



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2017 – 08:58 PM EDT

PDB ID : 2PKR  
Title : Crystal structure of (A+CTE)4 chimeric form of photosynthetic glyceraldehyd  
e-3-phosphate dehydrogenase, complexed with NADP  
Authors : Fermani, S.; Falini, G.; Ripamonti, A.  
Deposited on : unknown  
Resolution : 2.40 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

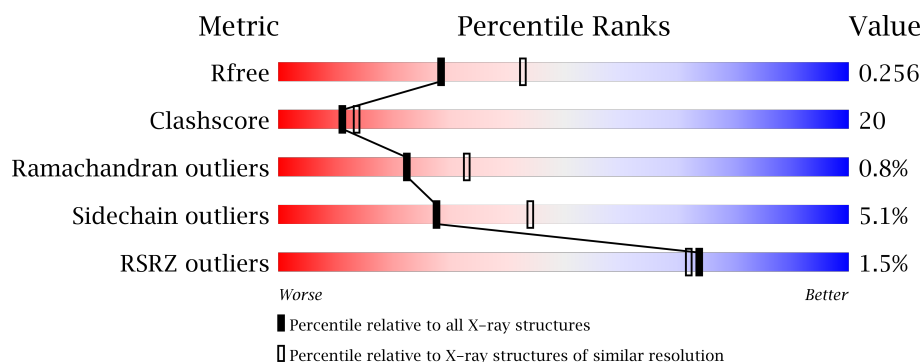
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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. 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	365	<div> <div>3%</div> <div>67% 19% 5% 8%</div> </div>
1	B	365	<div> <div>63% 24% 8%</div> </div>
1	C	365	<div> <div>68% 22% 8%</div> </div>
1	D	365	<div> <div>3%</div> <div>69% 20% 8%</div> </div>
1	H	365	<div> <div>2%</div> <div>68% 19% 8%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	I	365	
1	L	365	
1	M	365	
1	O	365	
1	P	365	
1	Q	365	
1	R	365	

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	SO4	B	364	-	-	X	-
2	SO4	B	377	-	-	-	X
2	SO4	D	364	-	-	X	-
2	SO4	H	364	-	-	X	X
2	SO4	I	366	-	-	X	X
2	SO4	I	379	-	-	-	X
2	SO4	M	370	-	-	X	X
2	SO4	M	380	-	-	-	X
2	SO4	O	364	-	-	X	-
2	SO4	O	371	-	-	X	X
2	SO4	P	364	-	-	X	-
2	SO4	Q	373	-	-	-	X
2	SO4	Q	374	-	-	-	X
2	SO4	R	376	-	-	X	-
3	NDP	A	376	X	-	-	-
3	NDP	B	378	X	-	-	-
3	NDP	C	365	X	-	-	-
3	NDP	D	365	X	-	-	-
3	NDP	H	379	X	-	-	-
3	NDP	I	380	X	-	-	-
3	NDP	L	369	X	-	-	-
3	NDP	M	381	X	-	-	-
3	NDP	O	372	X	-	-	-
3	NDP	P	373	X	-	-	-
3	NDP	Q	375	X	-	-	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NDP	R	377	X	-	-	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 31248 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 Glyceraldehyde-3-phosphate dehydrogenase Aor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	R	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	P	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	Q	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	A	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	B	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	C	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	D	336	Total	C	N	O	S	0	0	0
			2538	1597	444	486	11			
1	H	336	Total	C	N	O	S	0	0	0
			2538	1597	444	486	11			
1	I	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			
1	L	336	Total	C	N	O	S	0	0	0
			2538	1597	444	486	11			
1	M	337	Total	C	N	O	S	0	0	0
			2542	1599	445	487	11			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



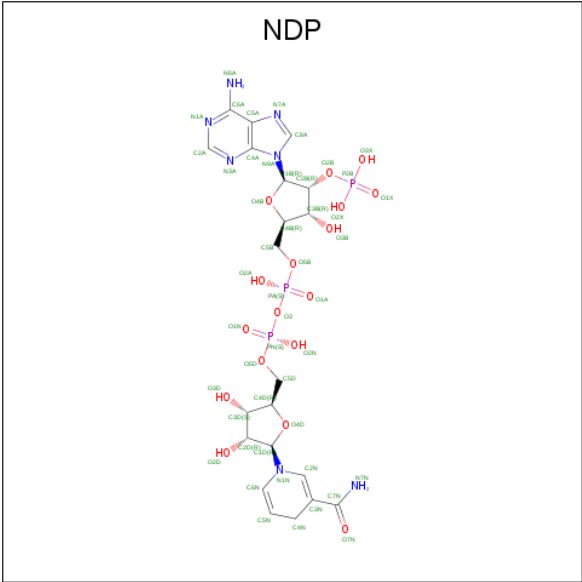
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	O	1	Total	O	S	0	0
			5	4	1		
2	O	1	Total	O	S	0	0
			5	4	1		
2	O	1	Total	O	S	0	0
			5	4	1		
2	R	1	Total	O	S	0	0
			5	4	1		
2	R	1	Total	O	S	0	0
			5	4	1		
2	R	1	Total	O	S	0	0
			5	4	1		
2	P	1	Total	O	S	0	0
			5	4	1		
2	P	1	Total	O	S	0	0
			5	4	1		
2	P	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	Q	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	H	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	I	1	Total	O	S	0	0
			5	4	1		
2	L	1	Total	O	S	0	0
			5	4	1		
2	L	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		
2	M	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is NADPH DIHYDRO-NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NDP) (formula: C<sub>21</sub>H<sub>30</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	O	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	R	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	P	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	Q	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	C	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	D	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	H	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	I	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	L	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
3	M	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 4 is water.

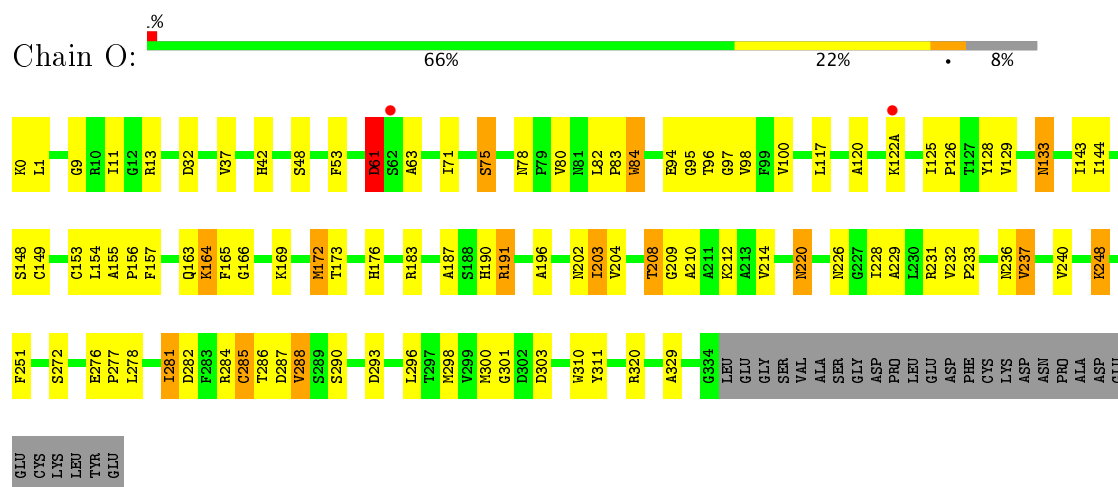


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	O	1	Total 1	O 1	0	0
4	H	2	Total 2	O 2	0	0
4	I	1	Total 1	O 1	0	0
4	L	4	Total 4	O 4	0	0
4	M	2	Total 2	O 2	0	0

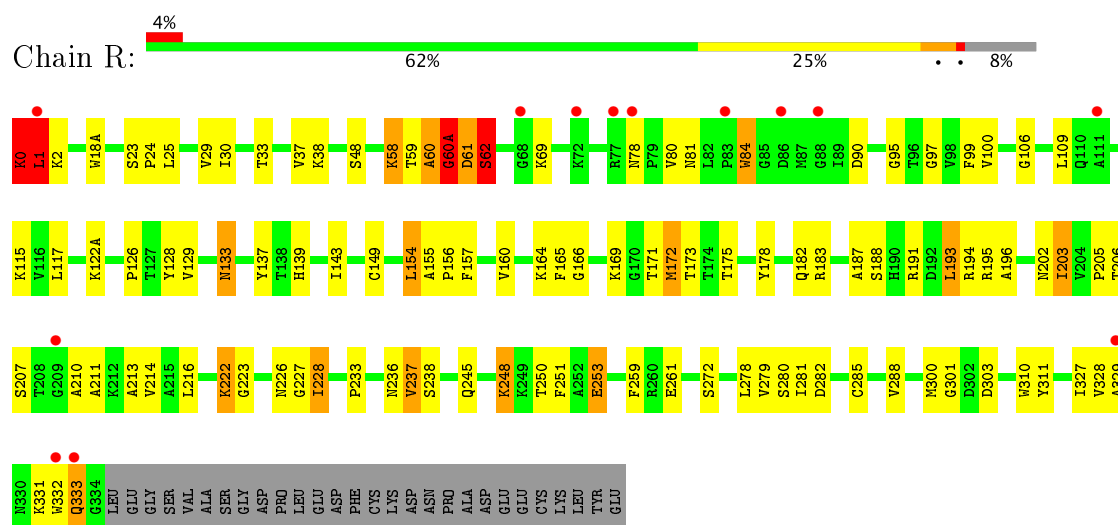
### 3 Residue-property plots

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 ( $RSRZ > 2$ ). 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: Glyceraldehyde-3-phosphate dehydrogenase Aor

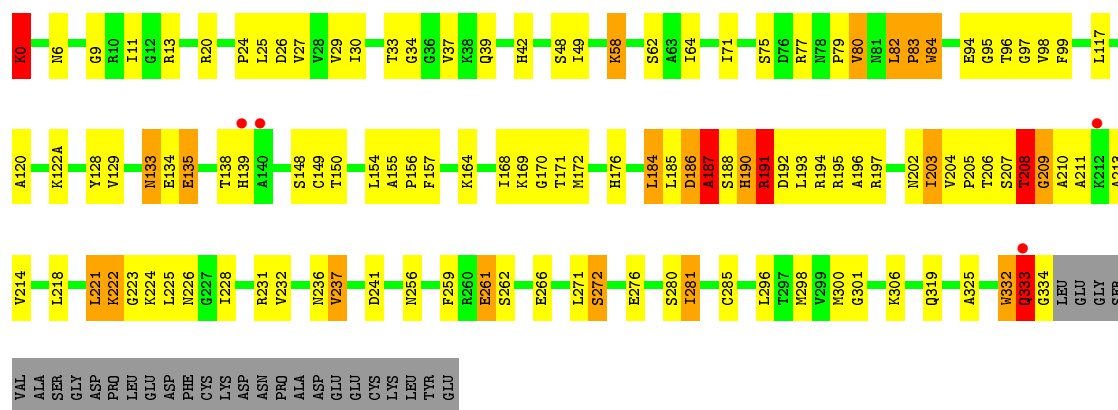


- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor

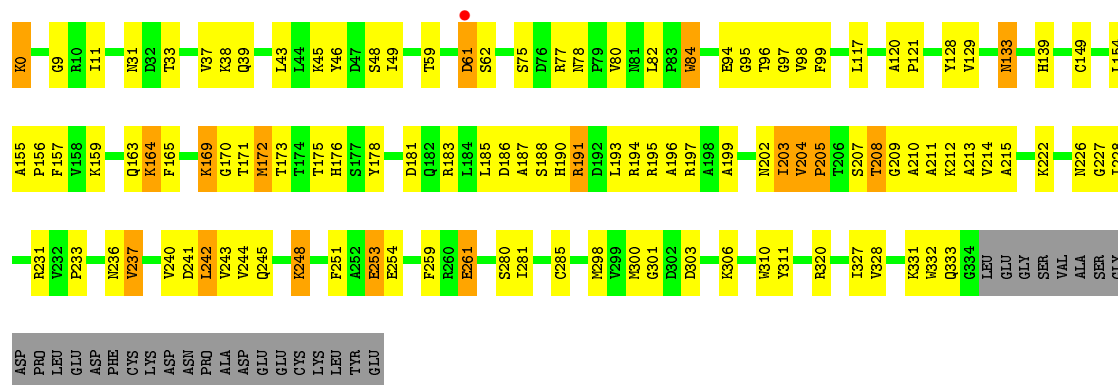


- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor

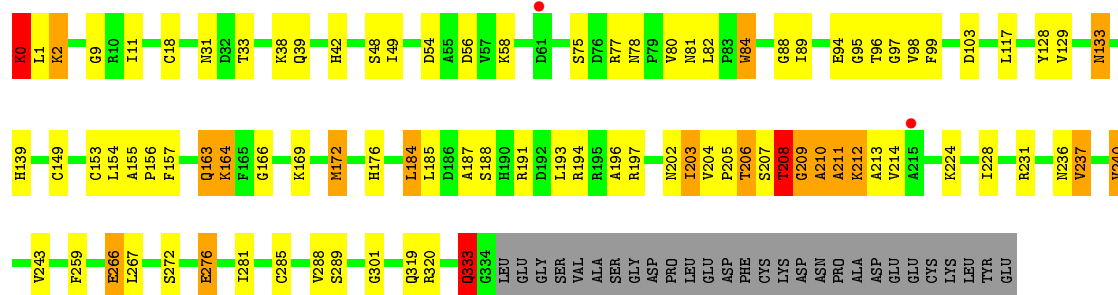




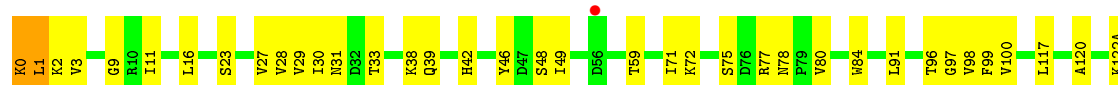
- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor



- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor



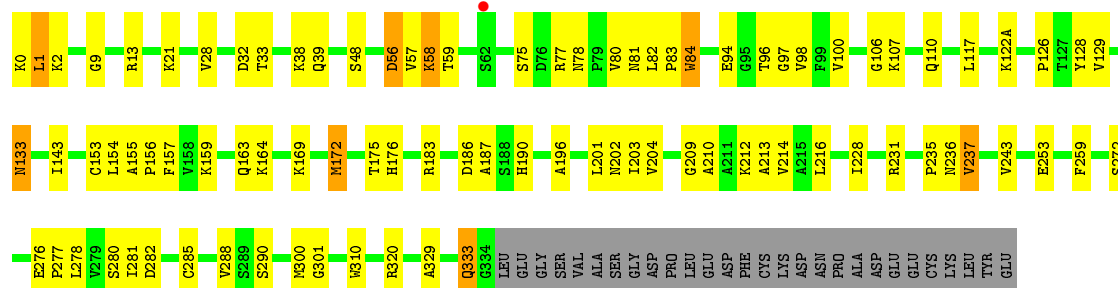
- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor





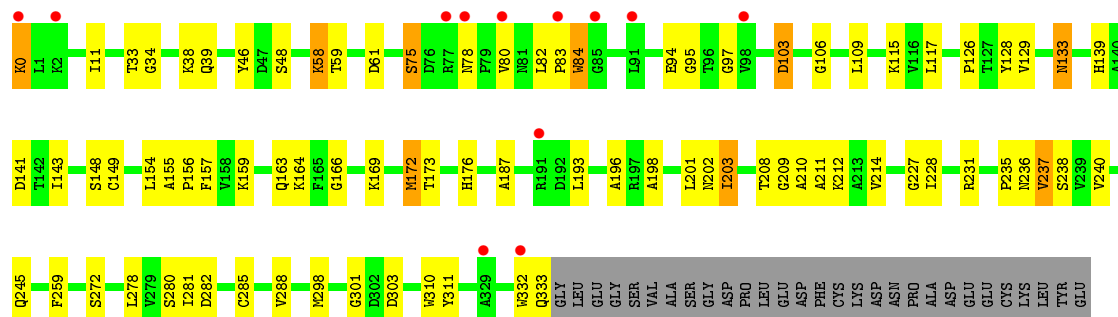
• Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor

Chain C: 68% 22% 8%



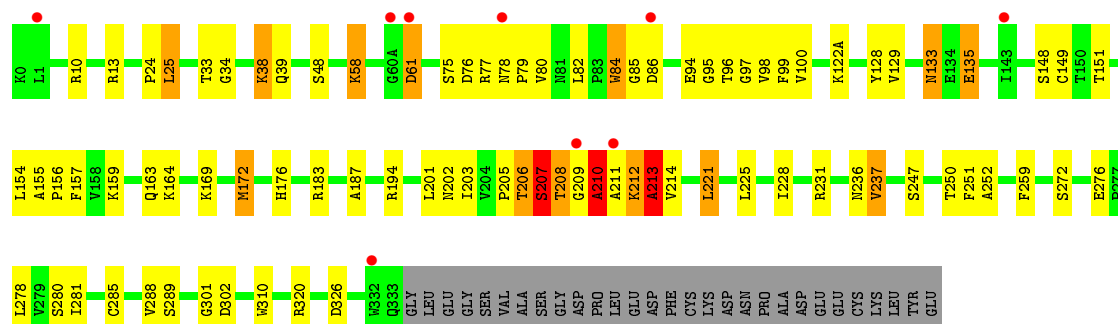
• Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor

Chain D: 3% 69% 20% 8%

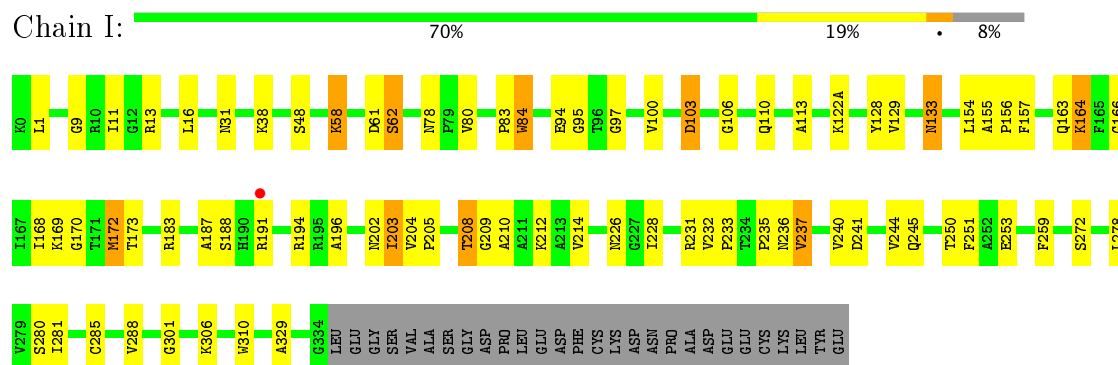


• Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor

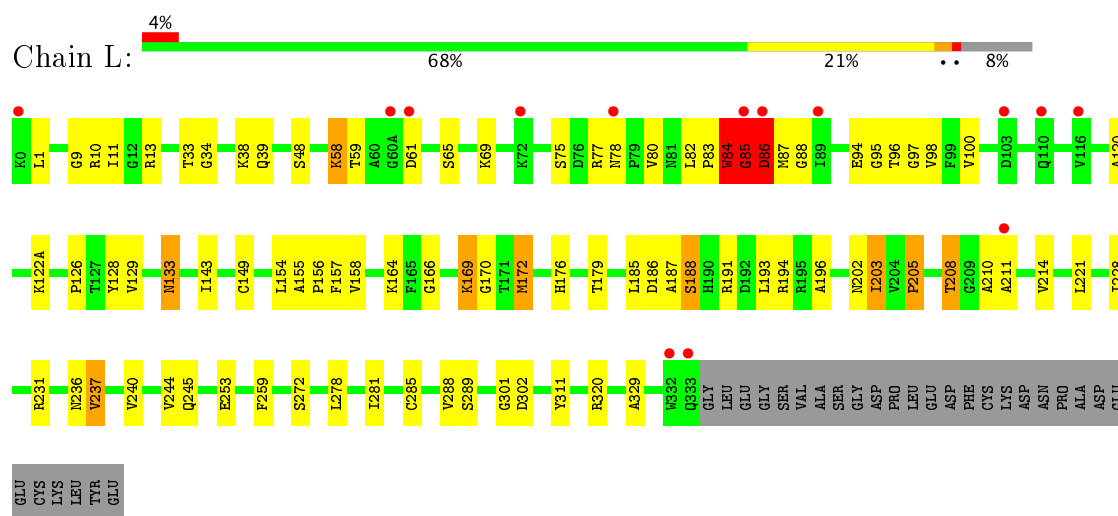
Chain H: 2% 68% 19% 8%



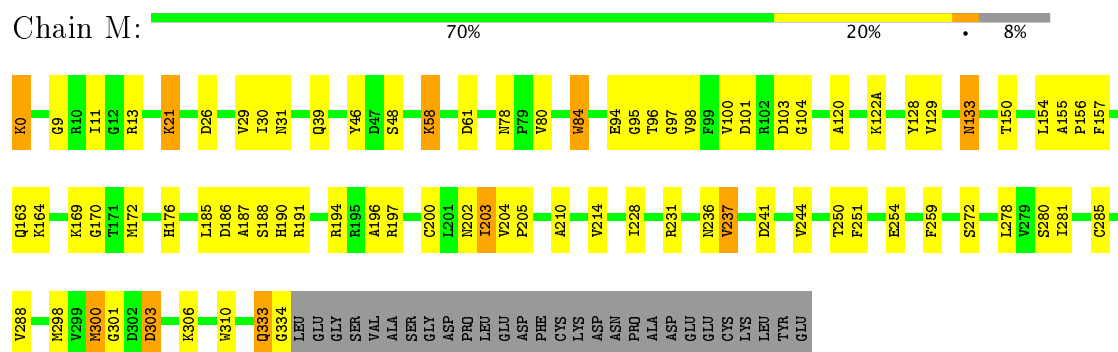
• Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor



- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor



- Molecule 1: Glyceraldehyde-3-phosphate dehydrogenase Aor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	105.38Å 217.71Å 114.79Å 90.00° 90.04° 90.00°	Depositor
Resolution (Å)	63.22 – 2.40 73.14 – 2.40	Depositor EDS
% Data completeness (in resolution range)	76.5 (63.22-2.40) 76.5 (73.14-2.40)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.70 (at 2.40Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.249 , 0.258 0.248 , 0.256	Depositor DCC
$R_{free}$ test set	7785 reflections (5.06%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.0	Xtriage
Anisotropy	0.168	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , -6.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.478 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	31248	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	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 20.39 % 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 8.7836e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: NDP, 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	A	0.65	3/2583 (0.1%)	0.96	9/3507 (0.3%)
1	B	0.74	4/2583 (0.2%)	0.95	11/3507 (0.3%)
1	C	0.60	0/2583	0.86	4/3507 (0.1%)
1	D	0.59	0/2579	0.82	3/3502 (0.1%)
1	H	0.66	3/2579 (0.1%)	0.97	12/3502 (0.3%)
1	I	0.57	0/2583	0.81	1/3507 (0.0%)
1	L	0.61	1/2579 (0.0%)	0.92	13/3502 (0.4%)
1	M	0.57	1/2583 (0.0%)	0.82	3/3507 (0.1%)
1	O	0.61	0/2583	0.88	4/3507 (0.1%)
1	P	0.66	1/2583 (0.0%)	1.04	16/3507 (0.5%)
1	Q	0.59	1/2583 (0.0%)	0.90	8/3507 (0.2%)
1	R	0.64	2/2583 (0.1%)	0.99	12/3507 (0.3%)
All	All	0.63	16/30984 (0.1%)	0.91	96/42069 (0.2%)

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	1	7
1	B	0	3
1	C	0	1
1	H	1	5
1	I	0	1
1	L	0	2
1	M	0	2
1	O	0	2
1	P	1	7
1	Q	0	6
1	R	0	8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
All	All	3	44

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	210	ALA	C-N	14.89	1.68	1.34
1	B	333	GLN	C-N	-11.20	1.12	1.33
1	A	0	LYS	C-O	10.76	1.43	1.23
1	B	205	PRO	C-N	9.79	1.56	1.34
1	L	205	PRO	C-N	-7.87	1.16	1.34

The worst 5 of 96 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	62	SER	O-C-N	-19.79	91.04	122.70
1	A	0	LYS	CA-C-O	-14.38	89.89	120.10
1	C	333	GLN	C-N-CA	13.54	150.73	122.30
1	H	206	THR	C-N-CA	12.83	153.76	121.70
1	A	206	THR	C-N-CA	11.39	150.18	121.70

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	P	208	THR	CA
1	A	208	THR	CA
1	H	212	LYS	CA

5 of 44 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	O	285	CYS	Mainchain
1	O	61	ASP	Mainchain
1	R	0	LYS	Mainchain,Peptide
1	R	60(A)	GLY	Mainchain,Peptide
1	R	62	SER	Mainchain

## 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	2542	0	2573	123	0
1	B	2542	0	2573	132	0
1	C	2542	0	2576	117	0
1	D	2538	0	2574	86	0
1	H	2538	0	2571	96	0
1	I	2542	0	2577	112	0
1	L	2538	0	2572	82	1
1	M	2542	0	2577	94	0
1	O	2542	0	2577	154	0
1	P	2542	0	2575	151	0
1	Q	2542	0	2577	167	1
1	R	2542	0	2575	154	0
2	A	15	0	0	0	0
2	B	15	0	0	2	0
2	C	10	0	0	0	0
2	D	10	0	0	2	0
2	H	15	0	0	4	0
2	I	15	0	0	6	0
2	L	10	0	0	1	0
2	M	15	0	0	3	0
2	O	15	0	0	4	0
2	P	15	0	0	6	0
2	Q	20	0	0	1	0
2	R	15	0	0	3	0
3	A	48	0	20	19	0
3	B	48	0	21	13	0
3	C	48	0	20	16	0
3	D	48	0	19	6	0
3	H	48	0	20	14	0
3	I	48	0	20	10	0
3	L	48	0	20	14	0
3	M	48	0	20	10	0
3	O	48	0	21	15	0
3	P	48	0	21	11	0
3	Q	48	0	21	16	0
3	R	48	0	20	5	0
4	H	2	0	0	0	0
4	I	1	0	0	0	0
4	L	4	0	0	0	0
4	M	2	0	0	1	0
4	O	1	0	0	0	0
All	All	31248	0	31140	1242	1

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 20.

The worst 5 of 1242 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)
3:D:365:NDP:C3D	3:D:365:NDP:C2D	1.76	1.64
3:M:381:NDP:C1D	3:M:381:NDP:C2D	1.76	1.62
3:I:380:NDP:C1D	3:I:380:NDP:C2D	1.77	1.61
3:C:365:NDP:C2D	3:C:365:NDP:C1D	1.79	1.59
3:B:378:NDP:C1D	3:B:378:NDP:C2D	1.78	1.55

All (1) 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)
1:Q:62:SER:OG	1:L:253:GLU:OE2[2_555]	2.19	0.01

## 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	335/365 (92%)	308 (92%)	23 (7%)	4 (1%)	15	21
1	B	335/365 (92%)	313 (93%)	17 (5%)	5 (2%)	12	16
1	C	335/365 (92%)	315 (94%)	19 (6%)	1 (0%)	44	60
1	D	334/365 (92%)	318 (95%)	13 (4%)	3 (1%)	20	29
1	H	334/365 (92%)	314 (94%)	17 (5%)	3 (1%)	20	29
1	I	335/365 (92%)	316 (94%)	16 (5%)	3 (1%)	20	29
1	L	334/365 (92%)	315 (94%)	15 (4%)	4 (1%)	15	21
1	M	335/365 (92%)	312 (93%)	22 (7%)	1 (0%)	44	60
1	O	335/365 (92%)	317 (95%)	16 (5%)	2 (1%)	28	41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	P	335/365 (92%)	312 (93%)	21 (6%)	2 (1%)	28	41
1	Q	335/365 (92%)	309 (92%)	25 (8%)	1 (0%)	44	60
1	R	335/365 (92%)	312 (93%)	19 (6%)	4 (1%)	15	21
All	All	4017/4380 (92%)	3761 (94%)	223 (6%)	33 (1%)	22	33

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	R	1	LEU
1	R	60(A)	GLY
1	P	209	GLY
1	A	209	GLY
1	A	211	ALA

### 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	279/303 (92%)	262 (94%)	17 (6%)	22	34
1	B	279/303 (92%)	269 (96%)	10 (4%)	40	60
1	C	279/303 (92%)	267 (96%)	12 (4%)	33	52
1	D	279/303 (92%)	269 (96%)	10 (4%)	40	60
1	H	279/303 (92%)	264 (95%)	15 (5%)	26	41
1	I	279/303 (92%)	267 (96%)	12 (4%)	33	52
1	L	279/303 (92%)	269 (96%)	10 (4%)	40	60
1	M	279/303 (92%)	266 (95%)	13 (5%)	30	48
1	O	279/303 (92%)	262 (94%)	17 (6%)	22	34
1	P	279/303 (92%)	256 (92%)	23 (8%)	13	20
1	Q	279/303 (92%)	265 (95%)	14 (5%)	28	45
1	R	279/303 (92%)	262 (94%)	17 (6%)	22	34
All	All	3348/3636 (92%)	3178 (95%)	170 (5%)	28	44

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	163	GLN
1	B	188	SER
1	M	0	LYS
1	A	184	LEU
1	A	285	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 87 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	152	ASN
1	C	163	GLN
1	M	81	ASN
1	B	202	ASN
1	C	81	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

46 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	A	363	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	A	364	-	4,4,4	0.40	0	6,6,6	0.09	0
2	SO4	A	375	-	4,4,4	0.32	0	6,6,6	0.09	0
3	NDP	A	376	-	43,52,52	4.89	26 (60%)	49,80,80	5.81	32 (65%)
2	SO4	B	363	-	4,4,4	0.28	0	6,6,6	0.08	0
2	SO4	B	364	-	4,4,4	0.38	0	6,6,6	0.09	0
2	SO4	B	377	-	4,4,4	0.27	0	6,6,6	0.08	0
3	NDP	B	378	-	43,52,52	4.15	24 (55%)	49,80,80	6.25	30 (61%)
2	SO4	C	363	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	C	364	-	4,4,4	0.40	0	6,6,6	0.09	0
3	NDP	C	365	-	43,52,52	4.13	23 (53%)	49,80,80	6.08	31 (63%)
2	SO4	D	363	-	4,4,4	0.32	0	6,6,6	0.09	0
2	SO4	D	364	-	4,4,4	0.41	0	6,6,6	0.09	0
3	NDP	D	365	-	43,52,52	4.23	24 (55%)	49,80,80	6.21	29 (59%)
2	SO4	H	363	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	H	364	-	4,4,4	0.41	0	6,6,6	0.09	0
2	SO4	H	378	-	4,4,4	0.31	0	6,6,6	0.10	0
3	NDP	H	379	-	43,52,52	4.17	23 (53%)	49,80,80	6.14	29 (59%)
2	SO4	I	365	-	4,4,4	0.30	0	6,6,6	0.08	0
2	SO4	I	366	-	4,4,4	0.39	0	6,6,6	0.09	0
2	SO4	I	379	-	4,4,4	0.40	0	6,6,6	0.09	0
3	NDP	I	380	-	43,52,52	4.11	24 (55%)	49,80,80	6.26	28 (57%)
2	SO4	L	367	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	L	368	-	4,4,4	0.39	0	6,6,6	0.10	0
3	NDP	L	369	-	43,52,52	4.31	22 (51%)	49,80,80	5.79	32 (65%)
2	SO4	M	369	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	M	370	-	4,4,4	0.37	0	6,6,6	0.08	0
2	SO4	M	380	-	4,4,4	0.31	0	6,6,6	0.09	0
3	NDP	M	381	-	43,52,52	4.11	23 (53%)	49,80,80	6.15	29 (59%)
2	SO4	O	363	-	4,4,4	0.30	0	6,6,6	0.09	0
2	SO4	O	364	-	4,4,4	0.39	0	6,6,6	0.09	0
2	SO4	O	371	-	4,4,4	0.39	0	6,6,6	0.09	0
3	NDP	O	372	-	43,52,52	4.55	23 (53%)	49,80,80	5.84	27 (55%)
2	SO4	P	363	-	4,4,4	0.31	0	6,6,6	0.09	0
2	SO4	P	364	-	4,4,4	0.38	0	6,6,6	0.09	0
2	SO4	P	372	-	4,4,4	0.38	0	6,6,6	0.09	0
3	NDP	P	373	-	43,52,52	4.23	26 (60%)	49,80,80	5.72	30 (61%)
2	SO4	Q	363	-	4,4,4	0.29	0	6,6,6	0.08	0
2	SO4	Q	364	-	4,4,4	0.40	0	6,6,6	0.09	0
2	SO4	Q	373	-	4,4,4	0.29	0	6,6,6	0.08	0
2	SO4	Q	374	-	4,4,4	0.40	0	6,6,6	0.09	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NDP	Q	375	-	43,52,52	4.09	24 (55%)	49,80,80	6.10	32 (65%)
2	SO4	R	363	-	4,4,4	0.32	0	6,6,6	0.09	0
2	SO4	R	364	-	4,4,4	0.39	0	6,6,6	0.09	0
2	SO4	R	376	-	4,4,4	0.39	0	6,6,6	0.09	0
3	NDP	R	377	-	43,52,52	4.30	25 (58%)	49,80,80	6.07	29 (59%)

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	SO4	A	363	-	-	0/0/0/0	0/0/0/0
2	SO4	A	364	-	-	0/0/0/0	0/0/0/0
2	SO4	A	375	-	-	0/0/0/0	0/0/0/0
3	NDP	A	376	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	B	363	-	-	0/0/0/0	0/0/0/0
2	SO4	B	364	-	-	0/0/0/0	0/0/0/0
2	SO4	B	377	-	-	0/0/0/0	0/0/0/0
3	NDP	B	378	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	C	363	-	-	0/0/0/0	0/0/0/0
2	SO4	C	364	-	-	0/0/0/0	0/0/0/0
3	NDP	C	365	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	D	363	-	-	0/0/0/0	0/0/0/0
2	SO4	D	364	-	-	0/0/0/0	0/0/0/0
3	NDP	D	365	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	H	363	-	-	0/0/0/0	0/0/0/0
2	SO4	H	364	-	-	0/0/0/0	0/0/0/0
2	SO4	H	378	-	-	0/0/0/0	0/0/0/0
3	NDP	H	379	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	I	365	-	-	0/0/0/0	0/0/0/0
2	SO4	I	366	-	-	0/0/0/0	0/0/0/0
2	SO4	I	379	-	-	0/0/0/0	0/0/0/0
3	NDP	I	380	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	L	367	-	-	0/0/0/0	0/0/0/0
2	SO4	L	368	-	-	0/0/0/0	0/0/0/0
3	NDP	L	369	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	M	369	-	-	0/0/0/0	0/0/0/0
2	SO4	M	370	-	-	0/0/0/0	0/0/0/0
2	SO4	M	380	-	-	0/0/0/0	0/0/0/0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NDP	M	381	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	O	363	-	-	0/0/0/0	0/0/0/0
2	SO4	O	364	-	-	0/0/0/0	0/0/0/0
2	SO4	O	371	-	-	0/0/0/0	0/0/0/0
3	NDP	O	372	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	P	363	-	-	0/0/0/0	0/0/0/0
2	SO4	P	364	-	-	0/0/0/0	0/0/0/0
2	SO4	P	372	-	-	0/0/0/0	0/0/0/0
3	NDP	P	373	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	Q	363	-	-	0/0/0/0	0/0/0/0
2	SO4	Q	364	-	-	0/0/0/0	0/0/0/0
2	SO4	Q	373	-	-	0/0/0/0	0/0/0/0
2	SO4	Q	374	-	-	0/0/0/0	0/0/0/0
3	NDP	Q	375	-	7/7/14/17	1/30/77/77	0/4/5/5
2	SO4	R	363	-	-	0/0/0/0	0/0/0/0
2	SO4	R	364	-	-	0/0/0/0	0/0/0/0
2	SO4	R	376	-	-	0/0/0/0	0/0/0/0
3	NDP	R	377	-	7/7/14/17	1/30/77/77	0/4/5/5

The worst 5 of 287 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	R	377	NDP	O2D-C2D	-14.33	1.10	1.43
3	H	379	NDP	O2D-C2D	-13.78	1.11	1.43
3	D	365	NDP	O2D-C2D	-13.58	1.11	1.43
3	C	365	NDP	O2D-C2D	-13.57	1.11	1.43
3	B	378	NDP	O2D-C2D	-13.47	1.12	1.43

The worst 5 of 358 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	376	NDP	O4D-C1D-C2D	-17.75	67.36	106.64
3	O	372	NDP	O4D-C1D-C2D	-13.93	75.80	106.64
3	A	376	NDP	C2D-C1D-N1N	-13.76	77.91	113.32
3	L	369	NDP	O4D-C1D-C2D	-12.54	78.89	106.64
3	A	376	NDP	C1D-N1N-C2N	-12.30	100.23	121.09

5 of 84 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	378	NDP	C1B

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atom
3	B	378	NDP	C3D
3	B	378	NDP	C2B
3	B	378	NDP	C4D
3	B	378	NDP	C2D

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	R	377	NDP	PA-O5B-C5B-C4B
3	A	376	NDP	PA-O5B-C5B-C4B
3	O	372	NDP	PA-O5B-C5B-C4B
3	D	365	NDP	PA-O5B-C5B-C4B
3	I	380	NDP	PA-O5B-C5B-C4B

There are no ring outliers.

29 monomers are involved in 181 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	376	NDP	19	0
2	B	364	SO4	2	0
3	B	378	NDP	13	0
3	C	365	NDP	16	0
2	D	364	SO4	2	0
3	D	365	NDP	6	0
2	H	363	SO4	1	0
2	H	364	SO4	2	0
2	H	378	SO4	1	0
3	H	379	NDP	14	0
2	I	366	SO4	5	0
2	I	379	SO4	1	0
3	I	380	NDP	10	0
2	L	367	SO4	1	0
3	L	369	NDP	14	0
2	M	370	SO4	3	0
3	M	381	NDP	10	0
2	O	364	SO4	2	0
2	O	371	SO4	2	0
3	O	372	NDP	15	0
2	P	363	SO4	1	0
2	P	364	SO4	4	0
2	P	372	SO4	1	0
3	P	373	NDP	11	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	Q	373	SO4	1	0
3	Q	375	NDP	16	0
2	R	363	SO4	1	0
2	R	376	SO4	2	0
3	R	377	NDP	5	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	2
1	A	1
1	L	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	210:ALA	C	211:ALA	N	1.68
1	A	206:THR	C	207:SER	N	1.18
1	L	205:PRO	C	206:THR	N	1.16
1	B	333:GLN	C	334:GLY	N	1.12

## 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	A	337/365 (92%)	0.12	2 (0%) 89 87	7, 24, 42, 53	1 (0%)
1	B	337/365 (92%)	0.02	1 (0%) 93 93	7, 18, 35, 54	1 (0%)
1	C	337/365 (92%)	-0.06	1 (0%) 93 93	6, 16, 31, 43	1 (0%)
1	D	336/365 (92%)	0.25	12 (3%) 43 42	7, 27, 56, 65	1 (0%)
1	H	336/365 (92%)	0.28	9 (2%) 55 52	9, 28, 55, 67	1 (0%)
1	I	337/365 (92%)	-0.00	1 (0%) 93 93	7, 22, 37, 50	1 (0%)
1	L	336/365 (92%)	0.32	14 (4%) 37 35	11, 30, 57, 73	1 (0%)
1	M	337/365 (92%)	0.01	0 100 100	10, 21, 38, 44	1 (0%)
1	O	337/365 (92%)	-0.04	2 (0%) 89 87	6, 17, 32, 47	1 (0%)
1	P	337/365 (92%)	0.18	4 (1%) 79 77	9, 25, 42, 54	1 (0%)
1	Q	337/365 (92%)	-0.04	1 (0%) 93 93	7, 18, 34, 44	1 (0%)
1	R	337/365 (92%)	0.30	13 (3%) 40 39	6, 26, 55, 68	1 (0%)
All	All	4041/4380 (92%)	0.11	60 (1%) 74 72	6, 22, 50, 73	12 (0%)

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	332	TRP	5.1
1	H	332	TRP	4.0
1	L	61	ASP	3.6
1	D	80	VAL	3.5
1	L	60(A)	GLY	3.3

### 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 ⓘ

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	O	371	5/5	0.49	0.88	27.18	148,148,148,148	5
2	SO4	M	370	5/5	0.78	0.48	8.58	70,71,71,71	5
2	SO4	M	380	5/5	0.92	0.41	7.90	65,65,66,66	5
2	SO4	I	366	5/5	0.76	0.38	6.14	69,69,70,70	5
2	SO4	Q	373	5/5	0.84	0.43	5.64	74,74,74,75	5
2	SO4	I	379	5/5	0.92	0.23	3.81	64,64,65,65	0
2	SO4	B	377	5/5	0.85	0.36	3.81	77,77,77,77	5
2	SO4	H	364	5/5	0.82	0.26	2.42	55,57,57,58	5
2	SO4	Q	374	5/5	0.95	0.24	2.30	51,52,52,52	5
2	SO4	O	364	5/5	0.93	0.21	2.00	56,56,56,57	5
2	SO4	D	364	5/5	0.89	0.22	1.60	54,54,55,55	5
2	SO4	Q	364	5/5	0.95	0.17	1.49	40,41,42,42	0
3	NDP	I	380	48/48	0.94	0.19	1.48	19,26,42,46	0
2	SO4	B	364	5/5	0.94	0.23	1.47	40,42,43,45	0
2	SO4	R	364	5/5	0.92	0.23	1.45	50,50,51,51	5
3	NDP	C	365	48/48	0.93	0.20	1.33	17,24,40,43	0
2	SO4	L	368	5/5	0.91	0.21	1.29	49,50,51,53	0
3	NDP	O	372	48/48	0.92	0.20	1.25	16,25,39,43	0
3	NDP	H	379	48/48	0.89	0.23	1.17	31,38,54,57	0
3	NDP	B	378	48/48	0.93	0.19	1.04	17,23,45,46	0
3	NDP	L	369	48/48	0.89	0.22	0.96	31,40,58,60	0
3	NDP	Q	375	48/48	0.93	0.19	0.88	25,30,49,51	0
3	NDP	M	381	48/48	0.93	0.18	0.85	21,26,41,43	0
2	SO4	C	364	5/5	0.87	0.17	0.80	57,57,59,59	0
3	NDP	R	377	48/48	0.90	0.23	0.73	30,43,53,54	0
2	SO4	A	364	5/5	0.95	0.21	0.60	48,49,51,51	0
3	NDP	D	365	48/48	0.92	0.22	0.49	30,40,55,57	0
3	NDP	P	373	48/48	0.94	0.17	0.29	11,26,45,47	0
3	NDP	A	376	48/48	0.93	0.17	0.26	19,25,47,48	0
2	SO4	P	364	5/5	0.96	0.18	0.16	44,44,45,45	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	R	376	5/5	0.93	0.17	0.10	62,63,63,63	0
2	SO4	A	375	5/5	0.97	0.16	-0.16	55,55,55,56	0
2	SO4	H	378	5/5	0.94	0.17	-0.31	55,56,56,56	0
2	SO4	P	372	5/5	0.97	0.13	-0.94	61,61,62,62	0
2	SO4	D	363	5/5	0.94	0.22	-	60,60,61,62	0
2	SO4	C	363	5/5	0.95	0.19	-	62,62,63,63	0
2	SO4	H	363	5/5	0.96	0.20	-	60,61,61,62	0
2	SO4	R	363	5/5	0.95	0.22	-	35,36,37,37	5
2	SO4	I	365	5/5	0.96	0.19	-	64,65,65,66	0
2	SO4	Q	363	5/5	0.94	0.19	-	51,51,52,53	0
2	SO4	P	363	5/5	0.95	0.20	-	60,60,61,62	0
2	SO4	B	363	5/5	0.94	0.23	-	52,53,54,54	0
2	SO4	A	363	5/5	0.87	0.28	-	56,56,57,58	5
2	SO4	L	367	5/5	0.95	0.22	-	63,63,64,65	0
2	SO4	O	363	5/5	0.94	0.25	-	50,51,51,52	5
2	SO4	M	369	5/5	0.93	0.23	-	60,60,61,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.