



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 03:23 am BST

PDB ID : 3N6Q
Title : Crystal structure of YghZ from E. coli
Authors : Zubieta, C.; Totir, M.; Echols, N.; May, A.; Alber, T.
Deposited on : 2010-05-26
Resolution : 1.80 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 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	:	2.11
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)	:	2.11

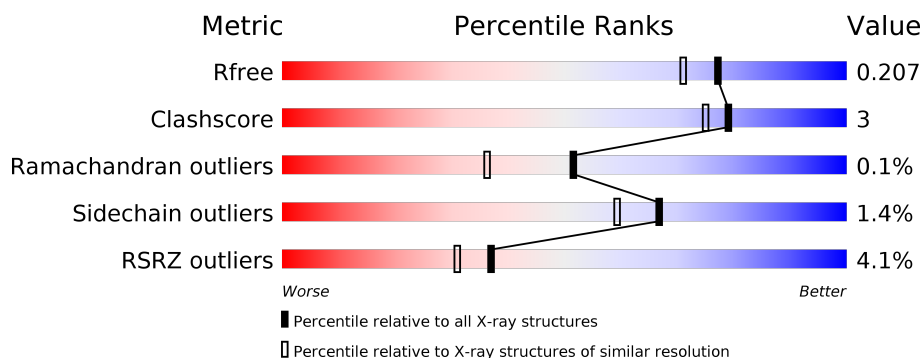
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 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	A	346	<div> <div>85%</div> <div>5%</div> <div>9%</div> </div>
1	B	346	<div> <div>86%</div> <div>5%</div> <div>9%</div> </div>
1	C	346	<div> <div>3%</div> <div>86%</div> <div>•</div> <div>10%</div> </div>
1	D	346	<div> <div>85%</div> <div>6%</div> <div>•</div> <div>9%</div> </div>
1	E	346	<div> <div>8%</div> <div>80%</div> <div>9%</div> <div>•</div> <div>11%</div> </div>
1	F	346	<div> <div>2%</div> <div>85%</div> <div>5%</div> <div>10%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	346	<div><div></div><div>6%</div><div>84%</div><div>6%</div><div>10%</div></div>
1	H	346	<div><div></div><div>8%</div><div>81%</div><div>7%</div><div>12%</div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 22603 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 YghZ aldo-keto reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	314	Total	C	N	O	S	0	14	0
			2556	1624	445	476	11			
1	B	315	Total	C	N	O	S	0	10	0
			2547	1618	445	473	11			
1	C	313	Total	C	N	O	S	0	4	0
			2491	1575	438	468	10			
1	D	315	Total	C	N	O	S	0	2	0
			2495	1580	436	469	10			
1	E	309	Total	C	N	O	S	0	4	0
			2457	1553	434	460	10			
1	F	313	Total	C	N	O	S	0	3	0
			2481	1569	436	466	10			
1	G	313	Total	C	N	O	S	0	4	0
			2489	1574	439	466	10			
1	H	306	Total	C	N	O	S	0	3	0
			2436	1544	429	453	10			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	1	Total	Mg	0	0
			1	1		
2	D	2	Total	Mg	0	0
			2	2		
2	E	1	Total	Mg	0	0
			1	1		
2	H	1	Total	Mg	0	0
			1	1		
2	B	1	Total	Mg	0	0
			1	1		
2	C	1	Total	Mg	0	0
			1	1		

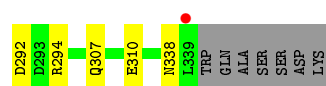
Continued on next page...

Continued from previous page...

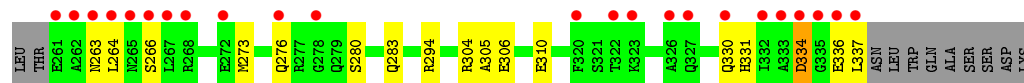
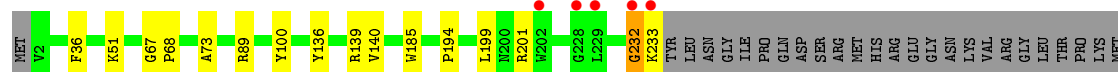
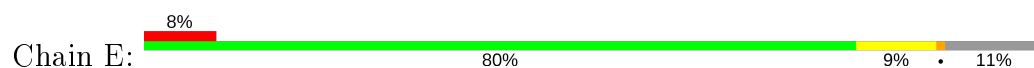
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Mg 1	0	0
2	F	1	Total 1	Mg 1	0	0

- Molecule 3 is water.

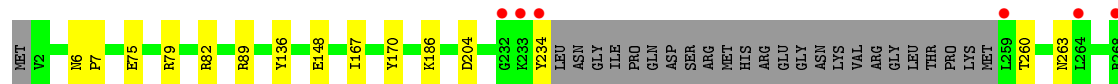
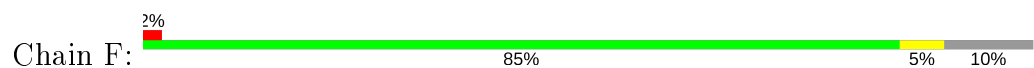
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	444	Total 444	O 444	0	0
3	B	382	Total 382	O 382	0	0
3	C	298	Total 298	O 298	0	0
3	D	343	Total 343	O 343	0	0
3	E	254	Total 254	O 254	0	0
3	F	362	Total 362	O 362	0	0
3	G	323	Total 323	O 323	0	0
3	H	236	Total 236	O 236	0	0



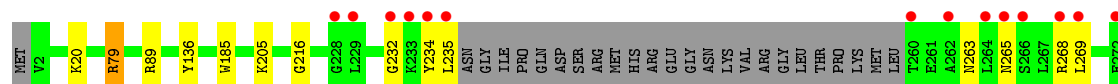
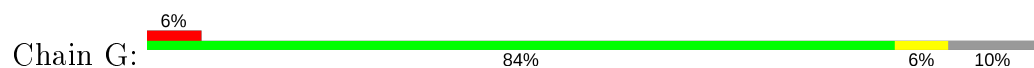
- Molecule 1: YghZ ald-keto reductase



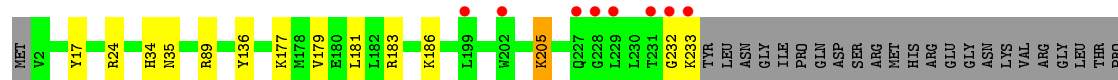
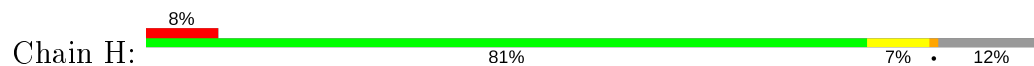
- Molecule 1: YghZ ald-keto reductase



- Molecule 1: YghZ ald-keto reductase



- Molecule 1: YghZ ald-keto reductase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	91.70Å 98.06Å 98.26Å 90.27° 92.97° 106.12°	Depositor
Resolution (Å)	94.07 – 1.80 67.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	96.2 (94.07-1.80) 96.2 (67.37-1.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.32 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.169 , 0.209 0.168 , 0.207	Depositor DCC
R_{free} test set	14824 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	24.2	Xtriage
Anisotropy	0.054	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 53.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	22603	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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: MG

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	0.97	1/2649 (0.0%)	0.83	2/3582 (0.1%)
1	B	1.03	0/2629	0.87	3/3557 (0.1%)
1	C	0.82	0/2554	0.76	1/3457 (0.0%)
1	D	0.88	1/2552 (0.0%)	0.76	1/3456 (0.0%)
1	E	0.73	0/2519	0.68	0/3408
1	F	0.95	1/2541 (0.0%)	0.82	0/3440
1	G	1.00	2/2552 (0.1%)	1.54	6/3455 (0.2%)
1	H	0.76	0/2497	0.72	0/3381
All	All	0.90	5/20493 (0.0%)	0.91	13/27736 (0.0%)

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	2
1	B	0	3
1	C	0	1
1	D	0	1
1	E	0	2
1	F	0	1
1	G	0	1
1	H	0	1
All	All	0	12

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	79[A]	ARG	CD-NE	-18.10	1.15	1.46
1	G	79[B]	ARG	CD-NE	-18.10	1.15	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	148	GLU	CB-CG	5.70	1.62	1.52
1	D	164	TYR	CD1-CE1	5.30	1.47	1.39
1	A	173	GLU	CG-CD	5.00	1.59	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	79[A]	ARG	NE-CZ-NH2	-41.88	99.36	120.30
1	G	79[B]	ARG	NE-CZ-NH2	-41.88	99.36	120.30
1	G	79[A]	ARG	NE-CZ-NH1	35.72	138.16	120.30
1	G	79[B]	ARG	NE-CZ-NH1	35.72	138.16	120.30
1	G	79[A]	ARG	CG-CD-NE	11.46	135.86	111.80

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	107	TYR	Sidechain
1	A	136	TYR	Peptide
1	B	107	TYR	Sidechain
1	B	136	TYR	Peptide
1	B	60	PHE	Sidechain

5.2 Too-close contacts

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	2556	0	2568	24	1
1	B	2547	0	2541	8	4
1	C	2491	0	2459	7	0
1	D	2495	0	2458	12	0
1	E	2457	0	2430	18	0
1	F	2481	0	2440	13	1
1	G	2489	0	2456	12	4
1	H	2436	0	2404	18	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	2	0	0	0	1
2	E	1	0	0	0	1
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	444	0	0	5	0
3	B	382	0	0	3	0
3	C	298	0	0	1	0
3	D	343	0	0	2	0
3	E	254	0	0	0	0
3	F	362	0	0	1	0
3	G	323	0	0	0	0
3	H	236	0	0	1	0
All	All	22603	0	19756	110	6

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 3.

The worst 5 of 110 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:327[B]:GLN:NE2	3:B:572:HOH:O	1.62	1.28
1:A:79[B]:ARG:HH11	1:A:79[B]:ARG:CG	1.60	1.11
1:A:82[B]:ARG:HH11	1:A:82[B]:ARG:HG3	1.02	1.10
1:A:79[B]:ARG:HH11	1:A:79[B]:ARG:HG3	1.27	0.96
1:A:82[B]:ARG:HH11	1:A:82[B]:ARG:CG	1.79	0.94

The worst 5 of 6 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:402:MG:MG	2:E:401:MG:MG[1_655]	1.28	0.92
1:A:79[B]:ARG:NH2	1:F:75:GLU:OE1[1_655]	1.52	0.68
1:B:79:ARG:NH2	1:G:79[B]:ARG:NH2[1_655]	2.00	0.20
1:B:79:ARG:CZ	1:G:79[B]:ARG:NH2[1_655]	2.04	0.16
1:B:79:ARG:NH1	1:G:79[B]:ARG:CZ[1_655]	2.08	0.12

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	324/346 (94%)	317 (98%)	7 (2%)	0	100	100
1	B	321/346 (93%)	311 (97%)	10 (3%)	0	100	100
1	C	313/346 (90%)	301 (96%)	12 (4%)	0	100	100
1	D	313/346 (90%)	304 (97%)	8 (3%)	1 (0%)	41	27
1	E	309/346 (89%)	296 (96%)	13 (4%)	0	100	100
1	F	312/346 (90%)	303 (97%)	9 (3%)	0	100	100
1	G	313/346 (90%)	301 (96%)	11 (4%)	1 (0%)	41	27
1	H	305/346 (88%)	292 (96%)	13 (4%)	0	100	100
All	All	2510/2768 (91%)	2425 (97%)	83 (3%)	2 (0%)	51	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	338	ASN
1	G	263	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	A	275/290 (95%)	274 (100%)	1 (0%)	91	89
1	B	272/290 (94%)	270 (99%)	2 (1%)	84	81
1	C	263/290 (91%)	260 (99%)	3 (1%)	73	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	262/290 (90%)	257 (98%)	5 (2%)	57	46
1	E	259/290 (89%)	252 (97%)	7 (3%)	44	31
1	F	260/290 (90%)	258 (99%)	2 (1%)	81	78
1	G	262/290 (90%)	259 (99%)	3 (1%)	73	68
1	H	256/290 (88%)	250 (98%)	6 (2%)	50	37
All	All	2109/2320 (91%)	2080 (99%)	29 (1%)	67	59

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	201	ARG
1	E	330	GLN
1	H	304	ARG
1	E	263	ASN
1	E	334	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	193	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	314/346 (90%)	-0.39	3 (0%) 82 80	11, 18, 34, 61	0
1	B	315/346 (91%)	-0.46	4 (1%) 77 74	12, 19, 34, 58	0
1	C	313/346 (90%)	-0.20	9 (2%) 51 46	15, 26, 49, 70	0
1	D	315/346 (91%)	-0.39	4 (1%) 77 74	15, 23, 40, 66	0
1	E	309/346 (89%)	-0.04	28 (9%) 9 7	18, 29, 67, 85	1 (0%)
1	F	313/346 (90%)	-0.34	7 (2%) 62 57	14, 22, 44, 70	3 (0%)
1	G	313/346 (90%)	-0.12	20 (6%) 19 15	15, 24, 61, 82	1 (0%)
1	H	306/346 (88%)	0.10	27 (8%) 10 7	19, 30, 65, 95	1 (0%)
All	All	2498/2768 (90%)	-0.23	102 (4%) 37 31	11, 24, 55, 95	6 (0%)

The worst 5 of 102 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	262	ALA	8.2
1	H	228	GLY	7.7
1	E	261	GLU	6.2
1	H	260	THR	6.2
1	B	258	MET	6.1

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 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. 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
2	MG	H	401	1/1	0.95	0.20	13,13,13,13	0
2	MG	D	401	1/1	0.96	0.19	17,17,17,17	0
2	MG	G	401	1/1	0.96	0.19	12,12,12,12	0
2	MG	B	401	1/1	0.97	0.21	12,12,12,12	0
2	MG	C	401	1/1	0.98	0.16	19,19,19,19	0
2	MG	F	401	1/1	0.98	0.17	10,10,10,10	0
2	MG	D	402	1/1	0.99	0.10	54,54,54,54	0
2	MG	A	401	1/1	0.99	0.15	12,12,12,12	0
2	MG	E	401	1/1	1.00	0.09	15,15,15,15	0

6.5 Other polymers [i](#)

There are no such residues in this entry.