wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 10:27 am BST

PDB ID	:	1XUP
Title	:	ENTEROCOCCUS CASSELIFLAVUS GLYCEROL KINASE COMPLEXED WITH GLYCEROL
Authors	:	Yeh, J.I.; Charrier, V.; Paulo, J.; Hou, L.; Darbon, E.; Hol, W.G.J.; Deutscher, J.
Deposited on	:	2004-10-26
Resolution	:	2.75 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 following versions of software and data (see references (i)) were used in the production of this report:

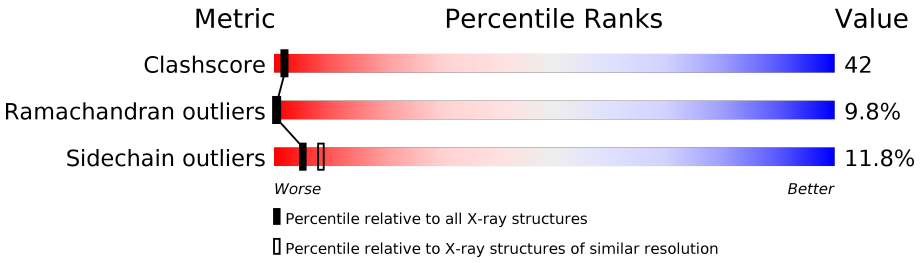
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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))
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	O	487	41% 46% 10% •
1	X	487	50% 41% 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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	O	500	-	-	X	-
2	GOL	X	501	-	-	X	-

2 Entry composition

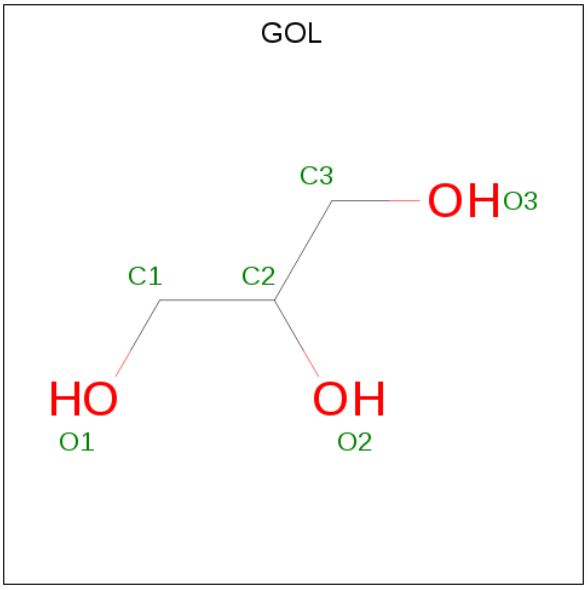
There are 2 unique types of molecules in this entry. The entry contains 7557 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Glycerol kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	487	Total	C	N	O	S	0	0	0
			3776	2392	629	741	14			
1	X	487	Total	C	N	O	S	0	0	0
			3769	2387	627	741	14			

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



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

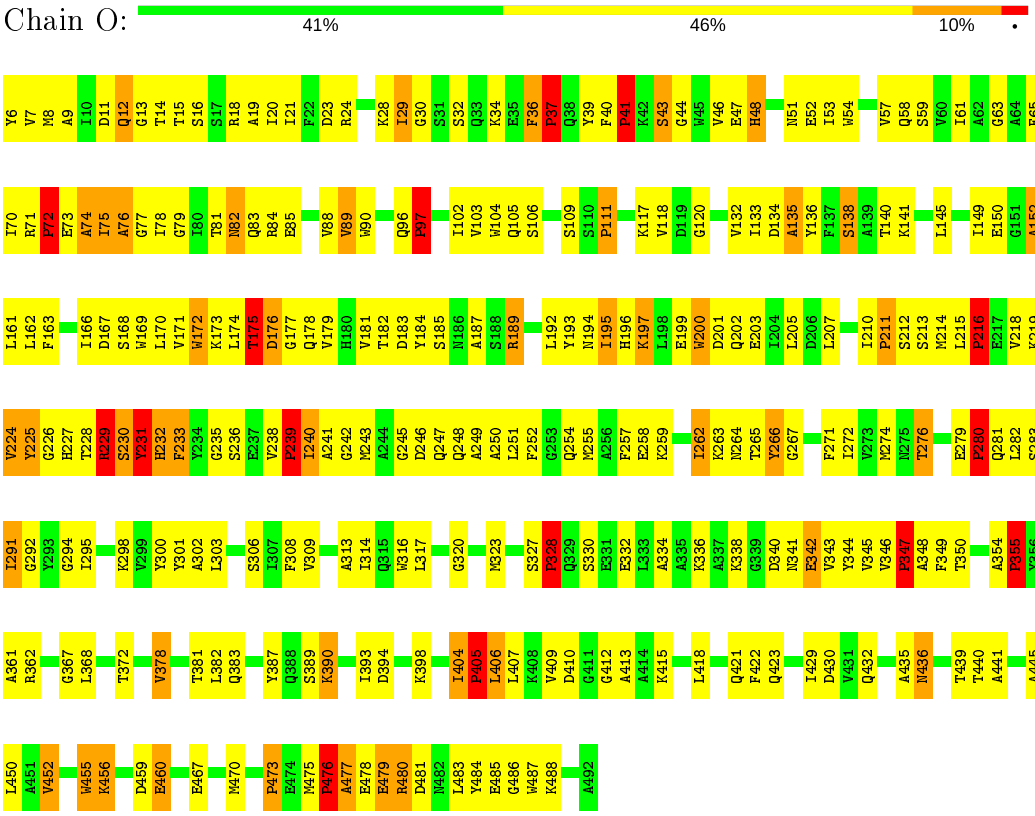
3 Residue-property plots (i)

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

Note EDS was not executed.

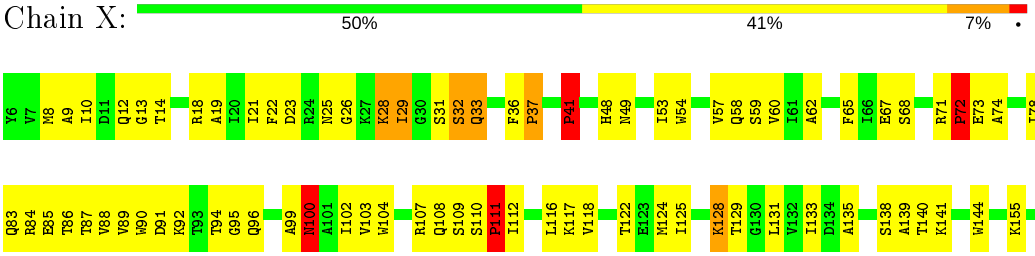
- Molecule 1: Glycerol kinase

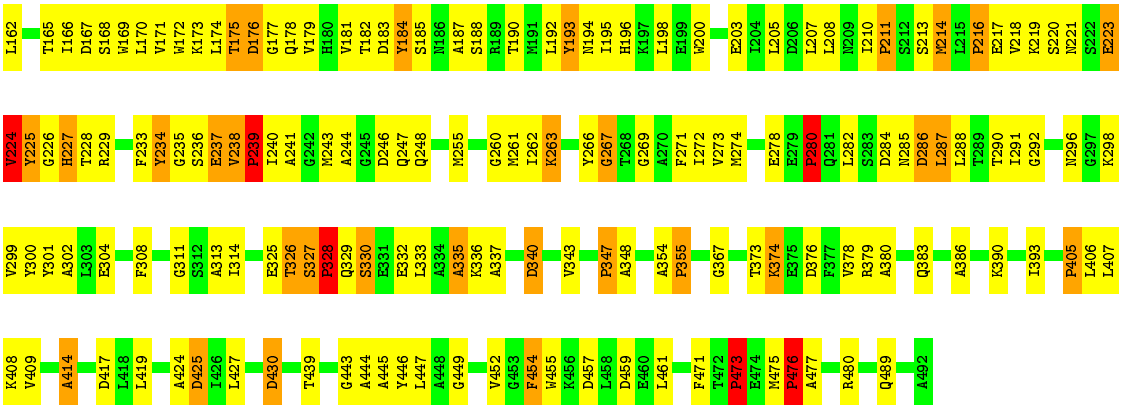
Chain O:



- Molecule 1: Glycerol kinase

Chain X:




L162 T165 I166 D167 S168 W169 L170 V171 W172 K173 L174 T175 D176 G177 Q178 V179 H180 V181 T182 D183 Y184 S185 N186 A187 S188 R189 T190 M191 L192 Y193 N194 I195 H196 K197 L198 E199 W200 E203 I204 L205 D206 L207 L208 N209 I210 P211 S212 S213 M214 L215 P216 E217 V218 K219 S220 M221 S222 E223
V224 V225 G226 H227 T228 R229 F233 Y234 G235 S236 E237 V238 P239 I240 A241 G242 M243 A244 G245 D246 Q247 Q248 M255 G260 M261 I262 K263 Y266 G267 T268 G269 A270 F271 I272 V273 M274 E278 E279 P280 Q281 L282 S283 D284 N285 D286 L287 L288 T289 T290 I291 G292 N296 G297 K298
V299 Y300 Y301 A302 L303 E304 F308 G311 S312 A313 I314 E325 T326 S327 P328 Q329 S330 E331 E332 L333 A334 A335 K336 A337 D340 V343 P347 A348 A354 P355 G367 T373 K374 E375 D376 F377 V378 R379 A380 Q383 A386 K390 I393 P405 L406 L407
K408 V409 A414 D417 L418 L419 A424 D425 T426 L427 D430 T439 G443 A444 A445 Y446 L447 A448 G449 V452 G453 F454 W455 K456 D457 L458 D459 E460 L461 F471 T472 P473 E474 M475 P476 A477 R480 Q489 A492

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	68.01Å 107.67Å 201.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.75	Depositor
% Data completeness (in resolution range)	87.0 (6.00-2.75)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.242 , 0.268	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7557	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL

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	O	0.48	0/3855	1.08	23/5229 (0.4%)
1	X	0.43	0/3847	0.96	19/5218 (0.4%)
All	All	0.45	0/7702	1.02	42/10447 (0.4%)

There are no bond length outliers.

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	O	355	PRO	CA-N-CD	-18.62	85.44	111.50
1	O	405	PRO	CA-N-CD	-17.28	87.30	111.50
1	O	347	PRO	CA-N-CD	-16.64	88.20	111.50
1	O	111	PRO	CA-N-CD	-15.73	89.48	111.50
1	X	405	PRO	CA-N-CD	-14.63	91.01	111.50

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	O	3776	0	3673	323	0
1	X	3769	0	3661	326	0
2	O	6	0	8	9	0
2	X	6	0	8	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7557	0	7350	626	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 42.

The worst 5 of 626 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:X:175:THR:HG21	1:X:179:VAL:CG1	1.41	1.50
1:X:175:THR:HB	1:X:227:HIS:CB	1.46	1.43
1:X:84:ARG:CG	2:X:501:GOL:H12	1.58	1.33
1:X:167:ASP:HA	1:X:243:MET:CE	1.60	1.31
1:X:32:SER:O	1:X:33:GLN:NE2	1.61	1.31

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	O	485/487 (100%)	330 (68%)	103 (21%)	52 (11%)	0	0
1	X	485/487 (100%)	330 (68%)	112 (23%)	43 (9%)	1	0
All	All	970/974 (100%)	660 (68%)	215 (22%)	95 (10%)	0	0

5 of 95 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	O	24	ARG
1	O	117	LYS
1	O	135	ALA
1	O	138	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	O	203	GLU

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	O	395/395 (100%)	345 (87%)	50 (13%)	4 7
1	X	393/395 (100%)	350 (89%)	43 (11%)	6 10
All	All	788/790 (100%)	695 (88%)	93 (12%)	5 8

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	406	LEU
1	X	37	PRO
1	X	406	LEU
1	O	422	PHE
1	O	467	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	O	254	GLN
1	O	421	GLN
1	X	108	GLN
1	O	247	GLN
1	X	202	GLN

5.3.3 RNA

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)

2 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	GOL	O	500	-	5,5,5	0.79	0	5,5,5	0.92	0
2	GOL	X	501	-	5,5,5	0.78	0	5,5,5	0.93	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	O	500	-	-	1/4/4/4	-
2	GOL	X	501	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	O	500	GOL	O1-C1-C2-C3
2	X	501	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	O	500	GOL	9	0
2	X	501	GOL	14	0

5.7 Other polymers

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.