



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 07:47 PM UTC

PDB ID : 5KQU / pdb_00005kqu
Title : Directed Evolution of Transaminases By Ancestral Reconstruction. Using Old Proteins for New Chemistries
Authors : Wilding, M.; Newman, J.; Peat, T.S.; Scott, C.
Deposited on : 2016-07-06
Resolution : 2.62 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

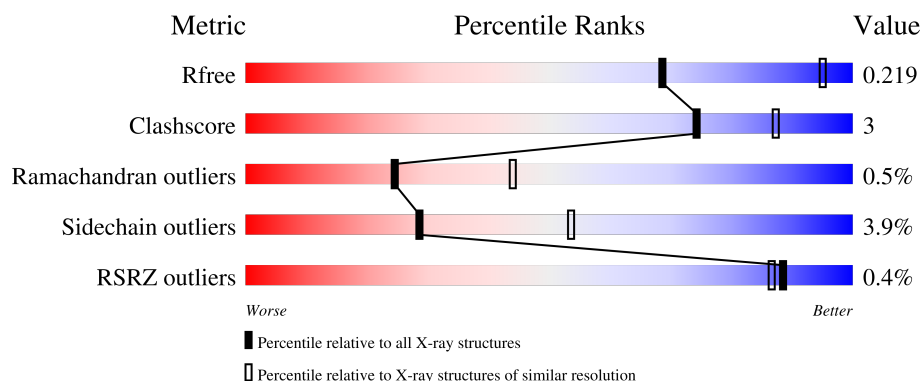
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.62 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}	180053	4951 (2.64-2.60)
Clashscore	190562	5303 (2.64-2.60)
Ramachandran outliers	187476	5217 (2.64-2.60)
Sidechain outliers	187428	5217 (2.64-2.60)
RSRZ outliers	180081	4950 (2.64-2.60)

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	479	 88% 7% • 5%
1	B	479	 88% 7% • •
1	C	479	 87% 8% • •
1	D	479	 % 85% 8% • 5%
1	E	479	 87% 8% • 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	479	

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	PLP	A	501	-	-	X	-
2	PLP	D	501	-	-	X	-
2	PLP	E	501	-	-	X	-

2 Entry composition [i](#)

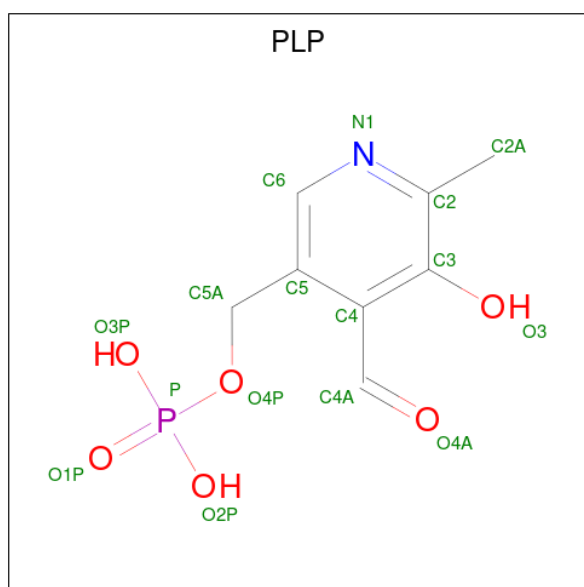
There are 3 unique types of molecules in this entry. The entry contains 21243 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 4-aminobutyrate transaminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	457	Total	C	N	O	S	0	1	0
			3493	2216	610	648	19			
1	B	458	Total	C	N	O	S	0	1	0
			3503	2221	611	652	19			
1	C	459	Total	C	N	O	S	0	1	0
			3508	2224	612	653	19			
1	D	457	Total	C	N	O	S	0	1	0
			3496	2217	610	650	19			
1	E	457	Total	C	N	O	S	0	1	0
			3496	2217	610	650	19			
1	F	459	Total	C	N	O	S	0	1	0
			3511	2226	612	653	20			

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (CCD ID: PLP) (formula: C₈H₁₀NO₆P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	E	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	F	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is water.

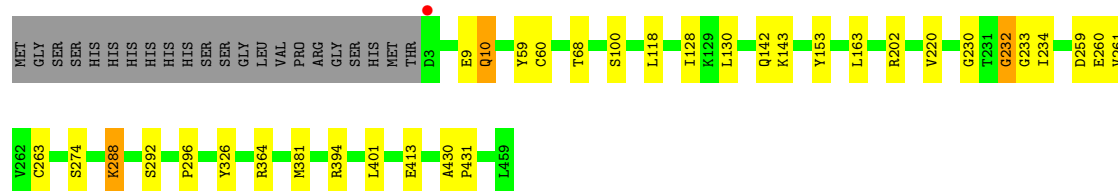
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	33	Total	O	0	0
			33	33		
3	B	31	Total	O	0	0
			31	31		
3	C	28	Total	O	0	0
			28	28		
3	D	20	Total	O	0	0
			20	20		
3	E	11	Total	O	0	0
			11	11		
3	F	23	Total	O	0	0
			23	23		

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.

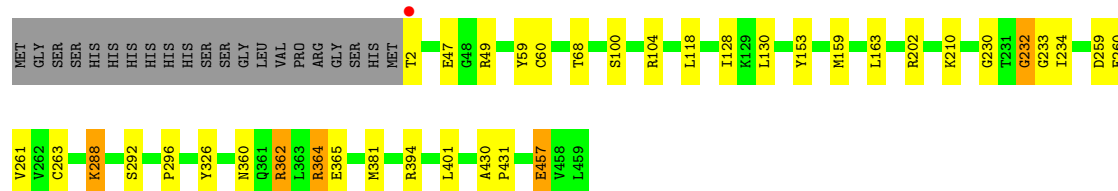
- Molecule 1: 4-aminobutyrate transaminase

Chain A: 



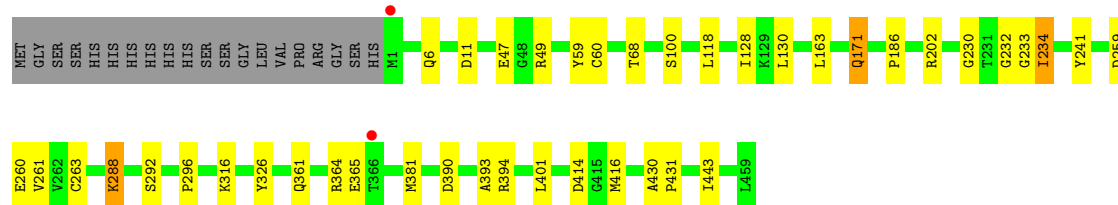
- Molecule 1: 4-aminobutyrate transaminase

Chain B: 



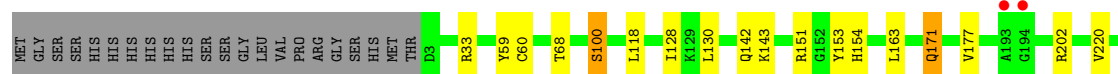
- Molecule 1: 4-aminobutyrate transaminase

Chain C: 



- Molecule 1: 4-aminobutyrate transaminase

Chain D: 





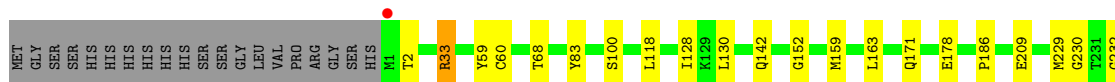
- Molecule 1: 4-aminobutyrate transaminase

Chain E: 87% 8% 5%



- Molecule 1: 4-aminobutyrate transaminase

Chain F: 87% 8% 5%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	74.10Å 146.89Å 131.78Å 90.00° 97.68° 90.00°	Depositor
Resolution (Å)	48.90 – 2.62 48.90 – 2.62	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.90-2.62) 99.6 (48.90-2.62)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.68 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.188 , 0.220 0.191 , 0.219	Depositor DCC
R_{free} test set	4128 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	38.6	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 19.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	21243	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% 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: 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	A	0.71	0/3573	0.89	1/4842 (0.0%)
1	B	0.69	0/3583	0.89	2/4856 (0.0%)
1	C	0.70	0/3588	0.90	0/4863
1	D	0.71	1/3576 (0.0%)	0.92	5/4846 (0.1%)
1	E	0.67	0/3576	0.90	1/4846 (0.0%)
1	F	0.69	0/3591	0.91	0/4866
All	All	0.70	1/21487 (0.0%)	0.90	9/29119 (0.0%)

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
1	B	0	2
1	C	0	1
1	D	0	3
1	E	0	1
1	F	0	2
All	All	0	10

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	362	ARG	C-O	-5.06	1.17	1.24

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	232	GLY	O-C-N	-5.67	115.33	122.70
1	D	448	LYS	CB-CG-CD	5.48	123.89	111.30
1	A	232	GLY	O-C-N	-5.29	115.82	122.70
1	D	232	GLY	CA-C-O	-5.23	111.47	120.57
1	E	209	GLU	CA-CB-CG	5.22	124.53	114.10

There are no chirality outliers.

5 of 10 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	232	GLY	Peptide
1	B	2	THR	Peptide
1	B	232	GLY	Peptide
1	C	232	GLY	Peptide
1	D	232	GLY	Mainchain

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	A	3493	0	3445	19	0
1	B	3503	0	3454	20	0
1	C	3508	0	3459	26	0
1	D	3496	0	3447	32	0
1	E	3496	0	3447	24	0
1	F	3511	0	3467	16	0
2	A	15	0	6	6	0
2	B	15	0	6	4	0
2	C	15	0	7	5	0
2	D	15	0	6	10	0
2	E	15	0	6	7	0
2	F	15	0	6	0	0
3	A	33	0	0	1	0
3	B	31	0	0	1	0
3	C	28	0	0	1	0
3	D	20	0	0	0	0
3	E	11	0	0	0	0
3	F	23	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	21243	0	20756	131	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.

The worst 5 of 131 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:D:259:ASP:OD2	2:D:501:PLP:C2A	1.69	1.38
1:D:259:ASP:CG	2:D:501:PLP:H2A1	1.63	1.21
1:C:416:MET:HE1	1:C:443:ILE:HA	1.24	1.17
1:E:259:ASP:CG	2:E:501:PLP:H2A2	1.76	1.09
1:D:259:ASP:CG	2:D:501:PLP:C2A	2.24	1.01

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	456/479 (95%)	435 (95%)	19 (4%)	2 (0%)	30	49
1	B	457/479 (95%)	436 (95%)	19 (4%)	2 (0%)	30	49
1	C	458/479 (96%)	437 (95%)	19 (4%)	2 (0%)	30	49
1	D	456/479 (95%)	434 (95%)	19 (4%)	3 (1%)	18	35
1	E	456/479 (95%)	435 (95%)	19 (4%)	2 (0%)	30	49
1	F	458/479 (96%)	434 (95%)	22 (5%)	2 (0%)	30	49
All	All	2741/2874 (95%)	2611 (95%)	117 (4%)	13 (0%)	24	44

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	457	GLU
1	A	288	LYS
1	D	288	LYS
1	E	288	LYS
1	F	288	LYS

5.3.2 Protein sidechains ⓘ

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	352/371 (95%)	340 (97%)	12 (3%)	32	58
1	B	354/371 (95%)	342 (97%)	12 (3%)	32	58
1	C	354/371 (95%)	342 (97%)	12 (3%)	32	58
1	D	353/371 (95%)	336 (95%)	17 (5%)	23	45
1	E	353/371 (95%)	340 (96%)	13 (4%)	30	55
1	F	355/371 (96%)	338 (95%)	17 (5%)	23	45
All	All	2121/2226 (95%)	2038 (96%)	83 (4%)	28	53

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	202	ARG
1	F	130	LEU
1	E	312	LYS
1	E	401	LEU
1	F	209	GLU

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

Mol	Chain	Res	Type
1	F	172	HIS
1	F	315	GLN
1	E	6	GLN
1	E	23	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	25	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

6 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	D	501	1	15,15,16	4.02	6 (40%)	21,22,23	3.47	9 (42%)
2	PLP	C	501	1	15,15,16	3.26	4 (26%)	21,22,23	2.36	7 (33%)
2	PLP	A	501	1	15,15,16	3.64	4 (26%)	21,22,23	3.10	9 (42%)
2	PLP	F	501	1	15,15,16	3.50	3 (20%)	21,22,23	2.14	7 (33%)
2	PLP	E	501	1	15,15,16	3.54	4 (26%)	21,22,23	2.70	9 (42%)
2	PLP	B	501	1	15,15,16	2.89	3 (20%)	21,22,23	1.70	5 (23%)

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	D	501	1	-	1/6/6/8	0/1/1/1
2	PLP	C	501	1	-	3/6/6/8	0/1/1/1
2	PLP	A	501	1	-	0/6/6/8	0/1/1/1
2	PLP	F	501	1	-	0/6/6/8	0/1/1/1
2	PLP	E	501	1	-	3/6/6/8	0/1/1/1
2	PLP	B	501	1	-	1/6/6/8	0/1/1/1

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	501	PLP	C5-C4	10.24	1.52	1.40
2	D	501	PLP	C5-C4	10.02	1.51	1.40
2	A	501	PLP	C5-C4	9.01	1.50	1.40
2	C	501	PLP	C5-C4	8.51	1.50	1.40
2	A	501	PLP	C3-C2	8.40	1.49	1.41

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	PLP	C2A-C2-C3	-7.89	111.56	120.80
2	D	501	PLP	C4A-C4-C3	7.43	132.91	120.52
2	D	501	PLP	C6-C5-C4	7.28	124.07	118.10
2	C	501	PLP	C2A-C2-C3	-6.77	112.87	120.80
2	D	501	PLP	C3-C4-C5	-6.43	110.88	118.59

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	PLP	C5A-O4P-P-O2P
2	C	501	PLP	C5A-O4P-P-O1P
2	C	501	PLP	C5A-O4P-P-O2P
2	C	501	PLP	C5A-O4P-P-O3P
2	E	501	PLP	C5A-O4P-P-O2P

There are no ring outliers.

5 monomers are involved in 32 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	501	PLP	10	0
2	C	501	PLP	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	PLP	6	0
2	E	501	PLP	7	0
2	B	501	PLP	4	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	457/479 (95%)	-0.41	1 (0%) 91 90	16, 35, 55, 73	1 (0%)
1	B	458/479 (95%)	-0.31	1 (0%) 91 90	19, 36, 58, 83	1 (0%)
1	C	459/479 (95%)	-0.28	2 (0%) 88 86	18, 35, 61, 72	1 (0%)
1	D	457/479 (95%)	-0.07	3 (0%) 84 82	22, 42, 72, 87	1 (0%)
1	E	457/479 (95%)	0.04	2 (0%) 88 86	23, 46, 71, 84	1 (0%)
1	F	459/479 (95%)	-0.24	1 (0%) 91 90	16, 39, 63, 83	1 (0%)
All	All	2747/2874 (95%)	-0.21	10 (0%) 88 86	16, 38, 64, 87	6 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1	MET	3.4
1	A	3	ASP	3.4
1	B	2	THR	3.0
1	D	193	ALA	2.4
1	E	232	GLY	2.4

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	PLP	C	501	15/16	0.90	0.12	43,51,56,58	0
2	PLP	D	501	15/16	0.90	0.13	41,62,69,73	0
2	PLP	E	501	15/16	0.90	0.12	46,59,63,66	0
2	PLP	B	501	15/16	0.93	0.09	39,42,49,50	0
2	PLP	F	501	15/16	0.93	0.11	42,44,48,51	0
2	PLP	A	501	15/16	0.94	0.10	37,44,47,47	0

6.5 Other polymers [i](#)

There are no such residues in this entry.