



wwPDB X-ray Structure Validation Summary Report

Jan 14, 2024 – 05:37 am GMT

PDB ID : 6YTN
Title : Magnesium chelatase H subunit (ChlH) E660W variant from *Synechocystis* sp.PCC6803
Authors : Bisson, C.; Hunter, C.N.
Deposited on : 2020-04-24
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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the  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](#) ) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

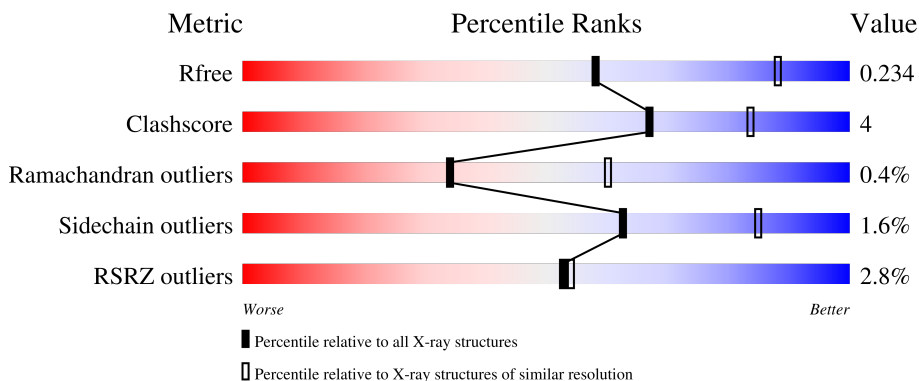
1 Overall quality at a glance

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}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (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 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 $\leq 5\%$. 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	AAA	1351	
1	BBB	1351	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 19771 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 Mg-chelatase subunit ChlH.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	1258	9913	6285	1684	1894	50	0	0	0
1	BBB	1251	9850	6240	1674	1886	50	0	0	0

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-19	MET	-	initiating methionine	UNP P73020
AAA	-18	GLY	-	expression tag	UNP P73020
AAA	-17	SER	-	expression tag	UNP P73020
AAA	-16	SER	-	expression tag	UNP P73020
AAA	-15	HIS	-	expression tag	UNP P73020
AAA	-14	HIS	-	expression tag	UNP P73020
AAA	-13	HIS	-	expression tag	UNP P73020
AAA	-12	HIS	-	expression tag	UNP P73020
AAA	-11	HIS	-	expression tag	UNP P73020
AAA	-10	HIS	-	expression tag	UNP P73020
AAA	-9	SER	-	expression tag	UNP P73020
AAA	-8	SER	-	expression tag	UNP P73020
AAA	-7	GLY	-	expression tag	UNP P73020
AAA	-6	LEU	-	expression tag	UNP P73020
AAA	-5	VAL	-	expression tag	UNP P73020
AAA	-4	PRO	-	expression tag	UNP P73020
AAA	-3	ARG	-	expression tag	UNP P73020
AAA	-2	GLY	-	expression tag	UNP P73020
AAA	-1	SER	-	expression tag	UNP P73020
AAA	0	HIS	-	expression tag	UNP P73020
AAA	660	TRP	GLU	engineered mutation	UNP P73020
BBB	-19	MET	-	initiating methionine	UNP P73020
BBB	-18	GLY	-	expression tag	UNP P73020
BBB	-17	SER	-	expression tag	UNP P73020
BBB	-16	SER	-	expression tag	UNP P73020

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	-15	HIS	-	expression tag	UNP P73020
BBB	-14	HIS	-	expression tag	UNP P73020
BBB	-13	HIS	-	expression tag	UNP P73020
BBB	-12	HIS	-	expression tag	UNP P73020
BBB	-11	HIS	-	expression tag	UNP P73020
BBB	-10	HIS	-	expression tag	UNP P73020
BBB	-9	SER	-	expression tag	UNP P73020
BBB	-8	SER	-	expression tag	UNP P73020
BBB	-7	GLY	-	expression tag	UNP P73020
BBB	-6	LEU	-	expression tag	UNP P73020
BBB	-5	VAL	-	expression tag	UNP P73020
BBB	-4	PRO	-	expression tag	UNP P73020
BBB	-3	ARG	-	expression tag	UNP P73020
BBB	-2	GLY	-	expression tag	UNP P73020
BBB	-1	SER	-	expression tag	UNP P73020
BBB	0	HIS	-	expression tag	UNP P73020
BBB	660	TRP	GLU	engineered mutation	UNP P73020

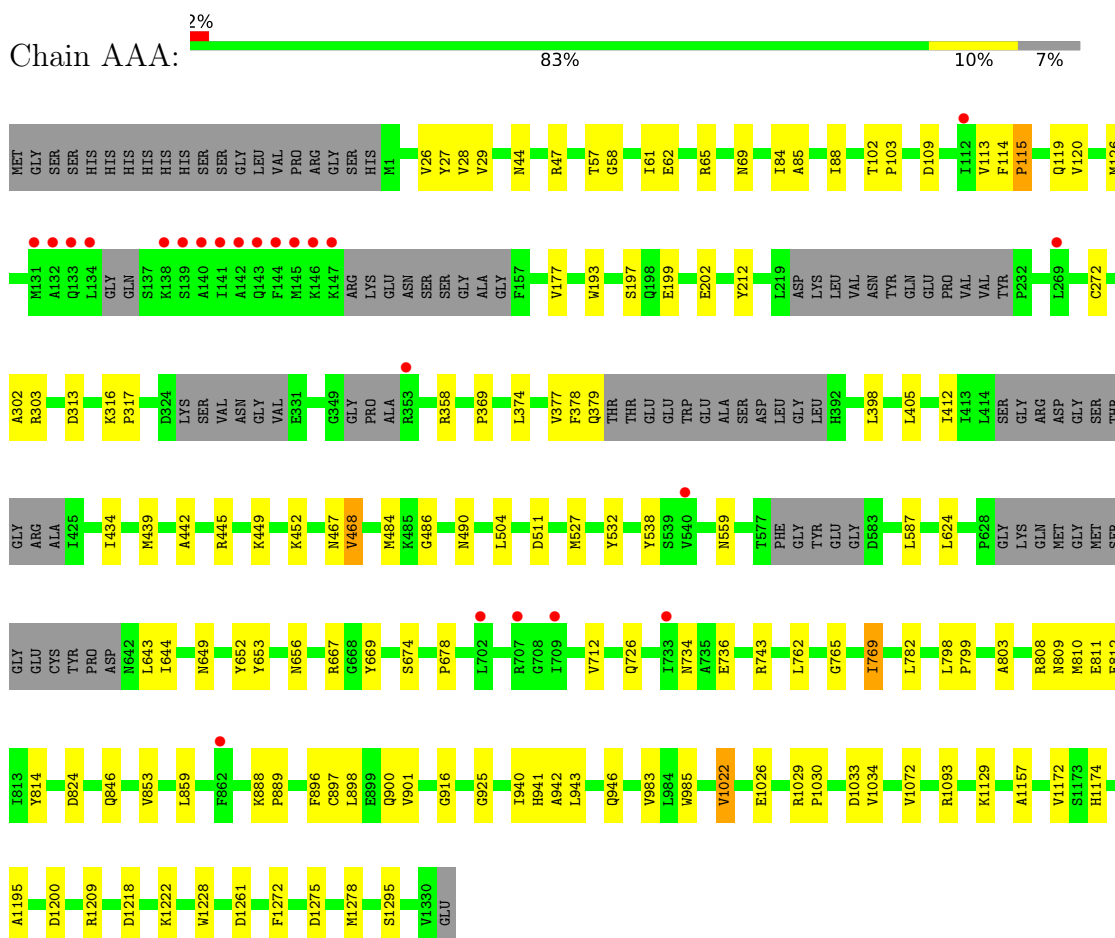
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	5	Total O 5 5	0	0
2	BBB	3	Total O 3 3	0	0

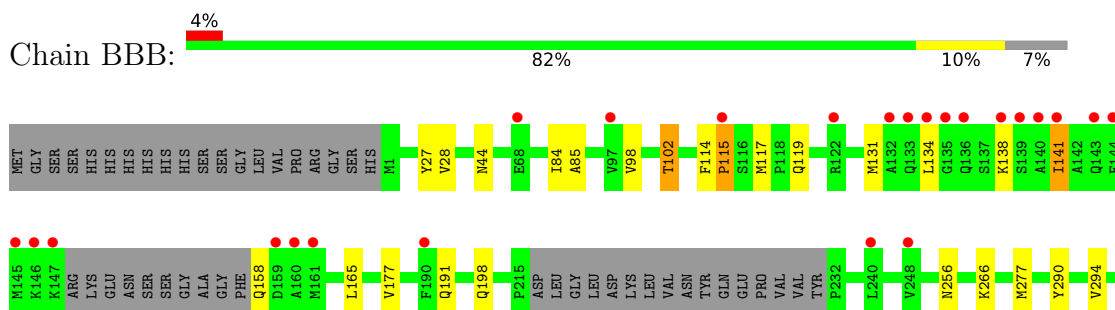
3 Residue-property plots i

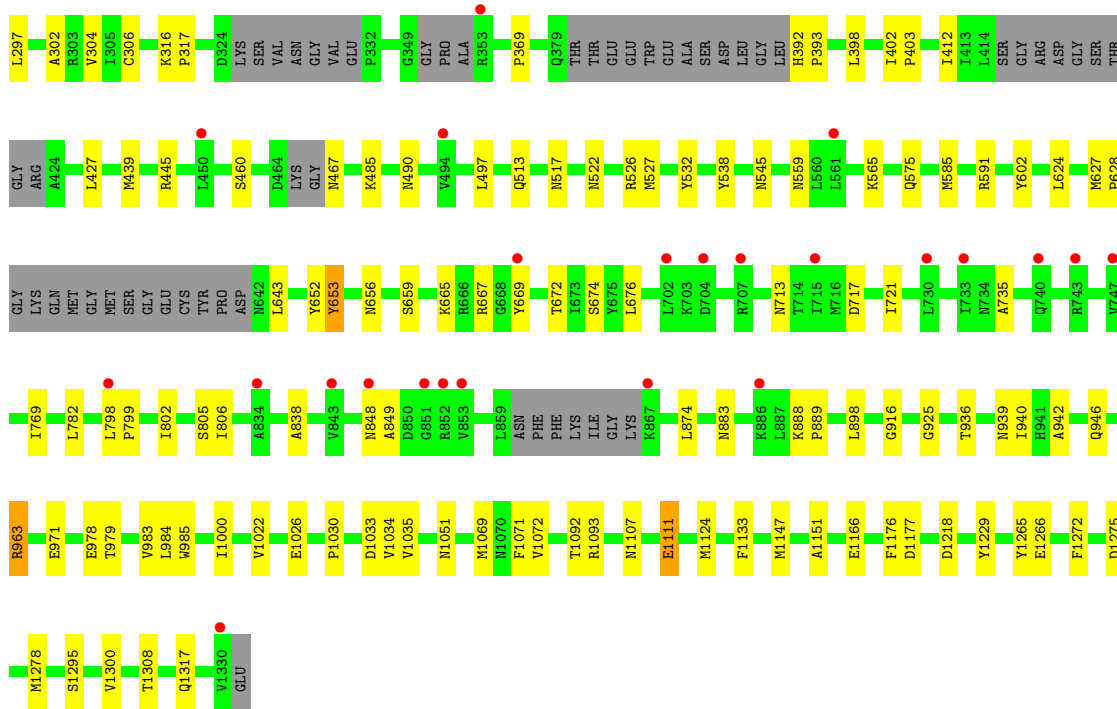
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Mg-chelatase subunit ChIH



- Molecule 1: Mg-chelatase subunit ChIH





4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	317.55Å 317.55Å 104.43Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	79.39 – 2.70 79.39 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (79.39-2.70) 99.9 (79.39-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.58 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.192 , 0.235 0.195 , 0.234	Depositor DCC
R_{free} test set	5365 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	61.4	Xtrriage
Anisotropy	0.311	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 41.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.016 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	19771	wwPDB-VP
Average B, all atoms (Å ²)	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.89% 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](#)

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	AAA	0.65	0/10110	0.76	0/13707
1	BBB	0.66	0/10046	0.74	0/13623
All	All	0.65	0/20156	0.75	0/27330

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	AAA	9913	0	9804	76	0
1	BBB	9850	0	9727	79	0
2	AAA	5	0	0	0	0
2	BBB	3	0	0	0	0
All	All	19771	0	19531	154	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 4.

The worst 5 of 154 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:AAA:109:ASP:HB3	1:AAA:212:TYR:O	1.84	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:644:ILE:HG23	1:AAA:649:ASN:HD22	1.53	0.73
1:BBB:526:ARG:HB3	1:BBB:559:ASN:HD22	1.61	0.65
1:AAA:538:TYR:OH	1:AAA:643:LEU:HD12	1.97	0.64
1:AAA:1022:VAL:HG13	1:AAA:1026:GLU:HB3	1.81	0.63

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	AAA	1238/1351 (92%)	1173 (95%)	61 (5%)	4 (0%)	41 66
1	BBB	1231/1351 (91%)	1164 (95%)	62 (5%)	5 (0%)	34 60
All	All	2469/2702 (91%)	2337 (95%)	123 (5%)	9 (0%)	34 60

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	468	VAL
1	AAA	942	ALA
1	BBB	115	PRO
1	BBB	942	ALA
1	AAA	467	ASN

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	AAA	1075/1147 (94%)	1060 (99%)	15 (1%)	67	86
1	BBB	1067/1147 (93%)	1048 (98%)	19 (2%)	59	83
All	All	2142/2294 (93%)	2108 (98%)	34 (2%)	62	85

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	1111	GLU
1	BBB	1176	PHE
1	BBB	1229	TYR
1	AAA	1129	LYS
1	AAA	1022	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

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 [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	AAA	1258/1351 (93%)	0.24	23 (1%) 68 70	36, 64, 110, 185	37 (2%)
1	BBB	1251/1351 (92%)	0.36	48 (3%) 40 39	45, 78, 116, 189	36 (2%)
All	All	2509/2702 (92%)	0.30	71 (2%) 53 54	36, 71, 113, 189	73 (2%)

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	145	MET	8.1
1	AAA	143	GLN	5.9
1	BBB	853	VAL	4.8
1	BBB	145	MET	4.7
1	BBB	139	SER	4.6

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.