26 (2%)	1 (0%)	51	67

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	591	GLN

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	480/615 (78%)	472 (98%)	8 (2%)	60	77
1	B	463/615 (75%)	452 (98%)	11 (2%)	49	67
All	All	943/1230 (77%)	924 (98%)	19 (2%)	55	72

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	484	GLU
1	B	626	LEU
1	B	697	ILE
1	B	597	LEU
1	B	25	GLU

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](#)

2 non-standard protein/DNA/RNA residues 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
1	MDO	B	203	1	12,13,14	3.24	5 (41%)	15,18,20	2.87	10 (66%)
1	MDO	A	203	1	12,13,14	3.15	6 (50%)	15,18,20	6.10	9 (60%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MDO	B	203	1	-	1/4/23/24	0/1/1/1
1	MDO	A	203	1	-	0/4/23/24	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	203	MDO	C1-N2	5.79	1.40	1.32
1	A	203	MDO	CA3-C3	5.70	1.68	1.49
1	B	203	MDO	CA2-C2	5.61	1.54	1.43
1	A	203	MDO	C1-N2	5.41	1.40	1.32
1	B	203	MDO	CA3-C3	4.77	1.65	1.49

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	203	MDO	C2-N3-C1	15.38	115.75	107.97
1	A	203	MDO	CA3-N3-C1	-10.97	114.00	127.16
1	A	203	MDO	N3-C1-N2	-8.31	105.70	111.45
1	B	203	MDO	O2-C2-CA2	-7.53	126.73	130.96
1	A	203	MDO	CA2-C2-N3	-5.70	100.68	103.37

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	203	MDO	C3-CA3-N3-C2

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	203	MDO	2	0
1	A	203	MDO	2	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry

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
2	K3Z	B	801	-	13,13,13	1.07	0	16,16,16	1.06	1 (6%)

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	K3Z	B	801	-	-	6/7/7/7	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	801	K3Z	C06-C05-C04	-2.88	120.33	126.91

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	801	K3Z	O03-C02-C04-C05
2	B	801	K3Z	O01-C02-C04-C05
2	B	801	K3Z	C12-C09-O10-C11
2	B	801	K3Z	C08-C09-O10-C11
2	B	801	K3Z	C04-C05-C06-C13

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	801	K3Z	1	0

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	A	627/740 (84%)	0.01	6 (0%) 82 80	37, 54, 77, 105	0
1	B	625/740 (84%)	0.02	6 (0%) 82 80	38, 57, 91, 110	0
All	All	1252/1480 (84%)	0.02	12 (0%) 82 80	37, 56, 86, 110	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	206	LEU	4.9
1	A	561	PHE	4.0
1	B	200	ILE	3.2
1	B	201	THR	3.1
1	B	348	GLN	3.1

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	MDO	B	203	13/14	0.88	0.21	47,54,61,65	0
1	MDO	A	203	13/14	0.93	0.17	46,51,61,61	0

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
2	K3Z	B	801	13/13	0.77	0.54	76,78,96,96	0

6.5 Other polymers [i](#)

There are no such residues in this entry.