



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:16 PM GMT

PDB ID : 4L28
Title : Crystal structure of delta516-525 human cystathionine beta-synthase D444N mutant containing C-terminal 6xHis tag
Authors : Ereno, J.; Majtan, T.; Oyenarte, I.; Kraus, J.P.; Martinez, L.A.
Deposited on : 2013-06-04
Resolution : 2.63 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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

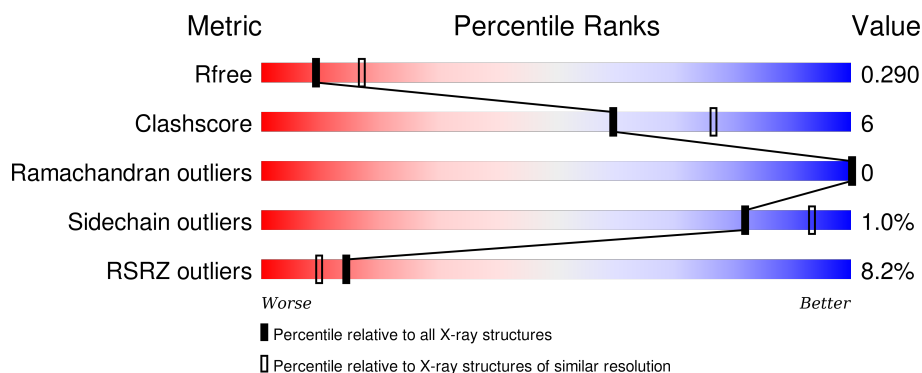
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.63 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}	91344	2700 (2.64-2.60)
Clashscore	102246	3065 (2.64-2.60)
Ramachandran outliers	100387	3015 (2.64-2.60)
Sidechain outliers	100360	3015 (2.64-2.60)
RSRZ outliers	91569	2706 (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 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	558	<div> <div>5%</div> <div>76%12%12%</div> </div>
1	B	558	<div> <div>8%</div> <div>78%11%11%</div> </div>
1	C	558	<div> <div>8%</div> <div>76%11%12%</div> </div>
1	D	558	<div> <div>7%</div> <div>77%11%12%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 15455 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 Cystathionine beta-synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	493	Total	C	N	O	S	0	0	0
			3803	2410	665	706	22			
1	B	494	Total	C	N	O	S	0	0	0
			3813	2414	669	708	22			
1	C	491	Total	C	N	O	S	0	0	0
			3791	2403	666	700	22			
1	D	493	Total	C	N	O	S	0	0	0
			3816	2416	668	710	22			

There are 36 discrepancies between the modelled and reference sequences:

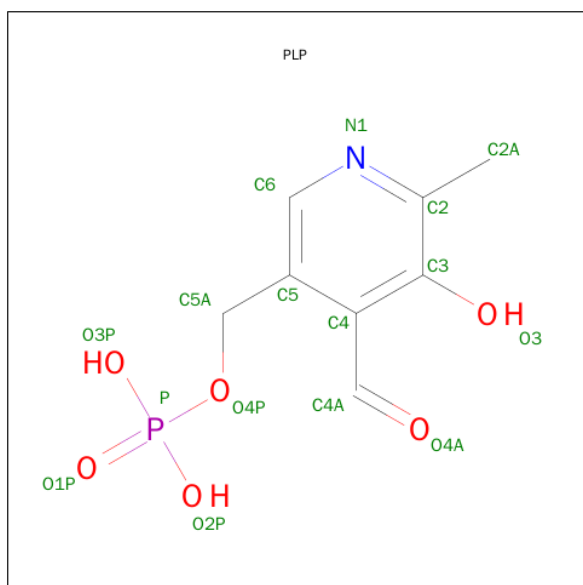
Chain	Residue	Modelled	Actual	Comment	Reference
A	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
A	552	LEU	-	EXPRESSION TAG	UNP P35520
A	553	GLU	-	EXPRESSION TAG	UNP P35520
A	554	HIS	-	EXPRESSION TAG	UNP P35520
A	555	HIS	-	EXPRESSION TAG	UNP P35520
A	556	HIS	-	EXPRESSION TAG	UNP P35520
A	557	HIS	-	EXPRESSION TAG	UNP P35520
A	558	HIS	-	EXPRESSION TAG	UNP P35520
A	559	HIS	-	EXPRESSION TAG	UNP P35520
B	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
B	552	LEU	-	EXPRESSION TAG	UNP P35520
B	553	GLU	-	EXPRESSION TAG	UNP P35520
B	554	HIS	-	EXPRESSION TAG	UNP P35520
B	555	HIS	-	EXPRESSION TAG	UNP P35520
B	556	HIS	-	EXPRESSION TAG	UNP P35520
B	557	HIS	-	EXPRESSION TAG	UNP P35520
B	558	HIS	-	EXPRESSION TAG	UNP P35520
B	559	HIS	-	EXPRESSION TAG	UNP P35520
C	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
C	552	LEU	-	EXPRESSION TAG	UNP P35520
C	553	GLU	-	EXPRESSION TAG	UNP P35520

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	554	HIS	-	EXPRESSION TAG	UNP P35520
C	555	HIS	-	EXPRESSION TAG	UNP P35520
C	556	HIS	-	EXPRESSION TAG	UNP P35520
C	557	HIS	-	EXPRESSION TAG	UNP P35520
C	558	HIS	-	EXPRESSION TAG	UNP P35520
C	559	HIS	-	EXPRESSION TAG	UNP P35520
D	444	ASN	ASP	ENGINEERED MUTATION	UNP P35520
D	552	LEU	-	EXPRESSION TAG	UNP P35520
D	553	GLU	-	EXPRESSION TAG	UNP P35520
D	554	HIS	-	EXPRESSION TAG	UNP P35520
D	555	HIS	-	EXPRESSION TAG	UNP P35520
D	556	HIS	-	EXPRESSION TAG	UNP P35520
D	557	HIS	-	EXPRESSION TAG	UNP P35520
D	558	HIS	-	EXPRESSION TAG	UNP P35520
D	559	HIS	-	EXPRESSION TAG	UNP P35520

- 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		

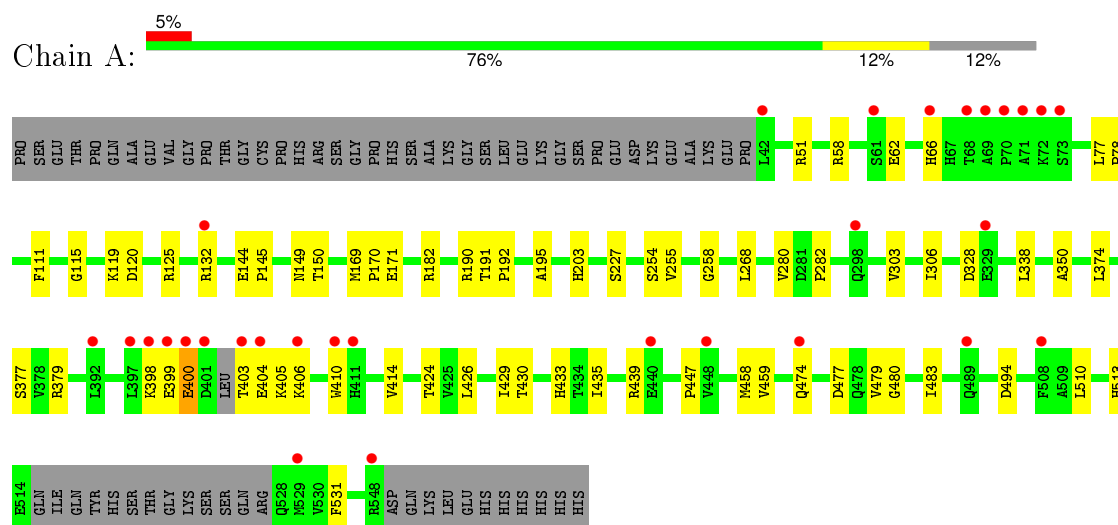
-
- Chemical structure of HEM (Heme) showing a central iron atom coordinated by four nitrogen atoms in a porphyrin-like ring, with various side chains and a central heme group.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	B	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	C	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
3	D	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

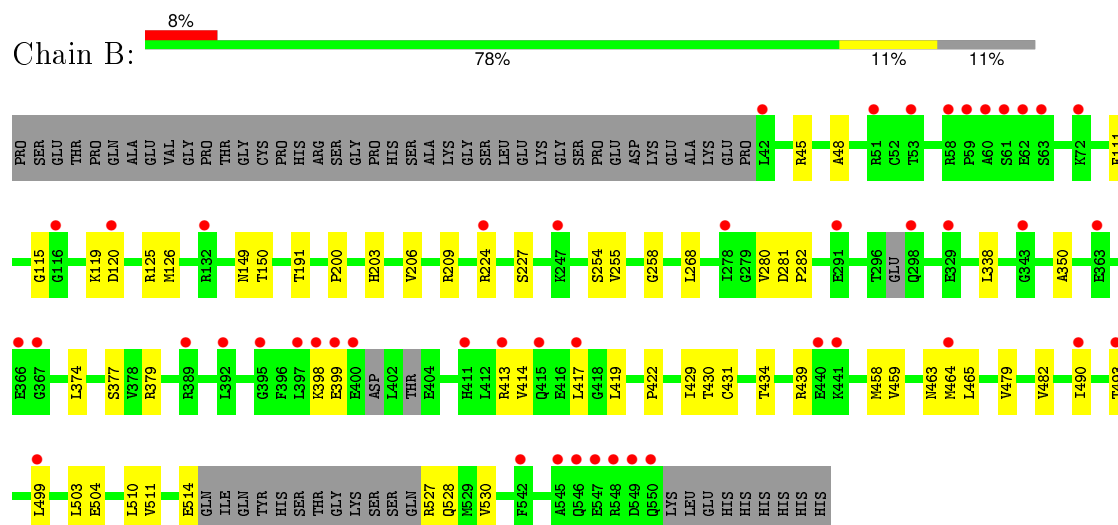
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

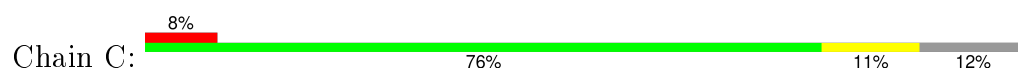
• Molecule 1: Cystathionine beta-synthase

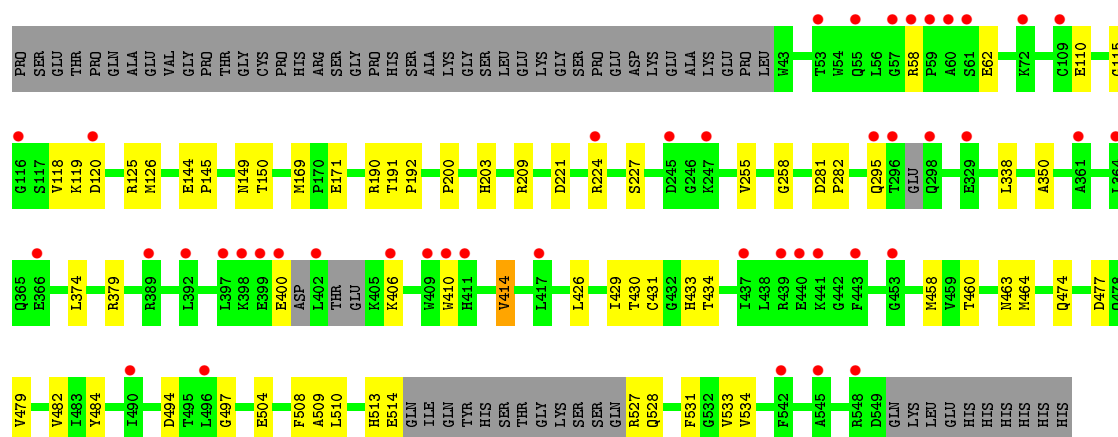


• Molecule 1: Cystathionine beta-synthase

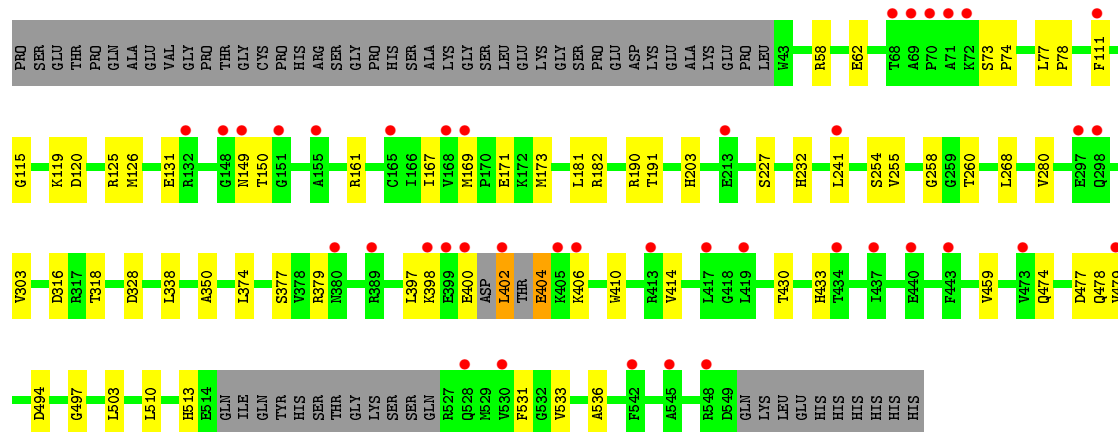
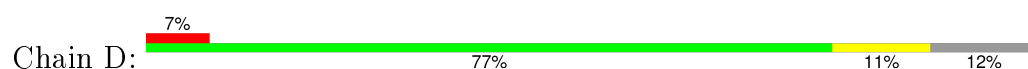


• Molecule 1: Cystathionine beta-synthase





• Molecule 1: Cystathionine beta-synthase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	109.48Å 131.07Å 207.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.39 – 2.63 49.41 – 2.63	Depositor EDS
% Data completeness (in resolution range)	98.1 (48.39-2.63) 98.3 (49.41-2.63)	Depositor EDS
R_{merge}	0.36	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.07 (at 2.61Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.256 , 0.281 0.267 , 0.290	Depositor DCC
R_{free} test set	4472 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	58.8	Xtriage
Anisotropy	0.415	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	5 of 88003 reflections (0.006%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	15455	wwPDB-VP
Average B, all atoms (Å ²)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.62 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.2651e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: HEM, 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	0.60	0/3874	0.52	0/5245
1	B	0.60	0/3882	0.52	0/5253
1	C	0.60	0/3860	0.52	0/5222
1	D	0.62	0/3886	0.53	0/5258
All	All	0.61	0/15502	0.52	0/20978

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	3803	0	3855	41	0
1	B	3813	0	3859	42	0
1	C	3791	0	3852	49	0
1	D	3816	0	3871	43	0
2	A	15	0	6	2	0
2	B	15	0	6	2	0
2	C	15	0	6	2	0
2	D	15	0	6	2	0
3	A	43	0	30	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	43	0	30	1	0
3	C	43	0	30	2	0
3	D	43	0	30	2	0
All	All	15455	0	15581	176	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 6.

The worst 5 of 176 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:48:ALA:HB3	1:B:224:ARG:HH12	1.19	1.07
1:A:403:THR:HA	1:A:406:LYS:HD2	1.43	0.99
1:D:513:HIS:HB2	1:D:531:PHE:HE2	1.45	0.80
1:B:431:CYS:HB3	1:B:464:MET:HE1	1.63	0.79
1:B:514:GLU:O	1:B:527:ARG:NH2	2.17	0.78

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	487/558 (87%)	480 (99%)	7 (1%)	0	100	100
1	B	485/558 (87%)	476 (98%)	9 (2%)	0	100	100
1	C	482/558 (86%)	475 (98%)	7 (2%)	0	100	100
1	D	486/558 (87%)	480 (99%)	6 (1%)	0	100	100
All	All	1940/2232 (87%)	1911 (98%)	29 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	412/472 (87%)	407 (99%)	5 (1%)	78	92
1	B	413/472 (88%)	410 (99%)	3 (1%)	88	96
1	C	411/472 (87%)	409 (100%)	2 (0%)	92	98
1	D	415/472 (88%)	409 (99%)	6 (1%)	74	90
All	All	1651/1888 (87%)	1635 (99%)	16 (1%)	82	93

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	399	GLU
1	C	338	LEU
1	D	398	LYS
1	B	338	LEU
1	D	400	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	66	HIS
1	D	528	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

8 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	A	601	1	15,15,16	3.08	8 (53%)	21,22,23	2.16	7 (33%)
3	HEM	A	602	1	30,50,50	2.34	5 (16%)	24,82,82	2.33	8 (33%)
2	PLP	B	601	1	15,15,16	3.22	8 (53%)	21,22,23	2.32	8 (38%)
3	HEM	B	602	1	30,50,50	2.13	5 (16%)	24,82,82	2.25	7 (29%)
2	PLP	C	601	1	15,15,16	3.18	8 (53%)	21,22,23	2.25	7 (33%)
3	HEM	C	602	1	30,50,50	2.29	7 (23%)	24,82,82	2.17	5 (20%)
2	PLP	D	601	1	15,15,16	3.24	7 (46%)	21,22,23	2.14	6 (28%)
3	HEM	D	602	1	30,50,50	2.27	6 (20%)	24,82,82	2.35	7 (29%)

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	A	601	1	-	0/6/6/8	0/1/1/1
3	HEM	A	602	1	-	0/10/54/54	0/0/8/8
2	PLP	B	601	1	-	0/6/6/8	0/1/1/1
3	HEM	B	602	1	-	0/10/54/54	0/0/8/8
2	PLP	C	601	1	-	0/6/6/8	0/1/1/1
3	HEM	C	602	1	-	0/10/54/54	0/0/8/8
2	PLP	D	601	1	-	0/6/6/8	0/1/1/1
3	HEM	D	602	1	-	0/10/54/54	0/0/8/8

The worst 5 of 54 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	602	HEM	C3B-C4B	-8.25	1.44	1.51
3	A	602	HEM	C3B-C4B	-8.13	1.44	1.51
3	D	602	HEM	C3B-C4B	-7.91	1.44	1.51
2	B	601	PLP	C4A-C4	-7.43	1.36	1.51
3	B	602	HEM	C3B-C4B	-7.33	1.45	1.51

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	601	PLP	O4P-P-O1P	-4.00	96.96	107.14
2	A	601	PLP	O3P-P-O4P	-3.97	95.13	106.56
2	A	601	PLP	O4P-P-O1P	-3.82	97.41	107.14
2	C	601	PLP	O4P-P-O1P	-3.72	97.68	107.14
2	B	601	PLP	O4P-P-O1P	-3.63	97.91	107.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	PLP	2	0
3	A	602	HEM	2	0
2	B	601	PLP	2	0
3	B	602	HEM	1	0
2	C	601	PLP	2	0
3	C	602	HEM	2	0
2	D	601	PLP	2	0
3	D	602	HEM	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	493/558 (88%)	0.53	30 (6%) 25 18	58, 82, 128, 160	0
1	B	494/558 (88%)	0.58	47 (9%) 10 6	60, 85, 130, 158	0
1	C	491/558 (87%)	0.58	44 (8%) 12 8	58, 83, 127, 162	0
1	D	493/558 (88%)	0.59	40 (8%) 15 10	55, 80, 127, 159	0
All	All	1971/2232 (88%)	0.57	161 (8%) 14 10	55, 83, 128, 162	0

The worst 5 of 161 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	400	GLU	7.2
1	C	400	GLU	6.8
1	A	71	ALA	6.1
1	B	550	GLN	5.7
1	C	402	LEU	5.5

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(\AA^2)	Q<0.9
2	PLP	B	601	15/16	0.94	0.23	1.03	61,74,84,85	0
3	HEM	A	602	43/43	0.95	0.21	0.63	71,72,72,72	0
2	PLP	A	601	15/16	0.95	0.23	0.37	53,69,79,80	0
2	PLP	C	601	15/16	0.93	0.20	0.30	60,72,75,78	0
3	HEM	C	602	43/43	0.89	0.25	0.05	90,91,91,92	0
3	HEM	D	602	43/43	0.95	0.18	-0.03	74,74,75,75	0
3	HEM	B	602	43/43	0.92	0.23	-0.14	102,102,103,103	0
2	PLP	D	601	15/16	0.96	0.20	-0.34	63,69,78,79	0

6.5 Other polymers [i](#)

There are no such residues in this entry.