



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 03:23 am BST

PDB ID : 1BOT  
Title : CRYSTAL STRUCTURE OF THE COMPLEX BETWEEN ESCHERICHIA COLI GLYCEROL KINASE AND THE ALLOSTERIC REGULATOR FRUCTOSE 1,6-BISPHOSPHATE.  
Authors : Ormo, M.; Bystrom, C.E.; Remington, S.J.  
Deposited on : 1998-08-05  
Resolution : 3.05 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](mailto: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
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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

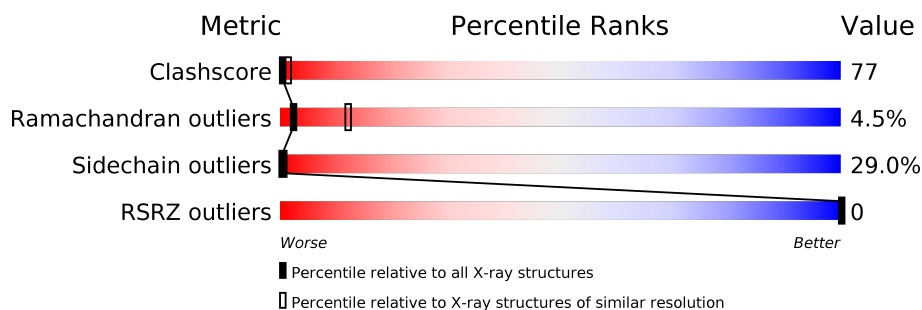
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

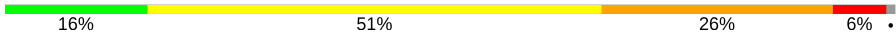
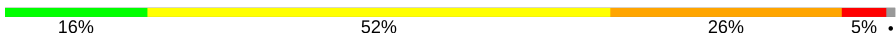
The reported resolution of this entry is 3.05 Å.

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	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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  $\geq 3$ , 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	O	501	
1	Z	501	

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
3	GOL	O	601	-	-	X	-

## 2 Entry composition [i](#)

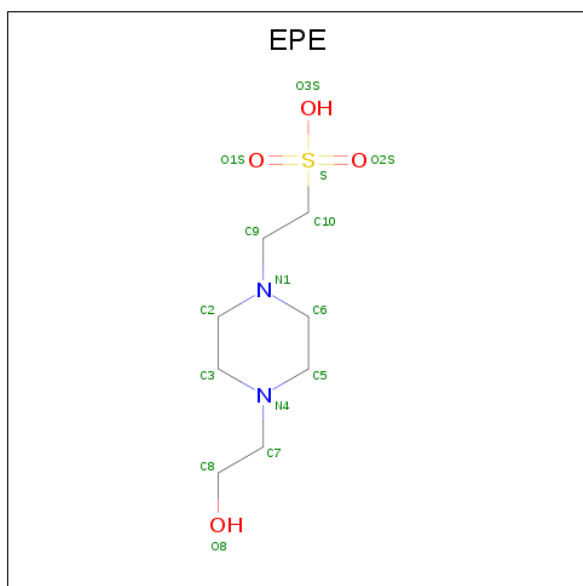
There are 3 unique types of molecules in this entry. The entry contains 7864 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 PROTEIN (GLYCEROL KINASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	497	Total	C	N	O	S	0	0	0
			3914	2469	686	740	19			
1	Z	498	Total	C	N	O	S	0	0	0
			3923	2474	687	743	19			

- Molecule 2 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	O	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



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

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 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 ( $\text{RSRZ} > 2$ ). 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.

[illegible]

Chain Z: 16% 52% 26% 5%

Category	Category
V63	THR
L64	E2
A65	K3
K66	K4
A67	Y5
D68	I6
I69	V7
S70	A8
S71	L9
D72	D10
Q73	Q11
I74	G12
A75	T13
I79	T14
T80	S15
N81	S16
Q82	R17
R83	A18
E84	V19
T85	W20
T86	M21
I87	D22
V88	H23
W89	D24
E90	A25
Q91	M26
E92	I27
T93	I28
Q94	S29
K95	V30
P96	S31
I97	Q32
Y98	R33
N99	E34
A100	F35
I101	E36
V102	Q37
W103	I38
C104	Y39
C105	P40
R106	K41
R107	P42
T108	G43
A109	W44
E110	V45
I111	E46
C112	H47
E113	D48
H114	P49
L115	M50
K116	L15
R117	I52
D118	W53
G119	A54
L120	T55
D121	Q56
D122	L60
Y123	V61
I124	R62

GLU	R436	N374	S312	L250	W485	R125
	E437	H375	I313	F251	A486	S126
	V438	I376	Q314	G252	S187	M127
	T439	I377	W315	Q253	R188	T128
	A440	R378	L316	L254	T189	G129
	L441	L381	R317	C255	M190	L130
	G442		D318	V256	L191	V131
	A443	E382	E319	K257	F192	I132
	A444	S383	M320	E258	M193	D133
	L445	I384	K321	G259	I194	P134
	L446	A385	L322	M260	H195	P135
	L449	Y386	I323	A261	T196	F136
		Q387	M324	K262	L197	S137
	A450	T388	D325	M263	D198	G138
	V451	R389	A326	T264	W199	T139
	G452	D390	Y327	Y265	D200	K140
	F453	V391	D328	G266	D201	V141
	W454	L392	S329	T267	K202	K142
	Q455	E393	E330	G268	M203	W143
	M456	A394	Y331	C269	L204	I144
	L457	M395	F332	F270	E205	L145
	D458	Q396	A333	M271	V206	D146
	E459	A397	T334	L272	L207	H147
	L460	D398	K335	M273	D208	V148
	Q461	S399	W336	M274	I209	E149
	E462	G400	Q337	T275	P210	G150
	K463	I401	M338	G276	R211	S151
	A464	R402	T339	E277	E212	R152
	V465	L403	M340	K278	M213	E153
	I466	L406	G341	A279	L214	R154
	E467		V342	V280	P215	A155
	R468	R407	Y343	K281	E216	R156
	E469	V408	V344	S282	V217	R157
	F470	D409	V345	E283	R218	G158
	R471	G410	P346	L286	R219	E159
	P472	G411	A347		S220	L160
	G473	A412	F348	S221	L161	L161
	I474	V413	T349	E222	F162	F162
	E475	A414	G350	V223	G163	G163
	T476	M415	L351	Y224	T164	T164
	T477	M416	G352	C292	G225	V165
	E478	F417	A353	G293	Q226	D166
	R479	L418	P354	P294	T227	T167
	M480	M419	Y355	T295	N228	W168
	Y481	Q420	W356	G296	I229	L169
	R482	F421	D357	E297	T235	L170
	Y483	Q422	P358	V298		W171
	W486	S423	Y359	N299	R236	K172
		D424	A360	Y300	I237	M173
	K487	I425	R361	A301	P238	T174
	K488	L426	G362	L302	I239	Q175
	A489	G427	A363	E303	S240	G176
	V490	T428	I364	G304	G241	R177
	K491	R429	F365	A305	I242	V178
	R492	V430	V306	V306	A243	H179
	E497	E431	T368	F307	G244	V180
		R432	V371	M308	D245	T181
	E498	P433		A309	Q246	D182
	H499	E434	N372	G310	A249	Y183
	ASP	V435	A373	A311		T184

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.80Å 168.80Å 202.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.05 19.98 – 2.90	Depositor EDS
% Data completeness (in resolution range)	81.0 (20.00-3.05) 73.8 (19.98-2.90)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 2.88Å)	Xtriage
Refinement program	TNT 5F	Depositor
R, $R_{free}$	0.219 , (Not available) 0.199 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.2	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 107.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	7864	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.30% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: GOL, EPE

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	1.20	34/3994 (0.9%)	1.57	57/5414 (1.1%)
1	Z	1.21	37/4003 (0.9%)	1.56	65/5426 (1.2%)
All	All	1.21	71/7997 (0.9%)	1.56	122/10840 (1.1%)

The worst 5 of 71 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	Z	34	GLU	CD-OE2	11.22	1.38	1.25
1	O	216	GLU	CD-OE2	10.60	1.37	1.25
1	O	437	GLU	CD-OE1	8.56	1.35	1.25
1	Z	475	GLU	CD-OE2	8.45	1.34	1.25
1	O	462	GLU	CD-OE1	8.39	1.34	1.25

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	O	357	ASP	C-N-CD	-13.79	90.26	120.60
1	Z	152	ARG	NE-CZ-NH1	12.76	126.68	120.30
1	O	200	ASP	CB-CG-OD1	-9.71	109.56	118.30
1	O	409	ASP	CB-CG-OD2	-9.62	109.64	118.30
1	Z	10	ASP	CB-CG-OD1	-8.85	110.34	118.30

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	O	3914	0	3853	631	0
1	Z	3923	0	3859	572	0
2	O	15	0	18	2	0
3	O	6	0	8	4	0
3	Z	6	0	8	2	0
All	All	7864	0	7746	1204	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 77.

The worst 5 of 1204 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:Z:329:SER:HB2	1:Z:381:LEU:HD11	1.28	1.14
1:Z:84:GLU:HB2	1:Z:103:TRP:HB3	1.15	1.12
1:O:145:LEU:HD12	1:O:151:SER:HB2	1.23	1.11
1:Z:228:ASN:HB2	1:Z:236:ARG:HD2	1.28	1.09
1:O:17:ARG:HD2	1:O:32:GLN:HE21	1.17	1.07

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	495/501 (99%)	370 (75%)	94 (19%)	31 (6%)	1	7
1	Z	496/501 (99%)	397 (80%)	85 (17%)	14 (3%)	5	21
All	All	991/1002 (99%)	767 (77%)	179 (18%)	45 (4%)	2	12

5 of 45 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	O	99	ASN
1	O	109	ALA
1	O	149	GLU
1	O	151	SER
1	O	358	PRO

### 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	O	408/412 (99%)	293 (72%)	115 (28%)	0	1
1	Z	409/412 (99%)	287 (70%)	122 (30%)	0	0
All	All	817/824 (99%)	580 (71%)	237 (29%)	0	1

5 of 237 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	O	463	LYS
1	Z	74	ILE
1	Z	456	ASN
1	O	476	THR
1	Z	17	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	Z	11	GLN
1	Z	104	GLN
1	Z	422	GLN
1	Z	47	HIS
1	Z	114	HIS

### 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](#)

3 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	EPE	O	602	-	15,15,15	2.10	3 (20%)	18,20,20	1.97	4 (22%)
3	GOL	Z	603	-	5,5,5	0.47	0	5,5,5	0.51	0
3	GOL	O	601	-	5,5,5	0.89	0	5,5,5	0.63	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	EPE	O	602	-	-	3/9/19/19	0/1/1/1
3	GOL	Z	603	-	-	0/4/4/4	-
3	GOL	O	601	-	-	2/4/4/4	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	602	EPE	O3S-S	5.74	1.67	1.47
2	O	602	EPE	C10-S	3.71	1.82	1.77
2	O	602	EPE	C5-N4	2.07	1.52	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	O	602	EPE	O3S-S-O1S	-4.38	100.58	111.27
2	O	602	EPE	O2S-S-C10	4.06	111.81	106.92
2	O	602	EPE	C2-C3-N4	2.93	116.65	110.64
2	O	602	EPE	O2S-S-O1S	2.44	122.41	113.95

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	O	602	EPE	C8-C7-N4-C5
3	O	601	GOL	C1-C2-C3-O3
3	O	601	GOL	O2-C2-C3-O3
2	O	602	EPE	C8-C7-N4-C3
2	O	602	EPE	N4-C7-C8-O8

There are no ring outliers.

3 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	O	602	EPE	2	0
3	Z	603	GOL	2	0
3	O	601	GOL	4	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	O	497/501 (99%)	-0.94	0 100 100	12, 33, 66, 90	0
1	Z	498/501 (99%)	-0.97	0 100 100	11, 32, 66, 87	0
All	All	995/1002 (99%)	-0.96	0 100 100	11, 32, 66, 90	0

There are no RSRZ outliers to report.

### 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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	EPE	O	602	15/15	0.92	0.28	84,84,84,84	0
3	GOL	O	601	6/6	0.97	0.11	10,11,69,71	0
3	GOL	Z	603	6/6	0.98	0.10	10,10,10,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.