wwPDB X-ray Structure Validation Summary Report (i)

Mar 19, 2025 – 04:05 PM EDT

PDB ID	:	4PPM
Title	:	Crystal structure of PigE: a transaminase involved in the biosynthesis of 2-methyl-3-n-amyl-pyrrole (MAP) from Serratia sp. FS14
Authors	:	Lou, X.D.; Ran, T.T.; Xu, D.Q.; Wang, W.W.
Deposited on	:	2014-02-27
Resolution	:	2.30 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:

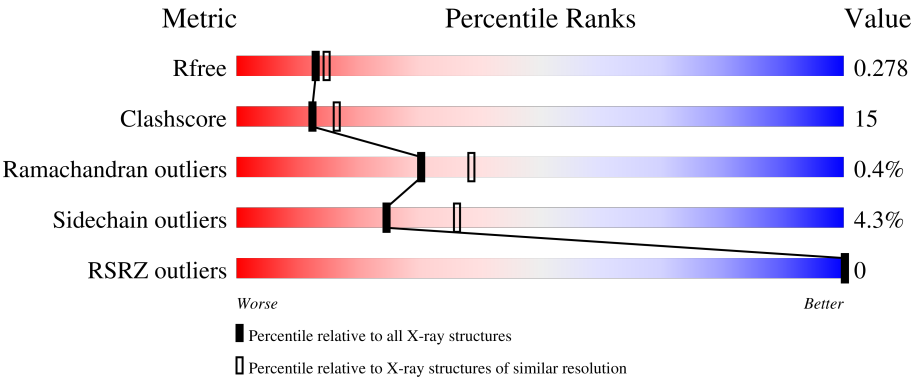
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.4

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 2.30 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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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	861	38% 15% • 46%
1	B	861	35% 17% • 46%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 7493 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 Aminotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	464	Total	C	N	O	S	0	0	0
			3563	2265	617	657	24			
1	B	462	Total	C	N	O	S	0	0	0
			3554	2263	615	652	24			

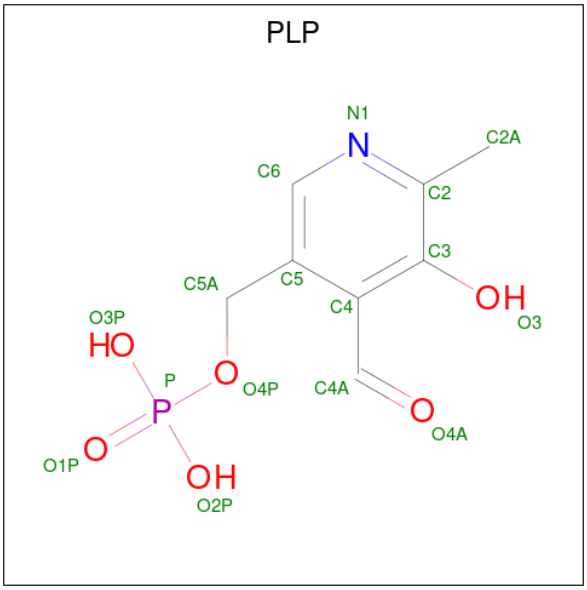
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	854	LEU	-	expression tag	UNP A0A059ZJX2
A	855	GLU	-	expression tag	UNP A0A059ZJX2
A	856	HIS	-	expression tag	UNP A0A059ZJX2
A	857	HIS	-	expression tag	UNP A0A059ZJX2
A	858	HIS	-	expression tag	UNP A0A059ZJX2
A	859	HIS	-	expression tag	UNP A0A059ZJX2
A	860	HIS	-	expression tag	UNP A0A059ZJX2
A	861	HIS	-	expression tag	UNP A0A059ZJX2
B	854	LEU	-	expression tag	UNP A0A059ZJX2
B	855	GLU	-	expression tag	UNP A0A059ZJX2
B	856	HIS	-	expression tag	UNP A0A059ZJX2
B	857	HIS	-	expression tag	UNP A0A059ZJX2
B	858	HIS	-	expression tag	UNP A0A059ZJX2
B	859	HIS	-	expression tag	UNP A0A059ZJX2
B	860	HIS	-	expression tag	UNP A0A059ZJX2
B	861	HIS	-	expression tag	UNP A0A059ZJX2

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Mg	0	0
			1	1		
2	B	1	Total	Mg	0	0
			1	1		

- Molecule 3 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
3	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 4 is water.

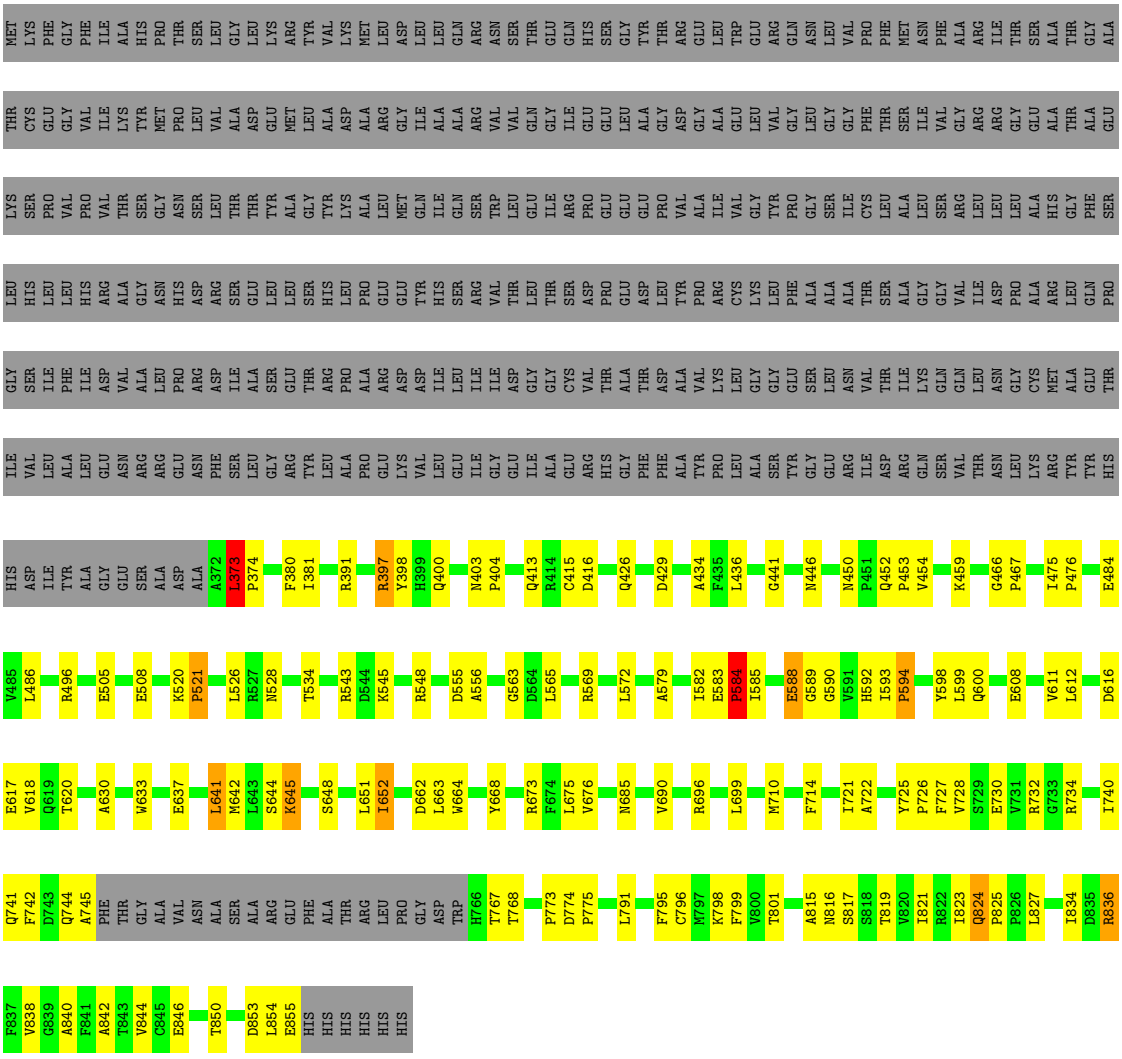
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	160	Total	O	0	0
			160	160		
4	B	184	Total	O	0	0
			184	184		

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.

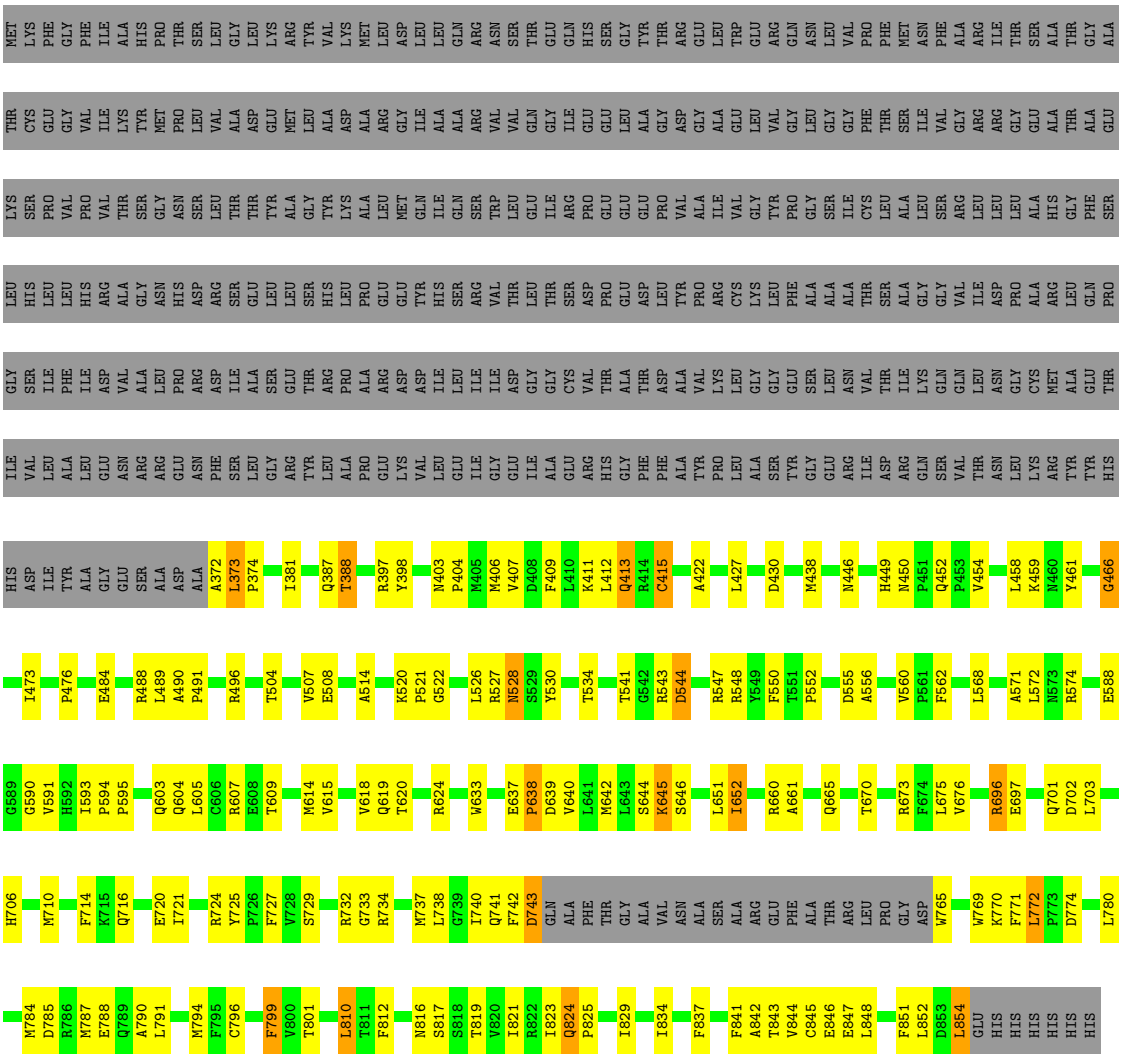
- Molecule 1: Aminotransferase

Chain A:



- Molecule 1: Aminotransferase

Chain B:


MET LYS PHE GLY PHE ILE ALA HIS PRO THR SER LEU GLY LEU LYS ARG TYR VAL LYS MET LEU ASP LEU LEU GLN ARG ASN SER THR GLU GLN HIS SER GLY TYR THR ARG GLU LEU GLN ARG GLN ASN LEU VAL PRO PHE MET ASN PHE ALA ARG ILE THR SER ALA THR GLY ALA
THR CYS GLU GLY VAL ILE LYS TYR MET PRO LEU VAL ALA ASP GLU MET LEU ALA ASP ALA ARG GLY ILE ALA ALA ARG VAL VAL GLN GLY ILE GLU GLU LEU ALA GLY ASP GLY ALA GLU LEU VAL GLY LEU GLY GLY PHE THR SER ILE VAL GLY ARG ARG GLY GLU ALA THR ALA GLU
LYS SER PRO VAL PRO VAL THR SER GLY ASN SER LEU THR THR TYR ALA GLY TYR LYS ALA LEU MET GLN ILE GLN SER TRP LEU GLU ILE ARG PRO GLU GLU GLU PRO VAL ALA ILE VAL GLY TYR PRO GLY SER ILE CYS LEU ALA LEU LEU LEU LEU ASP PRO ALA ARG LEU LEU GLN PRO
LEU HIS LEU LEU HIS ARG ALA GLY ASN HIS ASP ARG SER GLU LEU LEU SER HIS LEU PRO GLU GLU TYR HIS SER ARG VAL THR LEU THR SER ASP PRO GLU GLU ASP LEU TYR PRO ARG CYS LYS LEU PHE ALA ALA ALA THR SER ALA LEU GLY VAL ILE ASP PRO ALA ARG LEU GLN PRO
GLY SER ILE PHE ILE ASP VAL ALA LEU PRO ARG ASP ILE ALA SER GLU THR ARG PRO ALA ARG ASP ILE LEU LEU ILE ILE ASP GLY GLY CYS VAL THR ALA THR ASP ALA VAL LYS LEU GLY GLY GLU SER LEU ASN VAL THR ILE LYS GLN GLN LEU LEU ASN GLY CYS MET ALA GLU THR
ILE VAL LEU ALA LEU GLU ASN ARG ARG GLU ASN PHE SER LEU GLY ARG TYR LEU ALA PRO GLU LYS VAL LEU GLU ILE GLY GLU ILE ALA GLU ARG HIS GLY PHE PHE ALA TYR PRO LEU ALA SER TYR GLY GLU ARG ILE ASP ARG GLN SER VAL THR ASN LEU LYS ARG TYR TYR HIS
HIS ASP ILE TYR ALA GLY GLU SER ALA ASP ALA A372 L373 P374 I381 Q387 T388 R397 Y398 N403 P404 M405 M406 V407 D408 F409 L410 K411 L412 Q413 R414 C415 A422 L427 D430 M438 N446 H449 N450 P451 Q452 P453 V454 L458 K459 N460 Y461 G466
I473 P476 E484 R488 L489 A490 P491 R496 T504 V507 E508 A514 K520 P521 G522 L526 R527 N528 S529 Y530 T534 T541 G542 R543 D544 R547 R548 Y549 F550 T551 P552 D555 A556 V560 P561 F562 L568 A571 L572 N573 R574 E588
G589 G590 V591 H592 I593 P594 P595 Q603 Q604 L605 C606 R607 E608 T609 M614 V615 V618 Q619 T620 R624 W633 E637 P638 D639 V640 L641 M642 L643 S644 K645 S646 L651 I652 R660 A661 Q665 T670 R673 F674 L675 V676 R696 E697 Q701 D702 L703
H706 M710 F714 K715 Q716 E720 I721 R724 Y725 P726 F727 V728 S729 R732 G733 R734 M737 L738 G739 I740 Q741 F742 D743 GLN ALA PHE THR GLY ALA VAL ASN ALA SER ALA ARG GLU PHE ALA THR ARG LEU PRO GLY ASP W765 W769 K770 F771 L772 P773 D774 L780
M784 D785 R786 M787 E788 Q789 A790 L791 M794 F795 C796 F799 V800 T801 L810 T811 F812 N816 S817 S818 T819 V820 I821 R822 I823 Q824 P825 I829 I834 F837 F841 A842 T843 V844 C845 E846 E847 L848 F851 L852 D853 L854 GLU HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	228.10Å 228.10Å 67.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.89 – 2.30 19.89 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.89-2.30) 100.0 (19.89-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.77 (at 2.30Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.238 , 0.272 0.251 , 0.278	Depositor DCC
R_{free} test set	4440 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	37.2	Xtriage
Anisotropy	0.092	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 18.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.38$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	0.340 for -h,-k,l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	7493	wwPDB-VP
Average B, all atoms (Å^2)	39.0	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: MG, 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	1.41	7/3633 (0.2%)	0.94	19/4920 (0.4%)
1	B	1.36	6/3626 (0.2%)	0.95	20/4912 (0.4%)
All	All	1.39	13/7259 (0.2%)	0.95	39/9832 (0.4%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	505	GLU	CD-OE2	-5.65	1.19	1.25
1	A	508	GLU	CD-OE1	-5.60	1.19	1.25
1	A	508	GLU	CD-OE2	-5.57	1.19	1.25
1	B	404	PRO	N-CD	5.52	1.55	1.47
1	B	552	PRO	N-CD	5.44	1.55	1.47

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	702	ASP	CB-CG-OD1	7.09	124.68	118.30
1	A	590	GLY	N-CA-C	6.99	130.58	113.10
1	A	593	ILE	C-N-CD	6.29	141.61	128.40
1	B	725	TYR	C-N-CD	6.16	141.34	128.40
1	A	594	PRO	C-N-CD	6.02	141.04	128.40

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	A	3563	0	3567	100	0
1	B	3554	0	3558	132	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	15	0	6	0	0
3	B	15	0	6	0	0
4	A	160	0	0	5	0
4	B	184	0	0	6	0
All	All	7493	0	7137	216	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 15.

The worst 5 of 216 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:727:PHE:HB3	1:A:744:GLN:NE2	1.44	1.32
1:A:588:GLU:O	1:A:816:ASN:HB2	1.43	1.14
1:B:388:THR:HG22	4:B:1037:HOH:O	1.55	1.06
1:A:727:PHE:CB	1:A:744:GLN:HE21	1.69	1.05
1:A:727:PHE:HB3	1:A:744:GLN:HE21	0.89	1.04

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	460/861 (53%)	432 (94%)	26 (6%)	2 (0%)	30 39
1	B	458/861 (53%)	440 (96%)	16 (4%)	2 (0%)	30 39

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	918/1722 (53%)	872 (95%)	42 (5%)	4 (0%)	30 39

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	645	LYS
1	B	645	LYS
1	A	644	SER
1	B	644	SER

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	372/692 (54%)	360 (97%)	12 (3%)	34 50
1	B	371/692 (54%)	351 (95%)	20 (5%)	18 27
All	All	743/1384 (54%)	711 (96%)	32 (4%)	25 36

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	810	LEU
1	B	812	PHE
1	B	413	GLN
1	A	855	GLU
1	B	845	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	824	GLN
1	B	665	GLN
1	A	777	GLN
1	A	744	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	805	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 2 are 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	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	PLP	B	902	1	15,15,16	2.69	8 (53%)	21,22,23	1.65	5 (23%)
3	PLP	A	902	1	15,15,16	2.88	9 (60%)	21,22,23	1.91	6 (28%)

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	PLP	B	902	1	-	1/6/6/8	0/1/1/1
3	PLP	A	902	1	-	2/6/6/8	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	902	PLP	C5-C4	5.92	1.47	1.40
3	B	902	PLP	C5-C4	5.44	1.46	1.40
3	A	902	PLP	C3-C2	4.12	1.45	1.41
3	B	902	PLP	C3-C2	4.06	1.45	1.41
3	A	902	PLP	P-O3P	-3.91	1.40	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	902	PLP	C4A-C4-C5	5.04	126.13	120.94
3	B	902	PLP	O3P-P-O2P	3.13	119.54	107.80
3	B	902	PLP	O4P-C5A-C5	2.77	114.54	109.36
3	A	902	PLP	O3P-P-O2P	2.66	117.79	107.80
3	B	902	PLP	C4A-C4-C5	2.62	123.64	120.94

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	902	PLP	C6-C5-C5A-O4P
3	A	902	PLP	C4-C5-C5A-O4P
3	B	902	PLP	C5A-O4P-P-O1P

There are no ring outliers.

No monomer is involved in short contacts.

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

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	464/861 (53%)	-1.52	0	100	100	17, 33, 60, 93	0
1	B	462/861 (53%)	-1.49	0	100	100	17, 38, 68, 96	0
All	All	926/1722 (53%)	-1.50	0	100	100	17, 36, 65, 96	0

There are no RSRZ outliers to report.

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
2	MG	A	901	1/1	0.97	0.04	30,30,30,30	0
2	MG	B	901	1/1	0.99	0.04	30,30,30,30	0
3	PLP	A	902	15/16	1.00	0.02	21,26,33,35	0
3	PLP	B	902	15/16	1.00	0.03	24,29,33,34	0

6.5 Other polymers (i)

There are no such residues in this entry.