



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:38 am GMT

PDB ID : 1EQB
Title : X-RAY CRYSTAL STRUCTURE AT 2.7 ANGSTROMS RESOLUTION OF
TERNARY COMPLEX BETWEEN THE Y65F MUTANT OF E-COLI SER-
INE HYDROXYMETHYLTRANSFERASE, GLYCINE AND 5-FORMYL
TETRAHYDROFOLATE
Authors : Contestabile, R.; Angelaccio, S.; Bossa, F.; Wright, H.T.; Scarsdale, N.;
Kazanina, G.; Schirch, V.
Deposited on : 2000-04-03
Resolution : 2.70 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

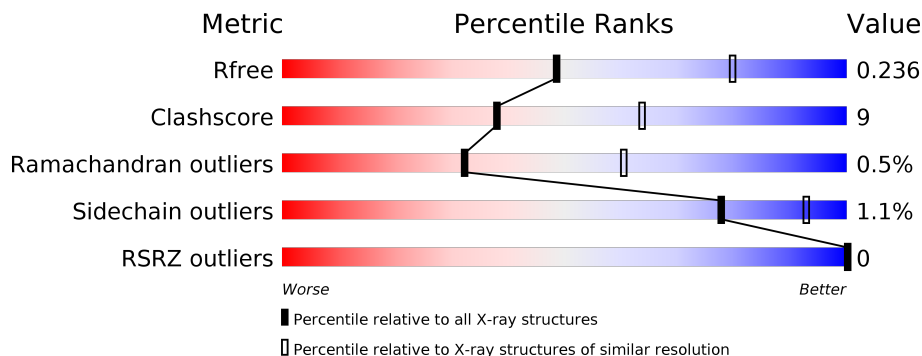
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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))
R_{free}	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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. 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	417	<div> <div>81%</div> <div>18%</div> <div>.</div> </div>
1	B	417	<div> <div>82%</div> <div>17%</div> <div>.</div> </div>
1	C	417	<div> <div>81%</div> <div>18%</div> <div>.</div> </div>
1	D	417	<div> <div>83%</div> <div>16%</div> <div>.</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	FFO	D	4293	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13171 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.

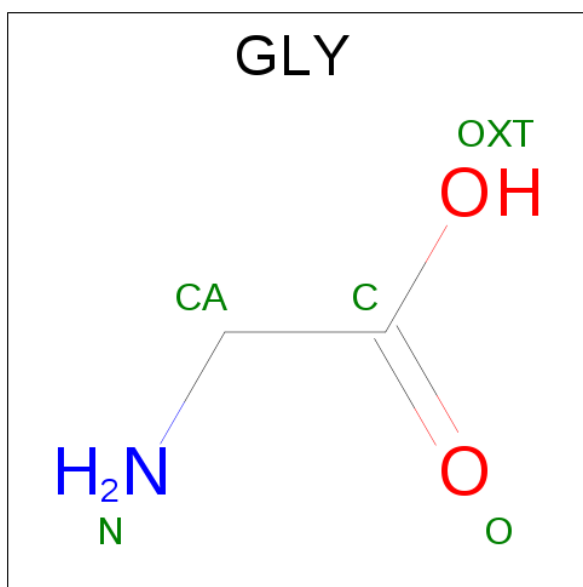
- Molecule 1 is a protein called SERINE HYDROXYMETHYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	S	0	0	0
			3161	2003	545	599	14			
1	B	417	Total	C	N	O	S	0	1	0
			3186	2017	547	607	15			
1	C	417	Total	C	N	O	S	0	0	0
			3161	2002	544	600	15			
1	D	417	Total	C	N	O	S	0	0	0
			3160	2003	541	601	15			

There are 4 discrepancies between the modelled and reference sequences:

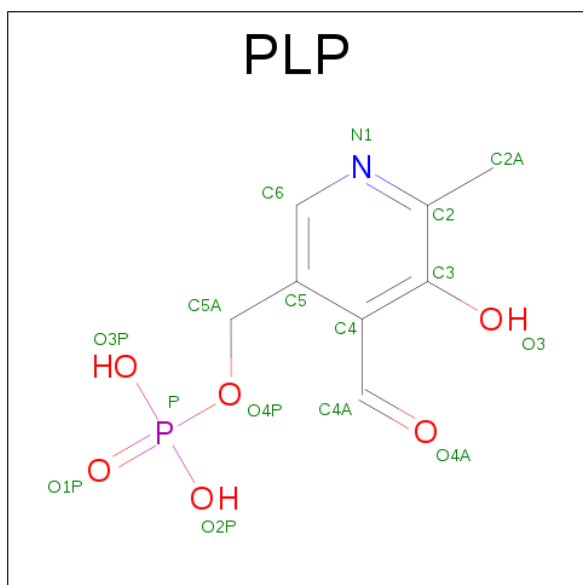
Chain	Residue	Modelled	Actual	Comment	Reference
A	65	PHE	TYR	ENGINEERED	UNP P00477
B	65	PHE	TYR	ENGINEERED	UNP P00477
C	65	PHE	TYR	ENGINEERED	UNP P00477
D	65	PHE	TYR	ENGINEERED	UNP P00477

- Molecule 2 is GLYCINE (three-letter code: GLY) (formula: C₂H₅NO₂).



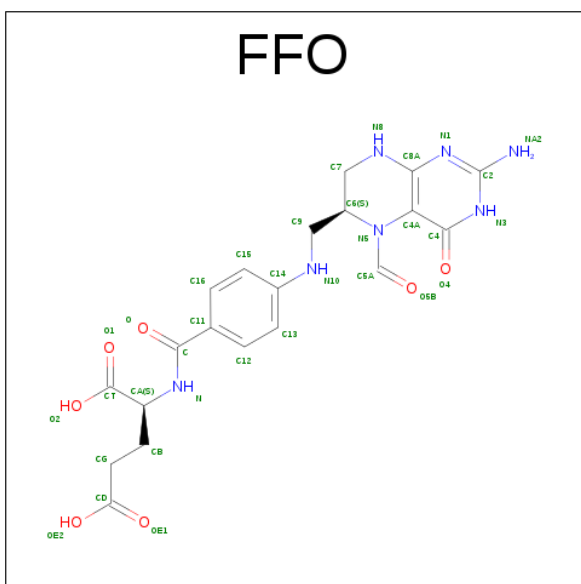
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			5	2	1	2		
2	B	1	Total	C	N	O	0	0
			5	2	1	2		
2	C	1	Total	C	N	O	0	0
			5	2	1	2		
2	D	1	Total	C	N	O	0	0
			5	2	1	2		

- Molecule 3 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: C₈H₁₀NO₆P).



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

- Molecule 4 is N-[4-({[(6S)-2-AMINO-5-FORMYL-4-OXO-3,4,5,6,7,8-HEXAHYDROPTERIDIN-6-YL]METHYL}AMINO)BENZOYL]-L-GLUTAMIC ACID (three-letter code: FFO) (formula: $C_{20}H_{23}N_7O_7$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 34	C 20	N 7	O 7	0	0
4	B	1	Total 34	C 20	N 7	O 7	0	0
4	C	1	Total 34	C 20	N 7	O 7	0	0
4	D	1	Total 34	C 20	N 7	O 7	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	70	Total O 70 70	0	0

Continued on next page...

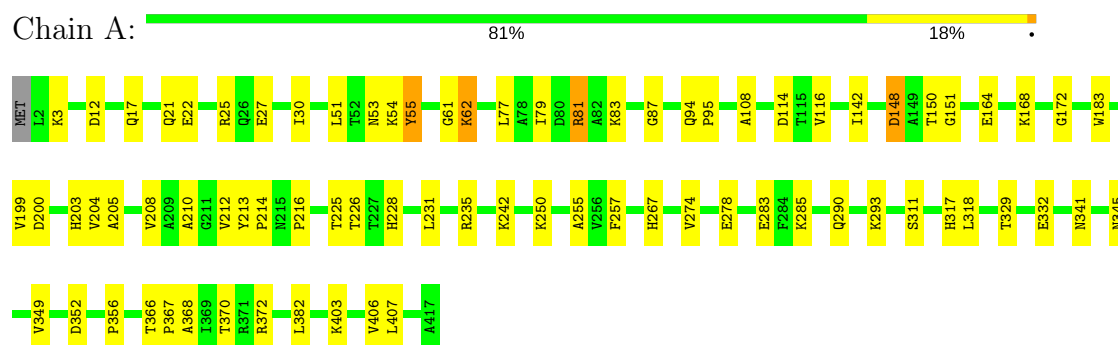
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	71	Total 71	O 71	0	0
5	C	73	Total 73	O 73	0	0
5	D	73	Total 73	O 73	0	0

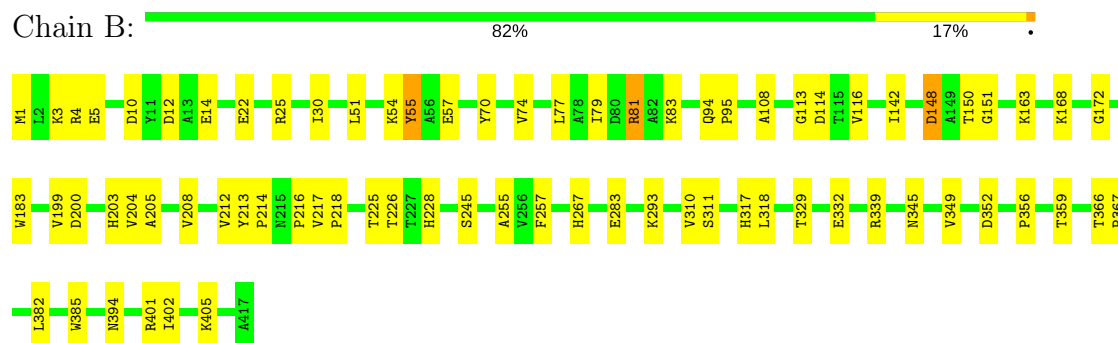
3 Residue-property plots [i](#)

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.

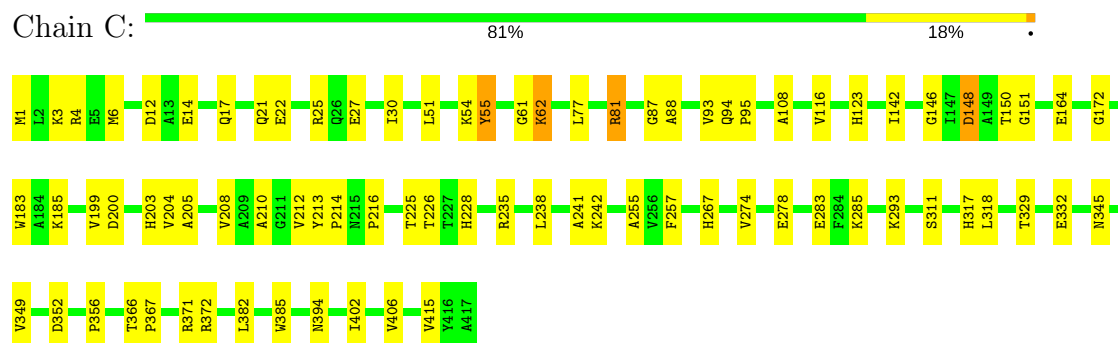
• Molecule 1: SERINE HYDROXYMETHYLTRANSFERASE



• Molecule 1: SERINE HYDROXYMETHYLTRANSFERASE

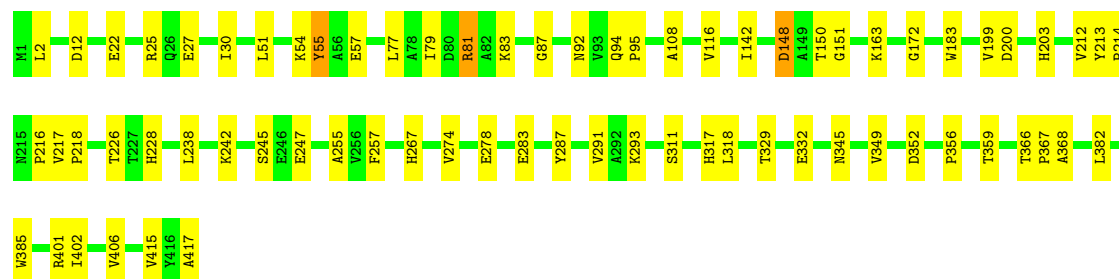


• Molecule 1: SERINE HYDROXYMETHYLTRANSFERASE



• Molecule 1: SERINE HYDROXYMETHYLTRANSFERASE

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	77.74Å 172.48Å 95.41Å 90.00° 104.28° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 92.46 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.6 (20.00-2.70) 94.6 (92.46-2.70)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.65 (at 2.69Å)	Xtriage
Refinement program	CNS 0.5	Depositor
R, R_{free}	0.205 , 0.236 0.204 , 0.236	Depositor DCC
R_{free} test set	6375 reflections (10.12%)	DCC
Wilson B-factor (Å ²)	35.1	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 45.5	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	0.92	EDS
Total number of atoms	13171	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.65% 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: FFO, 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	0.36	0/3229	0.59	0/4377
1	B	0.36	0/3254	0.59	0/4409
1	C	0.35	0/3229	0.59	0/4379
1	D	0.36	0/3228	0.59	0/4377
All	All	0.36	0/12940	0.59	0/17542

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 [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	3161	0	3107	57	0
1	B	3186	0	3132	57	0
1	C	3161	0	3097	63	0
1	D	3160	0	3097	56	0
2	A	5	0	2	1	0
2	B	5	0	2	1	0
2	C	5	0	2	1	0
2	D	5	0	2	1	0
3	A	15	0	6	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	15	0	6	1	0
3	C	15	0	6	1	0
3	D	15	0	6	1	0
4	A	34	0	21	4	0
4	B	34	0	21	2	0
4	C	34	0	21	3	0
4	D	34	0	21	2	0
5	A	70	0	0	1	0
5	B	71	0	0	3	0
5	C	73	0	0	3	0
5	D	73	0	0	4	0
All	All	13171	0	12549	221	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 9.

The worst 5 of 221 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:B:339:ARG:HD2	5:B:1431:HOH:O	1.76	0.86
1:A:108:ALA:HB2	1:A:255:ALA:HB2	1.75	0.68
1:B:77:LEU:O	1:B:81:ARG:HG2	1.94	0.68
1:D:108:ALA:HB2	1:D:255:ALA:HB2	1.76	0.66
1:C:108:ALA:HB2	1:C:255:ALA:HB2	1.77	0.66

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	A	414/417 (99%)	398 (96%)	14 (3%)	2 (0%)	32 60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	416/417 (100%)	402 (97%)	12 (3%)	2 (0%)	32	60
1	C	415/417 (100%)	399 (96%)	15 (4%)	1 (0%)	51	79
1	D	415/417 (100%)	398 (96%)	14 (3%)	3 (1%)	25	53
All	All	1660/1668 (100%)	1597 (96%)	55 (3%)	8 (0%)	32	60

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	LYS
1	A	148	ASP
1	B	148	ASP
1	C	148	ASP
1	D	2	LEU

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	323/329 (98%)	319 (99%)	4 (1%)	75	92
1	B	327/329 (99%)	324 (99%)	3 (1%)	82	94
1	C	322/329 (98%)	318 (99%)	4 (1%)	75	92
1	D	322/329 (98%)	319 (99%)	3 (1%)	82	94
All	All	1294/1316 (98%)	1280 (99%)	14 (1%)	78	93

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	345	ASN
1	C	55	TYR
1	D	55	TYR
1	B	81	ARG
1	C	345	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	394	ASN
1	C	203	HIS
1	D	289	GLN
1	C	17	GLN
1	C	219	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

12 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	$\# Z > 2$
3	PLP	A	1291	2	15,15,16	1.73	4 (26%)	20,22,23	2.41	8 (40%)
2	GLY	A	1292	3	1,4,4	0.42	0	0,4,4	0.00	-
4	FFO	A	1293	-	24,36,36	2.40	10 (41%)	29,50,50	3.69	13 (44%)
3	PLP	B	2291	2	15,15,16	1.75	5 (33%)	20,22,23	2.40	8 (40%)
2	GLY	B	2292	3	1,4,4	0.45	0	0,4,4	0.00	-
4	FFO	B	2293	-	24,36,36	2.37	10 (41%)	29,50,50	3.67	13 (44%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	PLP	C	3291	2	15,15,16	1.67	4 (26%)	20,22,23	2.41	8 (40%)
2	GLY	C	3292	3	1,4,4	0.44	0	0,4,4	0.00	-
4	FFO	C	3293	-	24,36,36	2.30	10 (41%)	29,50,50	3.69	14 (48%)
3	PLP	D	4291	2	15,15,16	1.66	4 (26%)	20,22,23	2.42	8 (40%)
2	GLY	D	4292	3	1,4,4	0.42	0	0,4,4	0.00	-
4	FFO	D	4293	-	24,36,36	2.42	10 (41%)	29,50,50	3.71	14 (48%)

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	A	1291	2	-	0/6/6/8	0/1/1/1
2	GLY	A	1292	3	-	0/0/2/2	0/0/0/0
4	FFO	A	1293	-	-	0/18/37/37	0/2/3/3
3	PLP	B	2291	2	-	0/6/6/8	0/1/1/1
2	GLY	B	2292	3	-	0/0/2/2	0/0/0/0
4	FFO	B	2293	-	-	0/18/37/37	0/2/3/3
3	PLP	C	3291	2	-	0/6/6/8	0/1/1/1
2	GLY	C	3292	3	-	0/0/2/2	0/0/0/0
4	FFO	C	3293	-	-	0/18/37/37	0/2/3/3
3	PLP	D	4291	2	-	0/6/6/8	0/1/1/1
2	GLY	D	4292	3	-	0/0/2/2	0/0/0/0
4	FFO	D	4293	-	-	0/18/37/37	0/2/3/3

The worst 5 of 57 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	3291	PLP	C4A-C4	-2.81	1.45	1.51
3	A	1291	PLP	O3-C3	-2.69	1.30	1.37
3	B	2291	PLP	C4A-C4	-2.57	1.46	1.51
3	D	4291	PLP	C4A-C4	-2.57	1.46	1.51
3	A	1291	PLP	C4A-C4	-2.46	1.46	1.51

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1293	FFO	N3-C2-N1	-6.84	114.36	125.45
4	B	2293	FFO	N3-C2-N1	-6.80	114.43	125.45
4	D	4293	FFO	N3-C2-N1	-6.78	114.46	125.45

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	3293	FFO	N3-C2-N1	-6.74	114.52	125.45
4	C	3293	FFO	C16-C15-C14	-5.41	113.97	120.30

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1291	PLP	1	0
2	A	1292	GLY	1	0
4	A	1293	FFO	4	0
3	B	2291	PLP	1	0
2	B	2292	GLY	1	0
4	B	2293	FFO	2	0
3	C	3291	PLP	1	0
2	C	3292	GLY	1	0
4	C	3293	FFO	3	0
3	D	4291	PLP	1	0
2	D	4292	GLY	1	0
4	D	4293	FFO	2	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	416/417 (99%)	-0.11	0 100 100	16, 34, 59, 93	0
1	B	417/417 (100%)	-0.09	0 100 100	13, 32, 59, 84	0
1	C	417/417 (100%)	-0.08	0 100 100	15, 37, 67, 87	0
1	D	417/417 (100%)	-0.11	0 100 100	15, 31, 62, 99	0
All	All	1667/1668 (99%)	-0.10	0 100 100	13, 33, 62, 99	0

There are no RSRZ outliers to report.

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
4	FFO	D	4293	34/34	0.92	0.21	2.25	46,46,46,46	0
2	GLY	B	2292	5/5	0.95	0.23	1.96	32,46,46,46	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	FFO	B	2293	34/34	0.92	0.22	1.34	49,49,49,49	0
3	PLP	A	1291	15/16	0.98	0.19	0.96	26,26,26,26	0
4	FFO	C	3293	34/34	0.93	0.21	0.80	43,43,43,43	0
3	PLP	B	2291	15/16	0.97	0.18	0.46	31,31,31,31	0
3	PLP	C	3291	15/16	0.97	0.19	0.40	32,32,32,32	0
2	GLY	A	1292	5/5	0.97	0.20	0.35	24,31,31,31	0
4	FFO	A	1293	34/34	0.95	0.20	0.30	37,37,37,37	0
3	PLP	D	4291	15/16	0.97	0.17	-0.19	34,34,34,34	0
2	GLY	C	3292	5/5	0.98	0.18	-0.37	19,31,31,31	0
2	GLY	D	4292	5/5	0.98	0.13	-2.75	26,39,39,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.