wwPDB X-ray Structure Validation Summary Report (i)

Mar 24, 2025 – 03:03 PM JST

PDB ID	:	8XND
Title	:	Crystal structure of serine hydroxymethyltransferase 1
Authors	:	Fan, S.; Wei, X.; Lv, R.; Wang, C.; Tang, M.; Jin, Y.; Yang, Z.
Deposited on	:	2023-12-29
Resolution	:	3.45 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 (1)) were used in the production of this report:

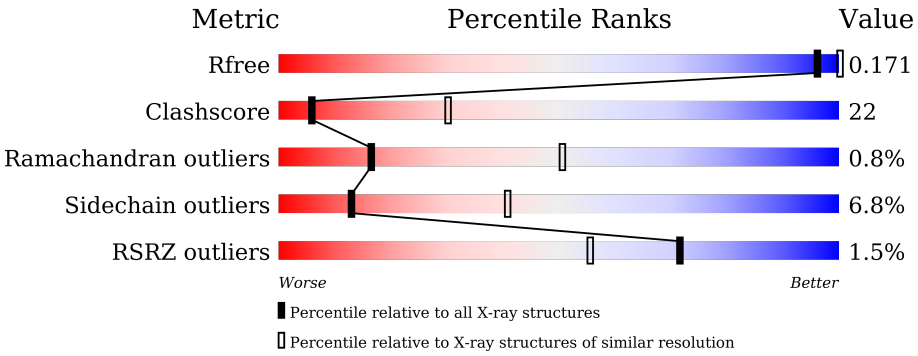
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.2

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 3.45 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	1597 (3.52-3.40)
Clashscore	180529	1041 (3.50-3.42)
Ramachandran outliers	177936	1026 (3.50-3.42)
Sidechain outliers	177891	1027 (3.50-3.42)
RSRZ outliers	164620	1596 (3.52-3.40)

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 <=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	472	61% 30% 7%
1	B	472	57% 35% 7%
1	C	472	59% 33% 5%
1	D	472	57% 35% 7%

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
2	PLP	C	500	-	-	X	-
2	PLP	D	500	-	-	X	-

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 14588 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.

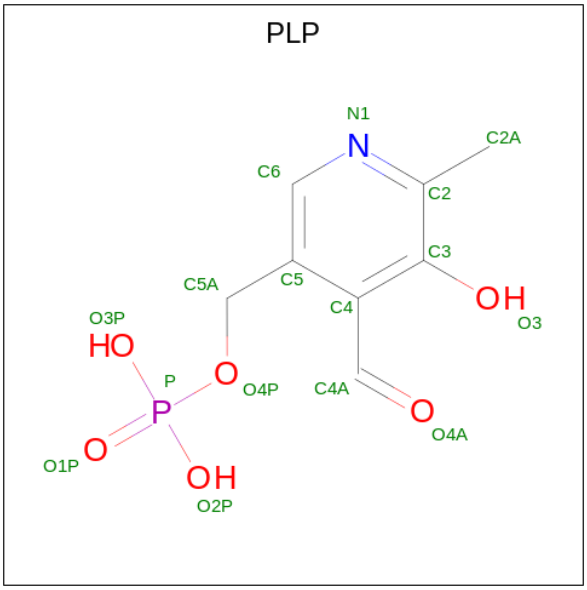
- Molecule 1 is a protein called Serine hydroxymethyltransferase, cytosolic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	472	Total	C	N	O	S	0	0	0
			3632	2290	637	687	18			
1	B	472	Total	C	N	O	S	0	0	0
			3632	2290	637	687	18			
1	C	472	Total	C	N	O	S	0	0	0
			3632	2290	637	687	18			
1	D	472	Total	C	N	O	S	0	0	0
			3632	2290	637	687	18			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	GLN	-	expression tag	UNP P34896
A	10	GLY	-	expression tag	UNP P34896
A	20	ALA	LYS	conflict	UNP P34896
A	279	ALA	LYS	conflict	UNP P34896
A	442	ALA	ARG	conflict	UNP P34896
B	9	GLN	-	expression tag	UNP P34896
B	10	GLY	-	expression tag	UNP P34896
B	20	ALA	LYS	conflict	UNP P34896
B	279	ALA	LYS	conflict	UNP P34896
B	442	ALA	ARG	conflict	UNP P34896
C	9	GLN	-	expression tag	UNP P34896
C	10	GLY	-	expression tag	UNP P34896
C	20	ALA	LYS	conflict	UNP P34896
C	279	ALA	LYS	conflict	UNP P34896
C	442	ALA	ARG	conflict	UNP P34896
D	9	GLN	-	expression tag	UNP P34896
D	10	GLY	-	expression tag	UNP P34896
D	20	ALA	LYS	conflict	UNP P34896
D	279	ALA	LYS	conflict	UNP P34896
D	442	ALA	ARG	conflict	UNP P34896

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



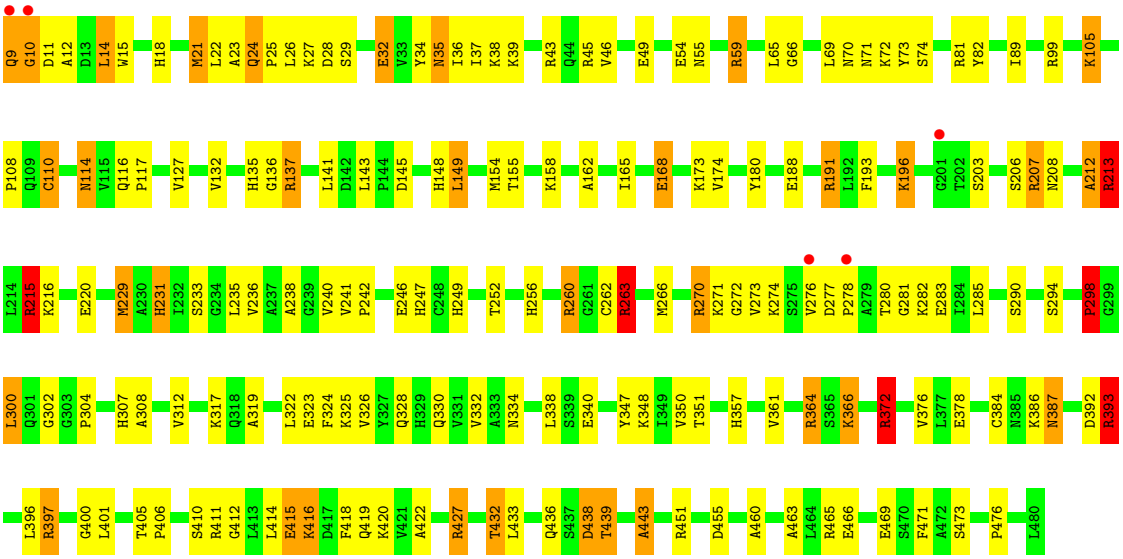
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	C	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	D	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

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.

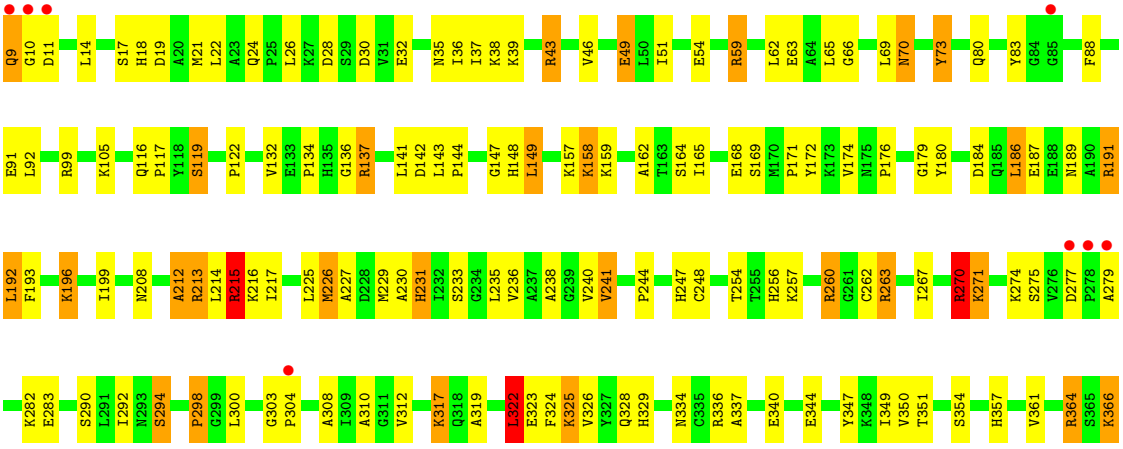
• Molecule 1: Serine hydroxymethyltransferase, cytosolic

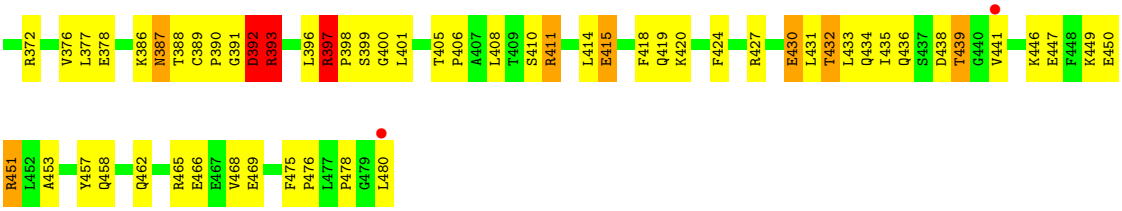
Chain A:



• Molecule 1: Serine hydroxymethyltransferase, cytosolic

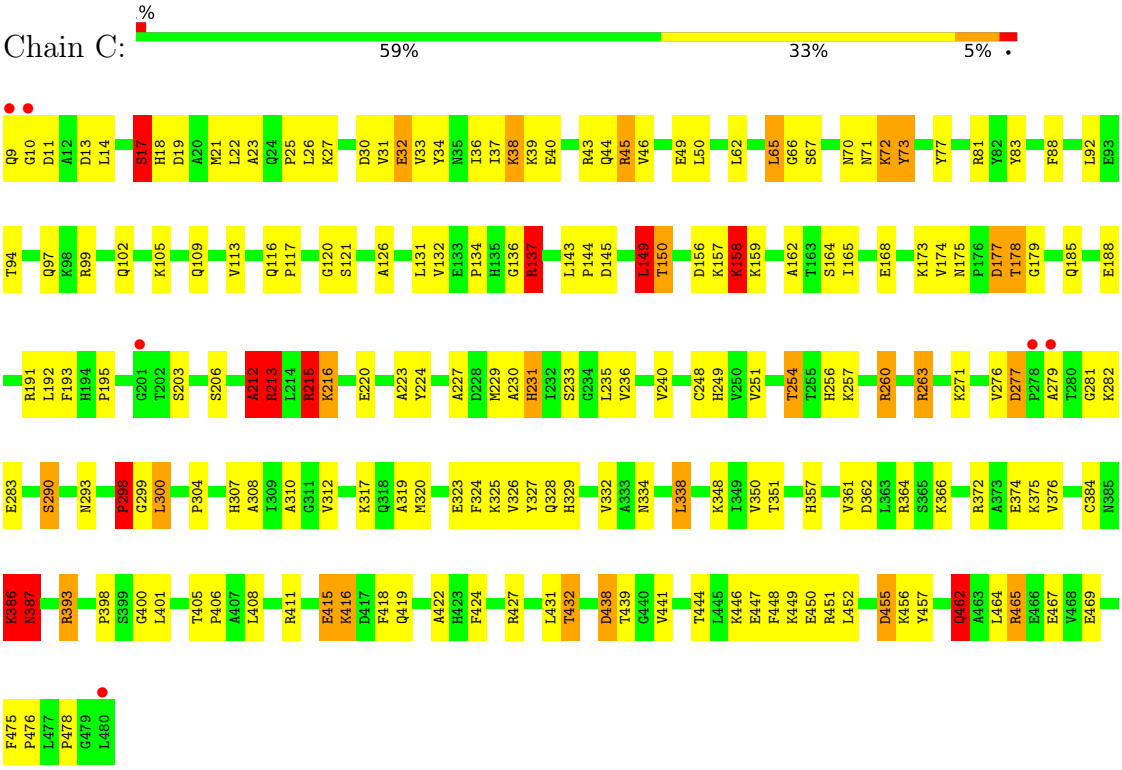
Chain B:





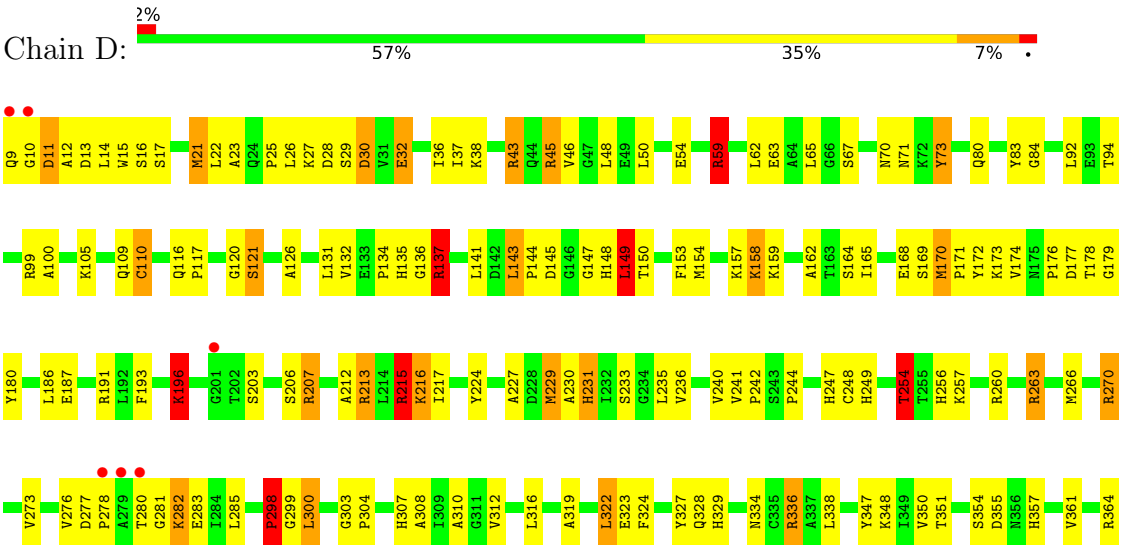
• Molecule 1: Serine hydroxymethyltransferase, cytosolic

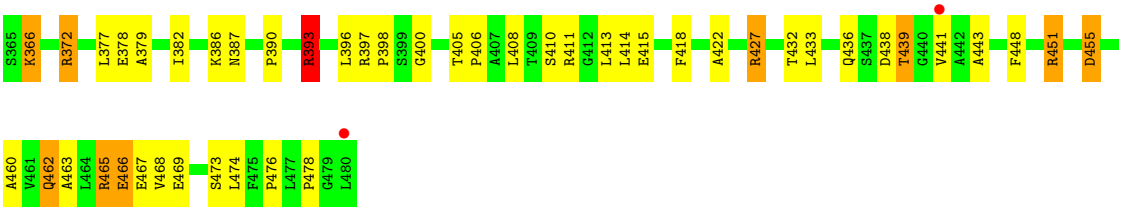
Chain C:



• Molecule 1: Serine hydroxymethyltransferase, cytosolic

Chain D:




S365
K366
R372
L377
E378
A379
I382
K386
N387
P390
R393
L396
R397
P398
S399
G400
T405
P406
A407
L408
T409
S410
R411
G412
L413
L414
E415
F418
A422
R427
T432
L433
Q436
S437
D438
T439
G440
V441
A442
A443
F448
R451
D455
A460
V461
Q462
A463
L464
R465
E466
E467
V468
E469
S473
L474
F475
P476
L477
P478
G479
L480

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	151.54Å 151.54Å 234.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.53 – 3.45 48.53 – 3.45	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.53-3.45) 99.9 (48.53-3.45)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.74 (at 3.48Å)	Xtriage
Refinement program	REFMAC 5.8.0419	Depositor
R, R_{free}	0.153 , 0.192 0.143 , 0.171	Depositor DCC
R_{free} test set	4079 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	95.2	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 63.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.268 for -h,-k,l 0.269 for h,-h-k,-l 0.268 for -k,-h,-l	Xtriage
Reported twinning fraction	0.228 for H, K, L 0.231 for K, H, -L 0.240 for -h,-k,l 0.301 for -K, -H, -L	Depositor
Outliers	0 of 79345 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	14588	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.36% 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: 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	0.63	2/3705 (0.1%)	1.42	58/5012 (1.2%)
1	B	0.63	2/3705 (0.1%)	1.38	53/5012 (1.1%)
1	C	0.61	0/3705	1.38	47/5012 (0.9%)
1	D	0.65	2/3705 (0.1%)	1.42	58/5012 (1.2%)
All	All	0.63	6/14820 (0.0%)	1.40	216/20048 (1.1%)

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	7
1	B	0	6
1	C	0	5
1	D	0	6
All	All	0	24

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	168	GLU	CD-OE2	6.89	1.33	1.25
1	D	378	GLU	CD-OE1	6.44	1.32	1.25
1	D	32	GLU	CD-OE2	5.58	1.31	1.25
1	A	168	GLU	CD-OE1	5.29	1.31	1.25
1	B	415	GLU	CD-OE1	5.22	1.31	1.25

The worst 5 of 216 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	59	ARG	CB-CA-C	-15.76	78.88	110.40
1	D	59	ARG	CB-CA-C	-14.04	82.32	110.40
1	D	215	ARG	NE-CZ-NH1	-13.45	113.57	120.30
1	C	188	GLU	N-CA-CB	12.48	133.06	110.60
1	D	143	LEU	CB-CG-CD2	-12.00	90.60	111.00

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	215	ARG	Sidechain
1	A	263	ARG	Sidechain
1	A	372	ARG	Sidechain
1	A	393	ARG	Sidechain
1	A	397	ARG	Sidechain

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	3632	0	3610	175	0
1	B	3632	0	3612	210	0
1	C	3632	0	3612	206	0
1	D	3632	0	3612	202	0
2	A	15	0	7	2	0
2	B	15	0	7	1	0
2	C	15	0	6	7	0
2	D	15	0	7	11	0
All	All	14588	0	14473	654	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 22.

The worst 5 of 654 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:257:LYS:NZ	2:D:500:PLP:C4A	1.80	1.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:257:LYS:HE3	2:C:500:PLP:C4A	1.61	1.31
1:D:439:THR:HG23	1:D:443:ALA:CB	1.66	1.23
1:D:257:LYS:HZ1	2:D:500:PLP:C4A	1.45	1.19
1:D:439:THR:CG2	1:D:443:ALA:HB3	1.73	1.18

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	470/472 (100%)	439 (93%)	27 (6%)	4 (1%)	14	48
1	B	470/472 (100%)	436 (93%)	29 (6%)	5 (1%)	12	44
1	C	470/472 (100%)	443 (94%)	22 (5%)	5 (1%)	12	44
1	D	470/472 (100%)	443 (94%)	26 (6%)	1 (0%)	44	76
All	All	1880/1888 (100%)	1761 (94%)	104 (6%)	15 (1%)	16	50

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	392	ASP
1	B	441	VAL
1	C	441	VAL
1	D	11	ASP
1	A	10	GLY

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	385/385 (100%)	361 (94%)	24 (6%)	15	44
1	B	385/385 (100%)	356 (92%)	29 (8%)	11	37
1	C	385/385 (100%)	358 (93%)	27 (7%)	12	40
1	D	385/385 (100%)	360 (94%)	25 (6%)	14	43
All	All	1540/1540 (100%)	1435 (93%)	105 (7%)	13	41

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	49	GLU
1	C	290	SER
1	D	427	ARG
1	C	65	LEU
1	C	177	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	231	HIS
1	C	462	GLN
1	C	458	GLN
1	D	70	ASN
1	B	24	GLN

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)

4 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	B	500	-	15,15,16	0.98	0	20,22,23	1.02	1 (5%)
2	PLP	C	500	-	15,15,16	1.10	2 (13%)	20,22,23	1.58	3 (15%)
2	PLP	A	500	-	15,15,16	0.97	1 (6%)	20,22,23	1.62	3 (15%)
2	PLP	D	500	-	15,15,16	1.06	1 (6%)	20,22,23	1.79	6 (30%)

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	B	500	-	-	5/6/6/8	0/1/1/1
2	PLP	C	500	-	-	1/6/6/8	0/1/1/1
2	PLP	A	500	-	-	1/6/6/8	0/1/1/1
2	PLP	D	500	-	-	2/6/6/8	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	500	PLP	P-O1P	2.56	1.58	1.50
2	D	500	PLP	C4A-C4	-2.32	1.46	1.51
2	A	500	PLP	C4A-C4	-2.08	1.47	1.51
2	C	500	PLP	C4A-C4	-2.05	1.47	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	500	PLP	O4P-C5A-C5	4.88	118.65	109.35
2	C	500	PLP	O3-C3-C2	4.09	126.42	117.49
2	C	500	PLP	O4P-C5A-C5	3.89	116.76	109.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	500	PLP	O2P-P-O4P	3.52	116.10	106.73
2	A	500	PLP	O3-C3-C2	3.51	125.14	117.49

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	500	PLP	C4-C5-C5A-O4P
2	B	500	PLP	C6-C5-C5A-O4P
2	B	500	PLP	C5A-O4P-P-O2P
2	B	500	PLP	C5A-O4P-P-O3P
2	D	500	PLP	C4-C5-C5A-O4P

There are no ring outliers.

4 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	500	PLP	1	0
2	C	500	PLP	7	0
2	A	500	PLP	2	0
2	D	500	PLP	11	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

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	472/472 (100%)	-0.57	5 (1%)	77	62	42, 90, 123, 181	0
1	B	472/472 (100%)	-0.52	10 (2%)	63	47	50, 90, 125, 174	0
1	C	472/472 (100%)	-0.61	6 (1%)	74	59	48, 89, 125, 195	0
1	D	472/472 (100%)	-0.58	8 (1%)	69	52	49, 88, 123, 179	0
All	All	1888/1888 (100%)	-0.57	29 (1%)	71	55	42, 89, 124, 195	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	9	GLN	8.9
1	D	10	GLY	8.2
1	C	9	GLN	5.7
1	A	9	GLN	5.2
1	B	10	GLY	5.0

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	PLP	A	500	15/16	0.99	0.06	69,83,118,120	0
2	PLP	B	500	15/16	0.99	0.05	74,84,103,123	0
2	PLP	C	500	15/16	0.99	0.04	49,65,95,97	0
2	PLP	D	500	15/16	0.99	0.04	62,70,93,116	0

6.5 Other polymers (i)

There are no such residues in this entry.