wwPDB X-ray Structure Validation Summary Report (i)

Jan 17, 2023 – 05:25 PM EST

PDB ID	:	1VKH
Title	:	CRYSTAL STRUCTURE OF A PUTATIVE SERINE HYDROLASE (YDR428C) FROM SACCHAROMYCES CEREVISIAE AT 1.85 A RESOLUTION
Authors	:	Joint Center for Structural Genomics (JCSG)
Deposited on	:	2004-05-20
Resolution	:	1.85 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:

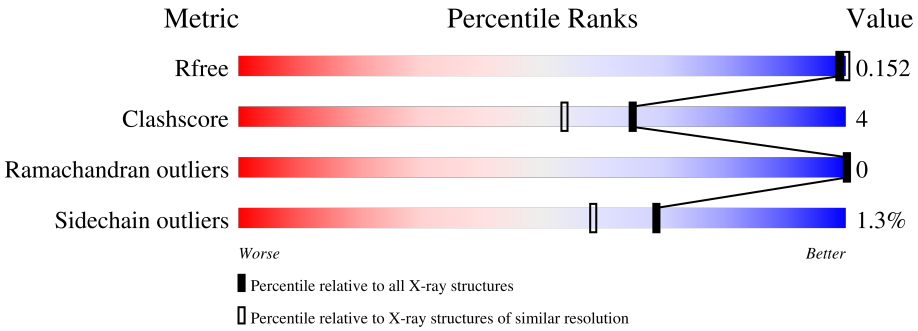
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)

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%

Mol	Chain	Length	Quality of chain
1	A	273	86% 8% • •
1	B	273	83% 8% • 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4659 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 putative serine hydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	261	Total	C	N	O	S	Se	0	7	0
			2120	1355	345	404	7	9			
1	B	252	Total	C	N	O	S	Se	0	10	0
			2067	1324	334	393	7	9			

There are 42 discrepancies between the modelled and reference sequences:

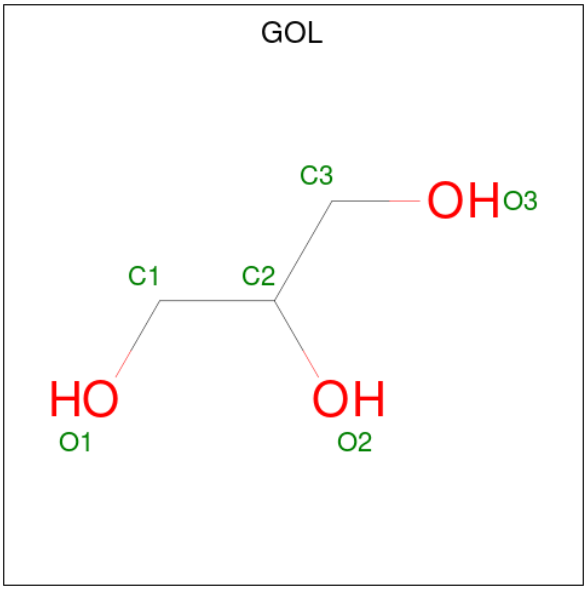
Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	expression tag	UNP Q04066
A	-10	GLY	-	expression tag	UNP Q04066
A	-9	SER	-	expression tag	UNP Q04066
A	-8	ASP	-	expression tag	UNP Q04066
A	-7	LYS	-	expression tag	UNP Q04066
A	-6	ILE	-	expression tag	UNP Q04066
A	-5	HIS	-	expression tag	UNP Q04066
A	-4	HIS	-	expression tag	UNP Q04066
A	-3	HIS	-	expression tag	UNP Q04066
A	-2	HIS	-	expression tag	UNP Q04066
A	-1	HIS	-	expression tag	UNP Q04066
A	0	HIS	-	expression tag	UNP Q04066
A	1	MSE	MET	modified residue	UNP Q04066
A	60	MSE	MET	modified residue	UNP Q04066
A	106	MSE	MET	modified residue	UNP Q04066
A	130	MSE	MET	modified residue	UNP Q04066
A	137	MSE	MET	modified residue	UNP Q04066
A	161	ILE	VAL	SEE REMARK 999	UNP Q04066
A	180	MSE	MET	modified residue	UNP Q04066
A	189	MSE	MET	modified residue	UNP Q04066
A	203	MSE	MET	modified residue	UNP Q04066
B	-11	MET	-	expression tag	UNP Q04066
B	-10	GLY	-	expression tag	UNP Q04066
B	-9	SER	-	expression tag	UNP Q04066
B	-8	ASP	-	expression tag	UNP Q04066

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-7	LYS	-	expression tag	UNP Q04066
B	-6	ILE	-	expression tag	UNP Q04066
B	-5	HIS	-	expression tag	UNP Q04066
B	-4	HIS	-	expression tag	UNP Q04066
B	-3	HIS	-	expression tag	UNP Q04066
B	-2	HIS	-	expression tag	UNP Q04066
B	-1	HIS	-	expression tag	UNP Q04066
B	0	HIS	-	expression tag	UNP Q04066
B	1	MSE	MET	modified residue	UNP Q04066
B	60	MSE	MET	modified residue	UNP Q04066
B	106	MSE	MET	modified residue	UNP Q04066
B	130	MSE	MET	modified residue	UNP Q04066
B	137	MSE	MET	modified residue	UNP Q04066
B	161	ILE	VAL	SEE REMARK 999	UNP Q04066
B	180	MSE	MET	modified residue	UNP Q04066
B	189	MSE	MET	modified residue	UNP Q04066
B	203	MSE	MET	modified residue	UNP Q04066

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	A	1	Total C O 6 3 3	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 2 2	0	1

- Molecule 4 is water.

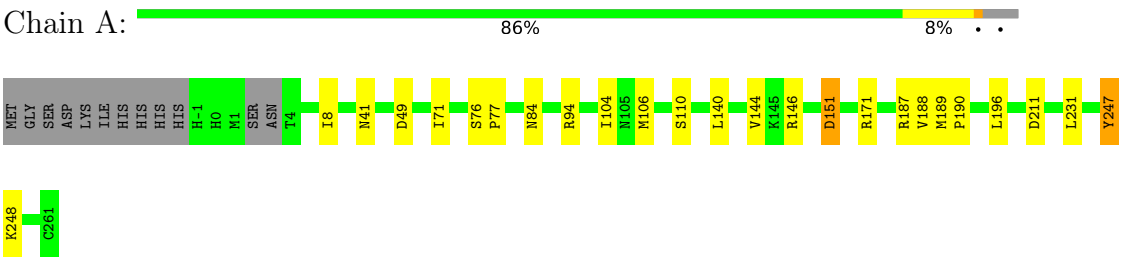
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	232	Total O 233 233	0	1
4	B	222	Total O 225 225	0	3

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. 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. 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.

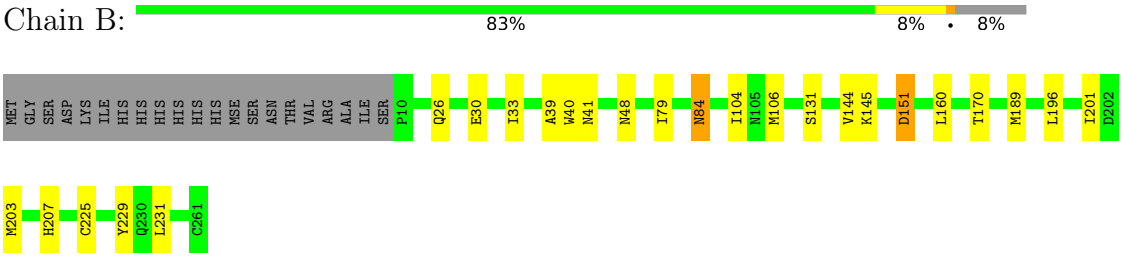
- Molecule 1: putative serine hydrolase

Chain A:



- Molecule 1: putative serine hydrolase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	41.97Å 85.96Å 70.48Å 90.00° 97.71° 90.00°	Depositor
Resolution (Å)	37.44 – 1.85 37.44 – 1.85	Depositor EDS
% Data completeness (in resolution range)	100.0 (37.44-1.85) 99.7 (37.44-1.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	3.03 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.1.9999	Depositor
R, R_{free}	0.143 , 0.189 0.154 , 0.152	Depositor DCC
R_{free} test set	2134 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	17.8	Xtriage
Anisotropy	0.578	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 42.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4659	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.35% 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: GOL, CL

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	0.93	2/2183 (0.1%)	0.93	9/2944 (0.3%)
1	B	0.87	0/2142	0.91	1/2890 (0.0%)
All	All	0.90	2/4325 (0.0%)	0.92	10/5834 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	188	VAL	CB-CG2	5.13	1.63	1.52
1	A	247	TYR	CE2-CZ	-5.09	1.31	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	171	ARG	NE-CZ-NH1	6.38	123.49	120.30
1	B	151	ASP	CB-CG-OD2	6.32	123.99	118.30
1	A	41[A]	ASN	N-CA-CB	-5.56	100.59	110.60
1	A	41[B]	ASN	N-CA-CB	-5.56	100.59	110.60
1	A	151	ASP	CB-CG-OD2	5.48	123.23	118.30

There are no chirality outliers.

There are no planarity outliers.

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	2120	0	2094	12	0
1	B	2067	0	2048	22	0
2	A	12	0	16	0	0
3	B	2	0	0	1	0
4	A	233	0	0	1	0
4	B	225	0	0	1	0
All	All	4659	0	4158	34	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 34 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:A:106[B]:MSE:HE2	1:A:144:VAL:HG21	1.53	0.90
1:B:106[B]:MSE:HE2	1:B:144[B]:VAL:HG21	1.60	0.83
1:B:33:ILE:HB	1:B:106[A]:MSE:HE1	1.71	0.73
1:B:196:LEU:HD21	1:B:203[B]:MSE:SE	2.43	0.69
1:B:30:GLU:OE2	1:B:145[A]:LYS:NZ	2.32	0.62

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	264/273 (97%)	258 (98%)	6 (2%)	0	100	100
1	B	260/273 (95%)	256 (98%)	4 (2%)	0	100	100
All	All	524/546 (96%)	514 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	243/242 (100%)	240 (99%)	3 (1%)	71 62
1	B	239/242 (99%)	236 (99%)	3 (1%)	69 58
All	All	482/484 (100%)	476 (99%)	6 (1%)	69 62

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	26	GLN
1	B	41	ASN
1	B	84	ASN
1	A	110	SER
1	A	84	ASN

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

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 4 ligands modelled in this entry, 2 are 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	$\#\|Z\| > 2$
2	GOL	A	456	-	5,5,5	0.55	0	5,5,5	0.49	0
2	GOL	A	457	-	5,5,5	0.50	0	5,5,5	0.58	0

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	GOL	A	456	-	-	2/4/4/4	-
2	GOL	A	457	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	456	GOL	O1-C1-C2-C3
2	A	457	GOL	O1-C1-C2-C3
2	A	456	GOL	O1-C1-C2-O2
2	A	457	GOL	O1-C1-C2-O2

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.