wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 06:38 AM EDT

PDB ID	:	1BU6
Title	:	CRYSTAL STRUCTURES OF ESCHERICHIA COLI GLYCEROL KINASE AND THE MUTANT A65T IN AN INACTIVE TETRAMER: CONFORMATIONAL CHANGES AND IMPLICATIONS FOR ALLOSTERIC REGULATION
Authors	:	Feese, M.D.; Faber, H.R.; Bystrom, C.E.; Pettigrew, D.W.; Remington, S.J.
Deposited on	:	1998-08-30
Resolution	:	2.37 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 (i)) were used in the production of this report:

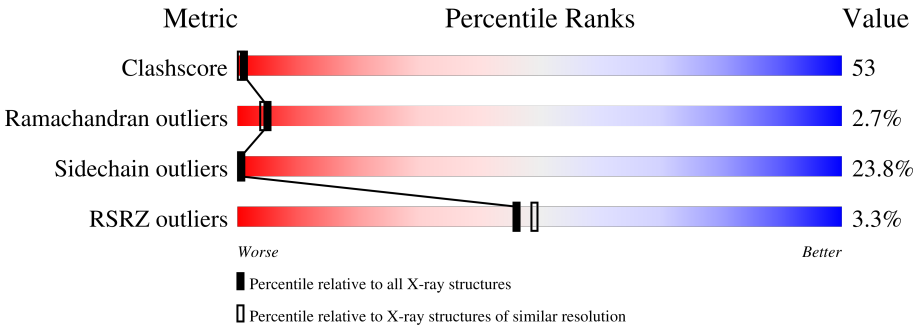
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.34
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 2.37 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	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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% 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	5% 26% 45% 22% 7%
1	X	501	3% 28% 45% 19% 7%
1	Y	501	2% 29% 44% 23%
1	Z	501	3% 26% 48% 21% 5%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	O	502	-	-	X	-
3	GOL	Z	504	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 15815 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.

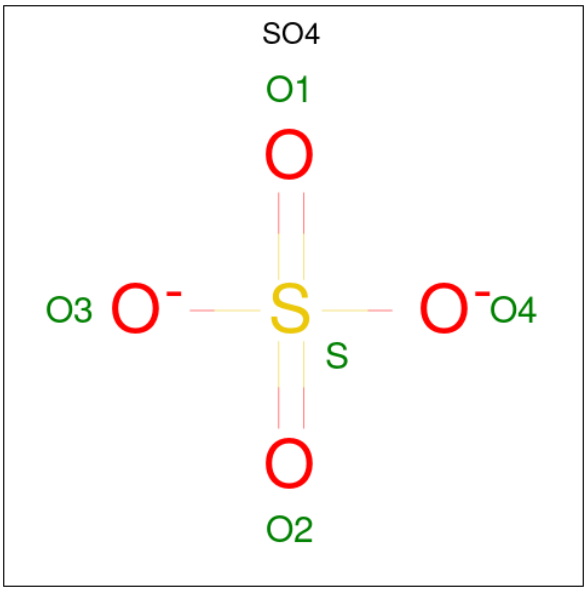
- Molecule 1 is a protein called PROTEIN (GLYCEROL KINASE).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	497	Total	C	N	O	S	0	0	0
			3891	2452	680	740	19			
1	Y	499	Total	C	N	O	S	0	0	0
			3921	2471	686	745	19			
1	Z	498	Total	C	N	O	S	0	0	0
			3913	2465	686	743	19			
1	X	498	Total	C	N	O	S	0	0	0
			3896	2456	681	740	19			

There are 4 discrepancies between the modelled and reference sequences:

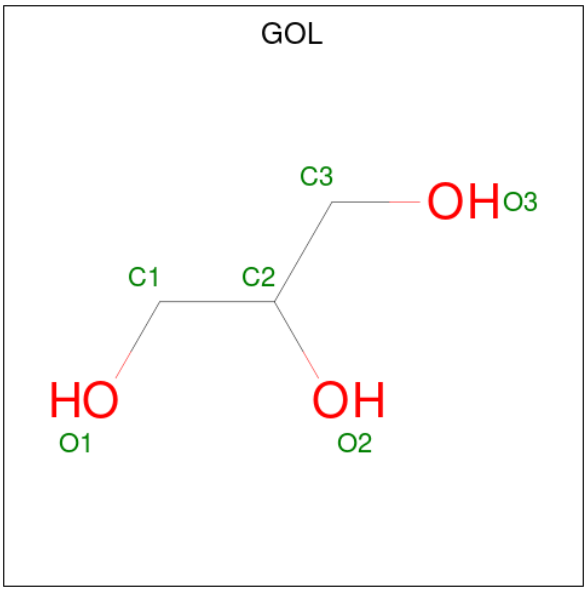
Chain	Residue	Modelled	Actual	Comment	Reference
O	65	THR	ALA	engineered mutation	UNP P0A6F3
Y	65	THR	ALA	engineered mutation	UNP P0A6F3
Z	65	THR	ALA	engineered mutation	UNP P0A6F3
X	65	THR	ALA	engineered mutation	UNP P0A6F3

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	O	1	Total O S 5 4 1	0	0
2	O	1	Total O S 5 4 1	0	0
2	Y	1	Total O S 5 4 1	0	0
2	Z	1	Total O S 5 4 1	0	0
2	Z	1	Total O S 5 4 1	0	0
2	X	1	Total O S 5 4 1	0	0

- 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 6 3 3	0	0
3	Y	1	Total C O 6 3 3	0	0
3	Z	1	Total C O 6 3 3	0	0
3	X	1	Total C O 6 3 3	0	0

- Molecule 4 is water.

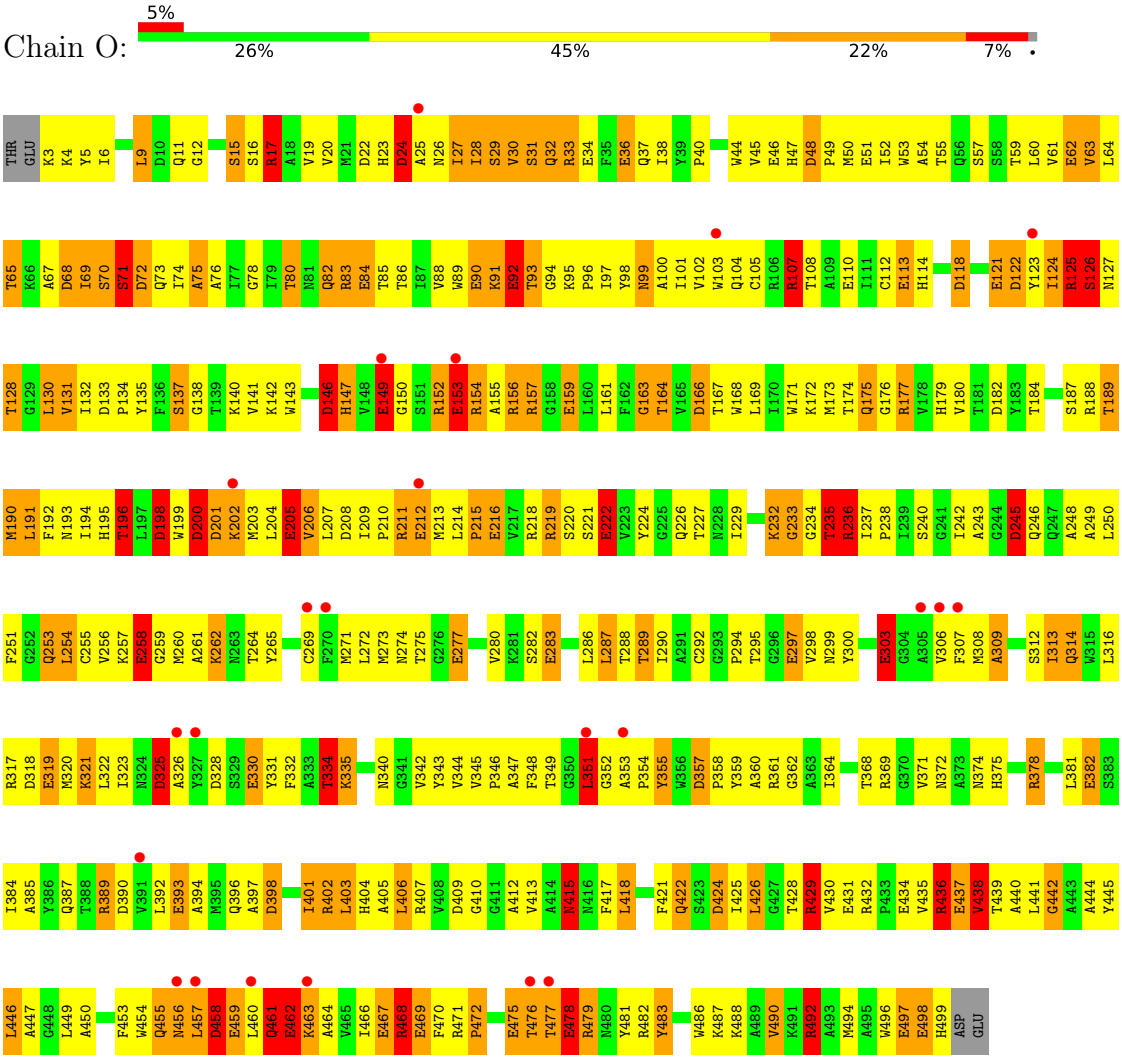
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	O	29	Total O 29 29	0	0
4	Y	38	Total O 38 38	0	0
4	Z	31	Total O 31 31	0	0
4	X	42	Total O 42 42	0	0

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.

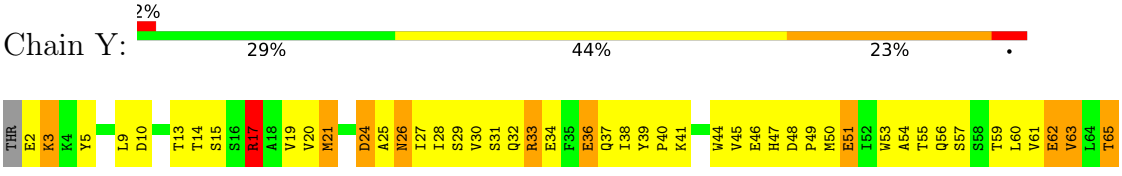
• Molecule 1: PROTEIN (GLYCEROL KINASE)

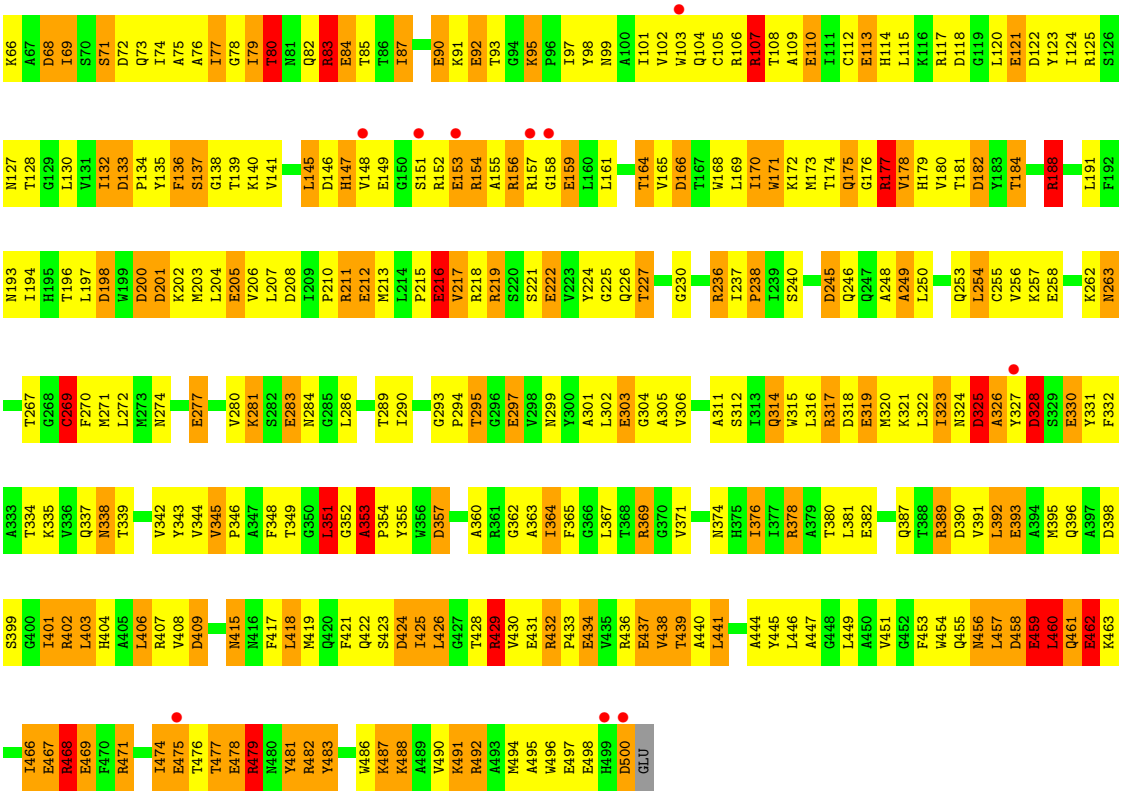
Chain O:



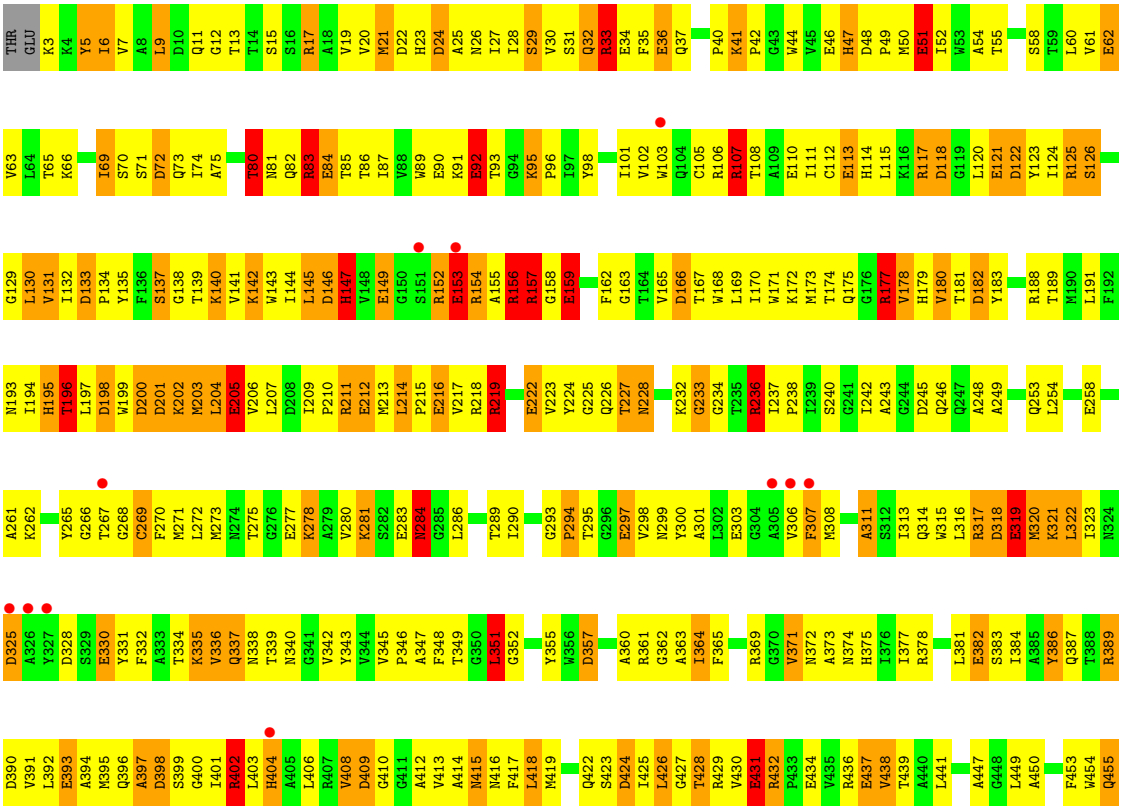
• Molecule 1: PROTEIN (GLYCEROL KINASE)

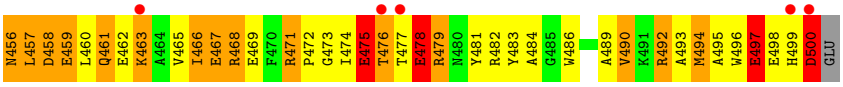
Chain Y:





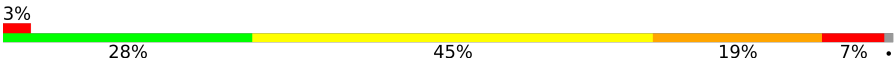
• Molecule 1: PROTEIN (GLYCEROL KINASE)

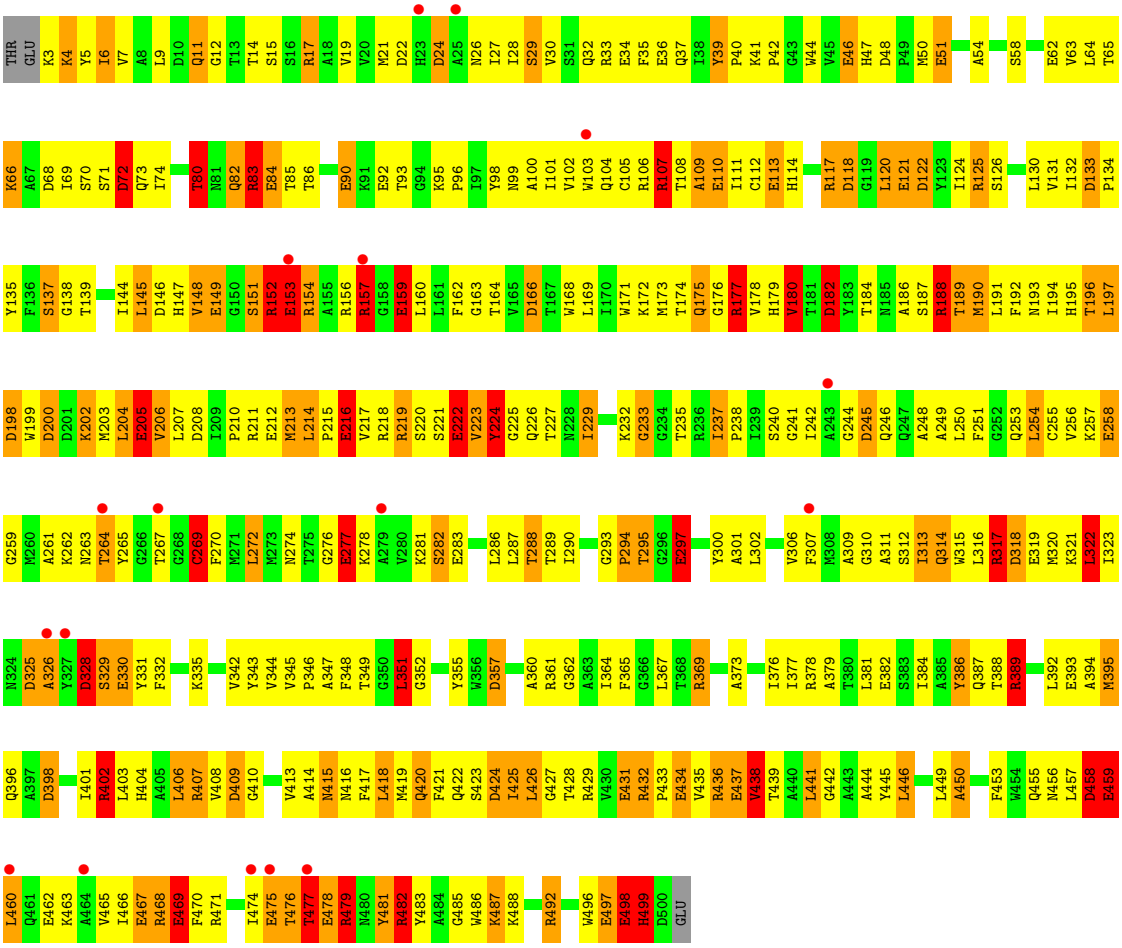




• Molecule 1: PROTEIN (GLYCEROL KINASE)

Chain X:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	91.90Å 119.00Å 109.30Å 90.00° 103.40° 90.00°	Depositor
Resolution (Å)	20.00 – 2.37 37.01 – 2.37	Depositor EDS
% Data completeness (in resolution range)	74.0 (20.00-2.37) 72.3 (37.01-2.37)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.49 (at 2.37Å)	Xtriage
Refinement program	TNT 5F-6	Depositor
R, R_{free}	0.167 , (Not available) 0.157 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	31.5	Xtriage
Anisotropy	0.469	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 128.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	15815	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 32.80 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.8628e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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, SO4

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	1.34	35/3970 (0.9%)	1.71	83/5387 (1.5%)
1	X	1.37	39/3975 (1.0%)	1.76	96/5393 (1.8%)
1	Y	1.34	40/4001 (1.0%)	1.75	102/5427 (1.9%)
1	Z	1.32	36/3992 (0.9%)	1.71	69/5413 (1.3%)
All	All	1.34	150/15938 (0.9%)	1.74	350/21620 (1.6%)

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	X	2	0
1	Z	1	0
All	All	3	0

The worst 5 of 150 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	X	475	GLU	CD-OE2	16.71	1.44	1.25
1	Y	475	GLU	CD-OE2	12.24	1.39	1.25
1	Z	36	GLU	CD-OE2	10.65	1.37	1.25
1	Y	84	GLU	CD-OE1	10.23	1.36	1.25
1	Y	434	GLU	CD-OE2	10.15	1.36	1.25

The worst 5 of 350 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	Z	107	ARG	NE-CZ-NH1	13.64	127.12	120.30
1	Z	479	ARG	NE-CZ-NH1	12.15	126.38	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	O	200	ASP	CB-CG-OD2	-11.71	107.76	118.30
1	X	479	ARG	NE-CZ-NH2	-10.50	115.05	120.30
1	Y	407	ARG	NE-CZ-NH2	-10.35	115.13	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	Z	149	GLU	CA
1	X	147	HIS	CA
1	X	404	HIS	CA

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	3891	0	3804	450	0
1	X	3896	0	3811	414	0
1	Y	3921	0	3839	406	0
1	Z	3913	0	3841	448	0
2	O	10	0	0	2	0
2	X	5	0	0	0	0
2	Y	5	0	0	0	0
2	Z	10	0	0	2	0
3	O	6	0	8	1	0
3	X	6	0	8	1	0
3	Y	6	0	8	3	0
3	Z	6	0	8	7	0
4	O	29	0	0	3	0
4	X	42	0	0	7	0
4	Y	38	0	0	4	0
4	Z	31	0	0	2	0
All	All	15815	0	15327	1658	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 53.

The worst 5 of 1658 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:Y:33:ARG:HH22	1:Z:58:SER:HB2	1.15	1.11
1:O:287:LEU:HD12	1:O:303:GLU:HG2	1.31	1.11
1:Y:456:ASN:ND2	1:Y:458:ASP:H	1.51	1.08
1:Z:468:ARG:HG3	1:Z:468:ARG:HH11	1.19	1.07
1:Z:154:ARG:HB3	1:Z:159:GLU:HG3	1.32	1.05

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%)	420 (85%)	58 (12%)	17 (3%)	3	3
1	X	496/501 (99%)	436 (88%)	42 (8%)	18 (4%)	3	2
1	Y	497/501 (99%)	442 (89%)	46 (9%)	9 (2%)	8	9
1	Z	496/501 (99%)	448 (90%)	39 (8%)	9 (2%)	8	9
All	All	1984/2004 (99%)	1746 (88%)	185 (9%)	53 (3%)	5	4

5 of 53 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	O	149	GLU
1	O	232	LYS
1	O	233	GLY
1	O	461	GLN
1	Y	175	GLN

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	404/413 (98%)	299 (74%)	105 (26%)	0	0
1	X	404/413 (98%)	309 (76%)	95 (24%)	1	0
1	Y	408/413 (99%)	322 (79%)	86 (21%)	1	1
1	Z	408/413 (99%)	308 (76%)	100 (24%)	0	0
All	All	1624/1652 (98%)	1238 (76%)	386 (24%)	0	0

5 of 386 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	Z	178	VAL
1	Z	478	GLU
1	Z	204	LEU
1	Z	364	ILE
1	X	72	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 66 such sidechains are listed below:

Mol	Chain	Res	Type
1	X	195	HIS
1	X	299	ASN
1	X	499	HIS
1	Y	299	ASN
1	Y	195	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

10 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	SO4	Z	502	-	4,4,4	1.06	0	6,6,6	0.25	0
3	GOL	Z	504	-	5,5,5	0.20	0	5,5,5	0.79	0
2	SO4	O	502	-	4,4,4	1.33	1 (25%)	6,6,6	0.36	0
2	SO4	O	503	-	4,4,4	1.76	2 (50%)	6,6,6	0.46	0
3	GOL	X	503	-	5,5,5	0.31	0	5,5,5	0.73	0
2	SO4	X	502	-	4,4,4	1.78	1 (25%)	6,6,6	0.23	0
3	GOL	Y	503	-	5,5,5	0.35	0	5,5,5	0.30	0
2	SO4	Z	503	-	4,4,4	1.74	1 (25%)	6,6,6	0.35	0
2	SO4	Y	502	-	4,4,4	1.56	1 (25%)	6,6,6	0.52	0
3	GOL	O	504	-	5,5,5	0.88	0	5,5,5	0.74	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
3	GOL	Z	504	-	-	1/4/4/4	-
3	GOL	Y	503	-	-	2/4/4/4	-
3	GOL	X	503	-	-	0/4/4/4	-
3	GOL	O	504	-	-	0/4/4/4	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	X	502	SO4	O1-S	3.00	1.62	1.46
2	O	503	SO4	O3-S	2.80	1.71	1.47
2	Y	502	SO4	O3-S	2.71	1.70	1.47
2	O	502	SO4	O1-S	-2.52	1.32	1.46
2	Z	503	SO4	O2-S	2.32	1.58	1.46

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	Y	503	GOL	O1-C1-C2-C3
3	Z	504	GOL	O1-C1-C2-C3
3	Y	503	GOL	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	Z	502	SO4	1	0
3	Z	504	GOL	7	0
2	O	502	SO4	2	0
3	X	503	GOL	1	0
3	Y	503	GOL	3	0
2	Z	503	SO4	1	0
3	O	504	GOL	1	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

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, 95^{th} 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	O	497/501 (99%)	0.11	23 (4%)	32	35	10, 36, 67, 75	0
1	X	498/501 (99%)	-0.01	17 (3%)	45	48	13, 34, 67, 75	0
1	Y	499/501 (99%)	-0.07	10 (2%)	65	66	13, 34, 65, 75	0
1	Z	498/501 (99%)	-0.09	16 (3%)	47	50	12, 33, 62, 75	0
All	All	1992/2004 (99%)	-0.02	66 (3%)	46	49	10, 34, 65, 75	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Y	158	GLY	5.6
1	Y	499	HIS	4.9
1	Y	327	TYR	4.9
1	O	460	LEU	4.8
1	X	460	LEU	4.4

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 monosaccharides in this entry.

6.4 Ligands

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^{th} 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	SO4	O	503	5/5	0.93	0.24	37,48,75,75	0
3	GOL	O	504	6/6	0.94	0.31	12,23,29,40	0
2	SO4	Y	502	5/5	0.95	0.27	36,67,75,75	0
2	SO4	Z	503	5/5	0.96	0.26	51,60,75,75	0
2	SO4	X	502	5/5	0.97	0.33	30,75,75,75	0
3	GOL	X	503	6/6	0.97	0.30	5,20,29,35	0
3	GOL	Z	504	6/6	0.98	0.26	5,16,20,54	0
3	GOL	Y	503	6/6	0.99	0.27	11,18,30,63	0
2	SO4	Z	502	5/5	0.99	0.10	27,48,75,75	0
2	SO4	O	502	5/5	0.99	0.11	11,36,68,75	0

6.5 Other polymers (i)

There are no such residues in this entry.