W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 11:53 AM EDT

PDB ID	:	1K8Z
Title	:	CRYSTAL STRUCTURE OF THE TRYPTOPHAN SYNTHASE BETA-SER178PRO MUTANT COMPLEXED WITH N-[1H-INDOL-3-YL-ACE TYL]GLYCINE ACID
Authors	:	Weyand, M.; Schlichting, I.; Marabotti, A.; Mozzarelli, A.
Deposited on	:	2001-10-26
Resolution	:	1.70 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
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:

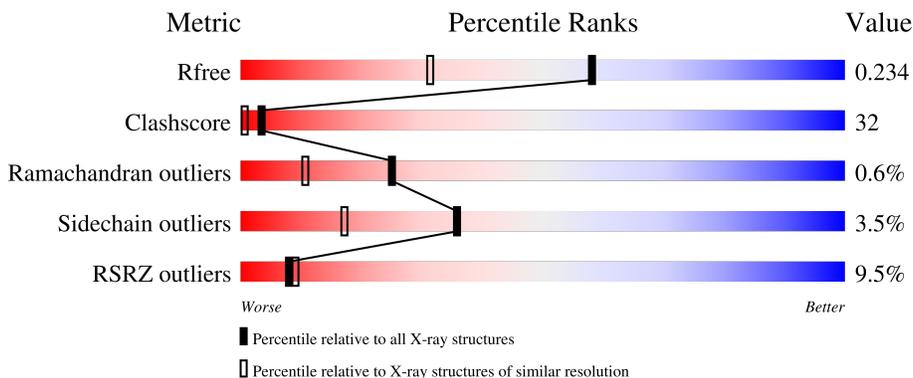|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 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	Percentile Ranks	Value
Rfree		0.234
Clashscore		32
Ramachandran outliers		0.6%
Sidechain outliers		3.5%
RSRZ outliers		9.5%

Worse *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (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% 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	268	7% 22% 56% 16% · ·
2	B	396	11% 7% 65% 27% ··

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 5352 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 TRYPTOPHAN SYNTHASE ALPHA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total 1955	C 1248	N 334	O 365	S 8	0	2	0

- Molecule 2 is a protein called TRYPTOPHAN SYNTHASE BETA CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	394	Total 2987	C 1877	N 527	O 564	S 19	0	1	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	34	SER	ARG	cloning artifact	UNP P0A2K1
B	178	PRO	SER	engineered mutation	UNP P0A2K1

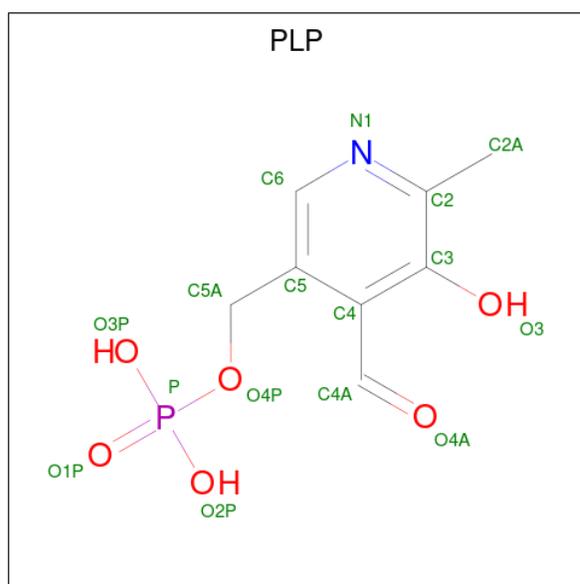- Molecule 3 is N-[1H-INDOL-3-YL-ACETYL]GLYCINE ACID (three-letter code: IAG) (formula: $C_{12}H_{12}N_2O_3$).
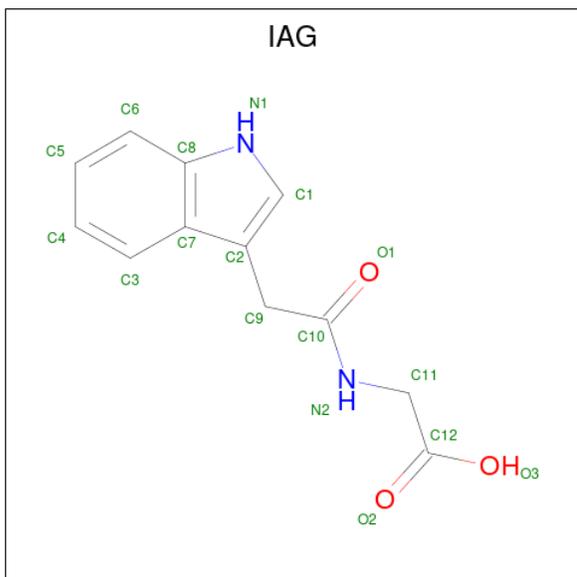
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total 17	C 12	N 2	O 3	0	0

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 1	Na 1	0	0

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

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	129	Total 129	O 129	0	0
6	B	248	Total 248	O 248	0	0

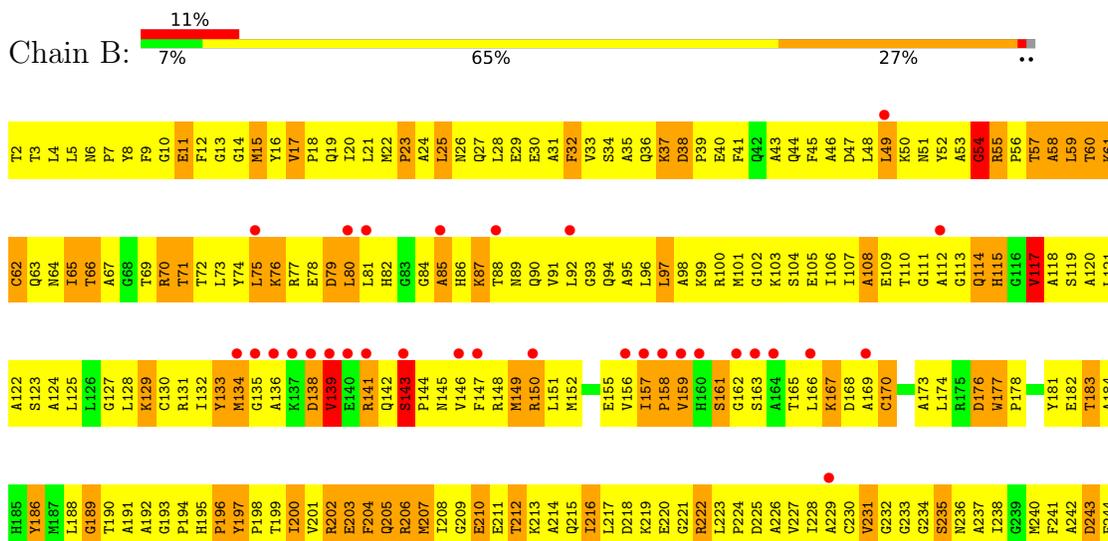# 3 Residue-property plots ⓘ

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: TRYPTOPHAN SYNTHASE ALPHA CHAIN
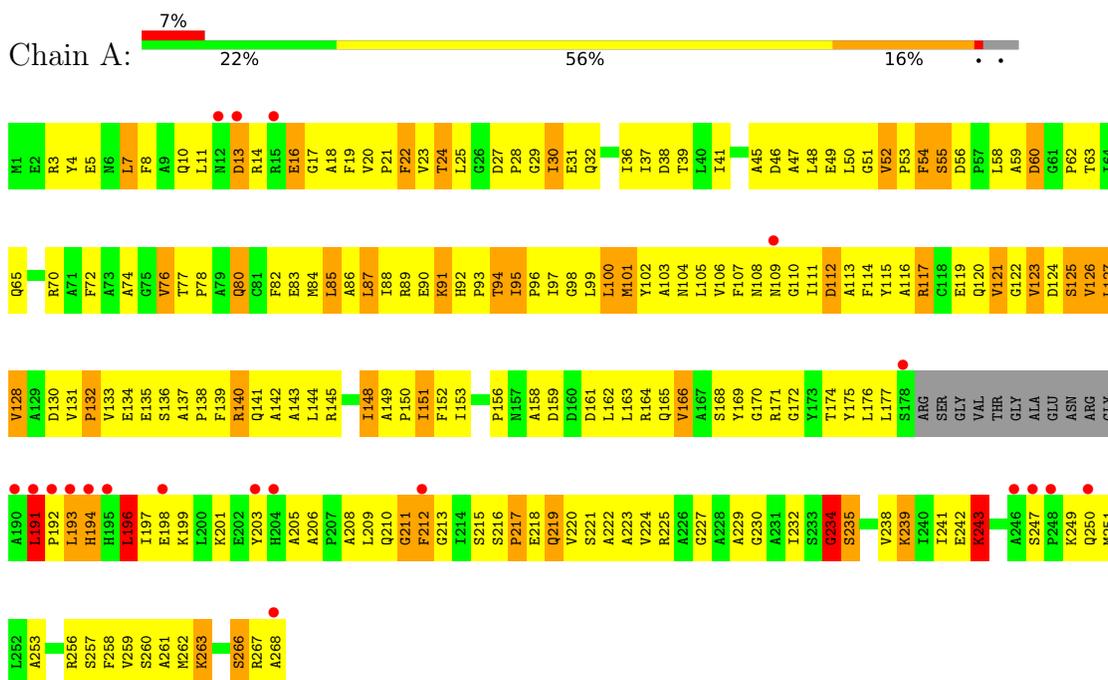
Chain A:

• Molecule 2: TRYPTOPHAN SYNTHASE BETA CHAIN

Chain B:

I245 N246 D247 T248 S249 V250 G251 L252 I253 G254 V255 E256 P257 G258 G259 H260 G261 I262 E263 T264 G265 E266 H267 G268 A269 P270 L271 K272 H273 G274 R275 V276 G277 I278 Y279 F280 G281 M282 K283 A284 P285 M286 M287 Q288 T289 A290 D291 G292 Q293 I294 E295 E296 S297 Y298 S299 I300 S301 A302 G303 L304

D305 F306 P307 S308 V309 G310 P311 Q312 H313 A314 Y315 L316 N317 S318 I319 G320 R321 A322 D323 Y324 V325 S326 I327 T328 D329 D330 E331 A332 L333 E334 A335 F336 K337 T338 L339 C340 R341 H342 E343 G344 I345 I346 P347 A348 L349 E350 S351 S352 H353 A354 L355 A356 H357 A358 L359 K360 M361 M362 R363 E364

Q365 P366 E367 K368 E369 Q370 L371 L372 V373 V374 N375 L376 S377 G378 R379 G380 D381 K382 D383 I384 F385 T386 V387 H388 D389 I390 L391 K392 A393 R394 G395 GLU ILE

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	184.42Å 61.03Å 67.53Å 90.00° 94.69° 90.00°	Depositor
Resolution (Å)	20.00 – 1.70 43.36 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.4 (20.00-1.70) 96.4 (43.36-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) >$ [1]	1.92 (at 1.70Å)	Xtriage
Refinement program	CNS, REFMAC	Depositor
R, R_{free}	0.209 , 0.265 0.188 , 0.234	Depositor DCC
R_{free} test set	4065 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	20.8	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 56.4	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5352	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% 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: IAG, PLP, NA

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	3.32	245/2004 (12.2%)	2.86	199/2722 (7.3%)
2	B	4.35	703/3052 (23.0%)	4.01	646/4123 (15.7%)
All	All	3.97	948/5056 (18.8%)	3.60	845/6845 (12.3%)

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
2	B	0	1
All	All	0	2

The worst 5 of 948 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	315	TYR	CE2-CZ	24.33	1.70	1.38
2	B	30	GLU	CD-OE2	20.54	1.48	1.25
2	B	8	TYR	CE1-CZ	19.55	1.64	1.38
2	B	291	ASP	CB-CG	-17.49	1.15	1.51
2	B	16	TYR	CE2-CZ	16.98	1.60	1.38

The worst 5 of 845 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(º)	Ideal(º)
2	B	55	ARG	NE-CZ-NH1	35.94	138.27	120.30
2	B	291	ASP	CB-CG-OD1	-34.65	87.12	118.30
2	B	55	ARG	NE-CZ-NH2	-29.76	105.42	120.30
1	A	140	ARG	NE-CZ-NH1	-28.07	106.26	120.30

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	222	ARG	NE-CZ-NH2	-26.74	106.93	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	267	ARG	Peptide
2	B	54	GLY	Mainchain

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	1955	0	1959	85	0
2	B	2987	0	2961	236	0
3	A	17	0	11	1	0
4	B	1	0	0	0	0
5	B	15	0	7	1	0
6	A	129	0	0	3	0
6	B	248	0	0	2	0
All	All	5352	0	4938	319	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 32.

The worst 5 of 319 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)
2:B:235:SER:CA	2:B:235:SER:CB	1.74	1.64
2:B:278:ILE:CG1	2:B:278:ILE:CD1	1.74	1.64
2:B:129:LYS:CD	2:B:129:LYS:CG	1.75	1.63
1:A:52:VAL:CB	1:A:52:VAL:CG1	1.75	1.63
1:A:126:VAL:CB	1:A:126:VAL:CG1	1.75	1.63

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	255/268 (95%)	250 (98%)	3 (1%)	2 (1%)	19	6
2	B	393/396 (99%)	371 (94%)	20 (5%)	2 (0%)	29	13
All	All	648/664 (98%)	621 (96%)	23 (4%)	4 (1%)	25	11

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	191	LEU
1	A	234	GLY
2	B	117	VAL
2	B	139	VAL

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	203/208 (98%)	195 (96%)	8 (4%)	32	13
2	B	309/310 (100%)	299 (97%)	10 (3%)	39	20
All	All	512/518 (99%)	494 (96%)	18 (4%)	36	17

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	216	ILE
2	B	394	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	347	PRO
2	B	65	ILE
2	B	207	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	109	ASN
1	A	219	GLN
2	B	246	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 1 is 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	IAG	A	501	-	17,18,18	1.87	4 (23%)	18,24,24	2.92	10 (55%)
5	PLP	B	502	2	15,15,16	3.62	9 (60%)	20,22,23	2.94	5 (25%)

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	IAG	A	501	-	-	2/9/9/9	0/2/2/2
5	PLP	B	502	2	-	0/6/6/8	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	502	PLP	P-O4P	-6.76	1.38	1.60
5	B	502	PLP	C3-C2	-5.71	1.35	1.40
5	B	502	PLP	C2A-C2	5.64	1.59	1.50
5	B	502	PLP	P-O3P	-4.32	1.38	1.54
5	B	502	PLP	C5-C4	4.07	1.45	1.40

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	502	PLP	O4P-C5A-C5	8.30	125.17	109.35
5	B	502	PLP	O3P-P-O4P	6.90	125.09	106.73
3	A	501	IAG	C6-C8-N1	5.28	145.41	130.80
3	A	501	IAG	C11-N2-C10	-4.72	108.22	121.53
3	A	501	IAG	C7-C8-N1	-4.68	97.53	107.92

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	IAG	O1-C10-C9-C2
3	A	501	IAG	N2-C10-C9-C2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	IAG	1	0
5	B	502	PLP	1	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	6
1	A	1

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	65:ILE	C	66:THR	N	1.19
1	B	321:ARG	C	322:ALA	N	1.19
1	B	332:ALA	C	333:LEU	N	1.19
1	B	343:GLU	C	344:GLY	N	1.19
1	B	362:MET	C	363:ARG	N	1.19

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, 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(Å²)	Q<0.9
1	A	257/268 (95%)	0.51	20 (7%)	13	15	13, 26, 48, 78	0
2	B	394/396 (99%)	0.66	42 (10%)	6	7	11, 19, 42, 73	0
All	All	651/664 (98%)	0.60	62 (9%)	8	9	11, 22, 44, 78	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	268	ALA	7.2
2	B	159	VAL	6.9
1	A	190	ALA	6.4
2	B	385	PHE	6.3
1	A	191	LEU	5.8

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, 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
3	IAG	A	501	17/17	0.84	0.15	22,35,50,52	0
5	PLP	B	502	15/16	0.98	0.12	15,18,28,30	0
4	NA	B	503	1/1	0.99	0.09	17,17,17,17	0

6.5 Other polymers (i)

There are no such residues in this entry.