Full wwPDB X-ray Structure Validation Report (i)

Oct 11, 2016 – 12:00 PM EDT

PDB ID : 5AZI

Title : Crystal structure of glycerol kinase from Trypanosoma brucei gambiense complexed with 4NP

Authors : Balogun, E.O.; Inaoka, D.K.; Shiba, T.; Tokuoka, S.M.; Tokumasu, F.; Sakamoto, K.; Michels, P.A.M.; Harada, S.; Kita, K.

Deposited on : 2015-10-08

Resolution : 2.45 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

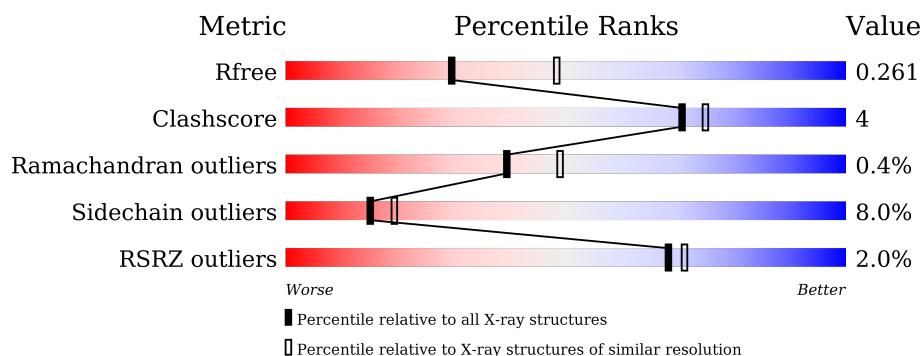|		
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

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 2.45 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	Percentile Ranks	Value
Rfree		0.261
Clashscore		4
Ramachandran outliers		0.4%
Sidechain outliers		8.0%
RSRZ outliers		2.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	4776 (2.50-2.42)
Clashscore	102246	1030 (2.48-2.44)
Ramachandran outliers	100387	1024 (2.48-2.44)
Sidechain outliers	100360	1024 (2.48-2.44)
RSRZ outliers	91569	4787 (2.50-2.42)

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. 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	A	518	.% 81% 16% ••
1	C	518	.% 87% 11% ••
2	B	518	3% 82% 15% ••
2	D	518	4% 82% 16% ••

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
3	4NP	A	601	-	-	-	X
3	4NP	C	601	-	-	-	X
4	GOL	C	602	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16014 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	A	513	Total	C	N	O	S	0	0	0
			3951	2493	694	731	33			
1	C	513	Total	C	N	O	S	0	0	0
			3951	2493	694	731	33			

There are 12 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP D3KVM3
A	-4	ILE	-	expression tag	UNP D3KVM3
A	-3	ASP	-	expression tag	UNP D3KVM3
A	-2	PRO	-	expression tag	UNP D3KVM3
A	-1	PHE	-	expression tag	UNP D3KVM3
A	0	THR	-	expression tag	UNP D3KVM3
C	-5	GLY	-	expression tag	UNP D3KVM3
C	-4	ILE	-	expression tag	UNP D3KVM3
C	-3	ASP	-	expression tag	UNP D3KVM3
C	-2	PRO	-	expression tag	UNP D3KVM3
C	-1	PHE	-	expression tag	UNP D3KVM3
C	0	THR	-	expression tag	UNP D3KVM3

- Molecule 2 is a protein called Glycerol kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	513	Total	C	N	O	P	S	0	0	0
			3955	2493	694	734	1	33			
2	D	513	Total	C	N	O	P	S	0	0	0
			3955	2493	694	734	1	33			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	GLY	-	expression tag	UNP D3KVM3
B	-4	ILE	-	expression tag	UNP D3KVM3
B	-3	ASP	-	expression tag	UNP D3KVM3
B	-2	PRO	-	expression tag	UNP D3KVM3
B	-1	PHE	-	expression tag	UNP D3KVM3
B	0	THR	-	expression tag	UNP D3KVM3
D	-5	GLY	-	expression tag	UNP D3KVM3
D	-4	ILE	-	expression tag	UNP D3KVM3
D	-3	ASP	-	expression tag	UNP D3KVM3
D	-2	PRO	-	expression tag	UNP D3KVM3
D	-1	PHE	-	expression tag	UNP D3KVM3
D	0	THR	-	expression tag	UNP D3KVM3

- Molecule 3 is 4-NITROPHENYL PHOSPHATE (three-letter code: 4NP) (formula: $C_6H_6NO_6P$).
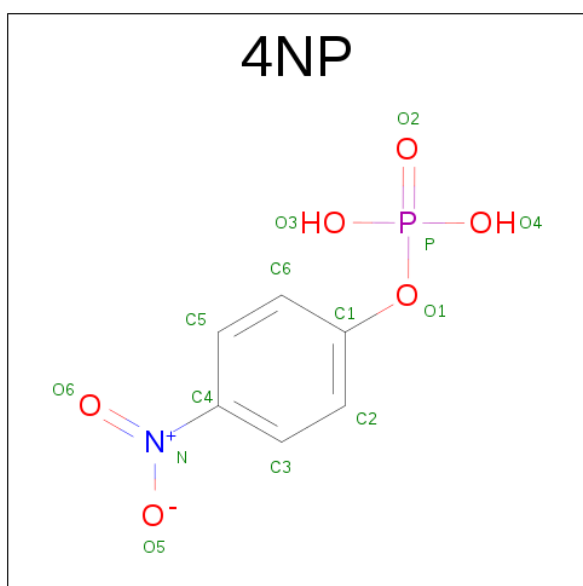
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			14	6	1	6	1		
3	C	1	Total	C	N	O	P	0	0
			14	6	1	6	1		

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

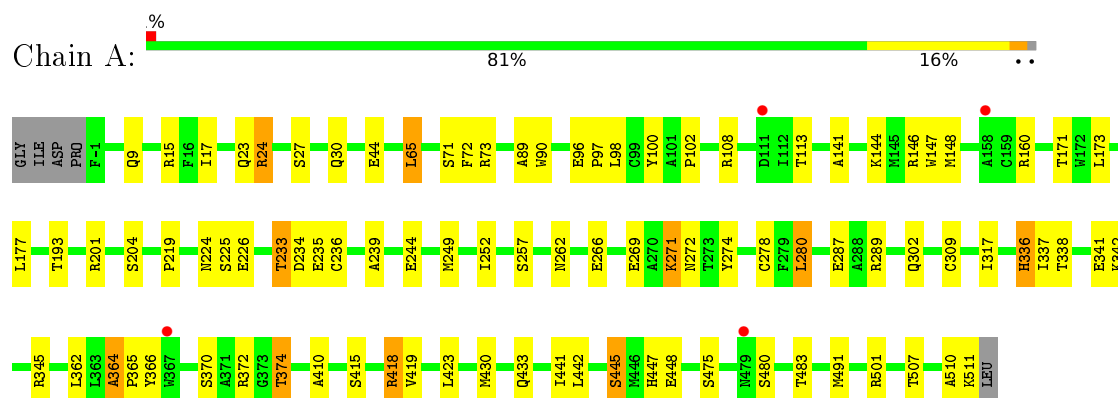| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total 6	C 3	O 3	0	0
4	B	1	Total 6	C 3	O 3	0	0
4	C	1	Total 6	C 3	O 3	0	0
4	D	1	Total 6	C 3	O 3	0	0

- Molecule 5 is water.

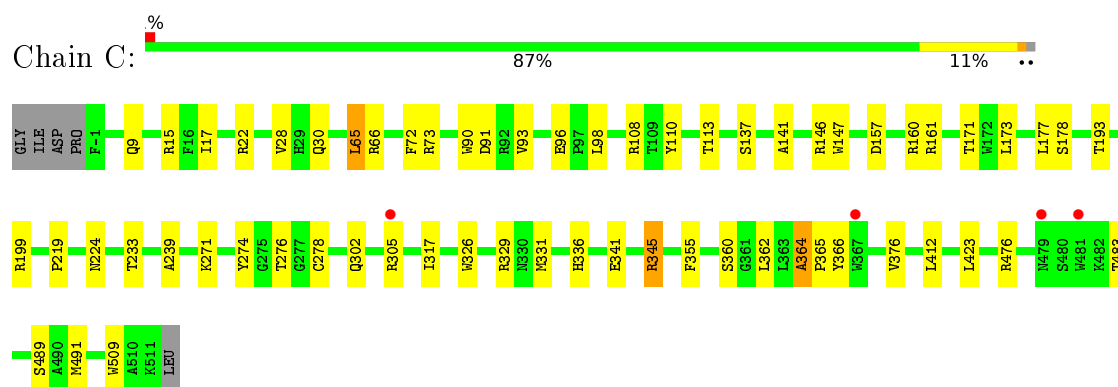| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	47	Total 47	O 47	0	0
5	B	24	Total 24	O 24	0	0
5	C	49	Total 49	O 49	0	0
5	D	30	Total 30	O 30	0	0

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 errors 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.
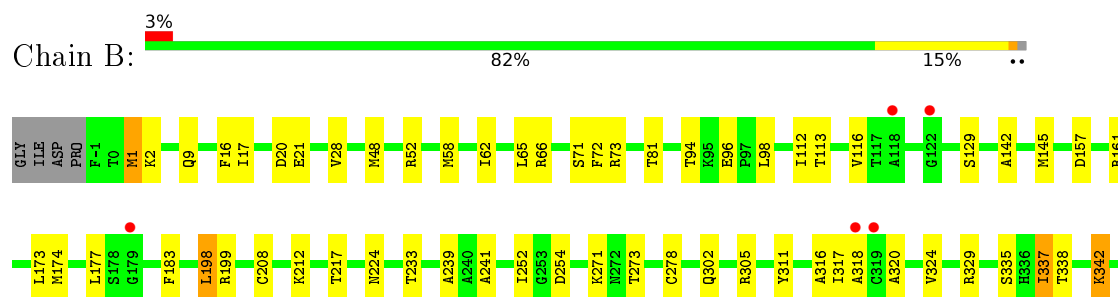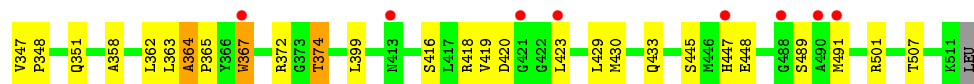
- Molecule 1: Glycerol kinase

Chain A:

- Molecule 1: Glycerol kinase

Chain C:

- Molecule 2: Glycerol kinase

Chain B:

V347 P348 Q351 A358 L362 L363 A364 P365 Y366 W367 R372 G373 T374 L399 N413 S416 L417 R418 V419 D420 G421 G422 L423 L429 M430 Q433 S445 M446 H447 E448 G488 S489 A490 M491 R501 T507 K511 LEU

• Molecule 2: Glycerol kinase

4%

Chain D:

82% 16%

GLY ILE ASP PRO F-1 Q9 R15 F16 I17 V28 K56 C57 M58 L65 R66 R73 K74 Q83 W90 D91 R92 V93 T94 K95 E96 P97 L98 C99 T113 K114 A118 G122 F127 K130 A142 M145 M148 A158 C159 R160 L173

L177 S178 G179 F183 D186 V187 T188 D197 L198 R201 E209 E216 T217 N224 D234 E244 R245 E269 A270 K271 N272 T273 Y274 G275 L280 V284 G294 Q302 L313 I317 A318 C319 V324 H336 I337 T338 K342 R345 S346 V347

P348 A358 L362 L363 A364 P365 Y366 W367 R372 G373 T374 G377 L399 L412 N413 L414 L417 R418 V419 D420 G421 G422 L423 L429 M430 Q433 V439 D440 I441 L442 M446 H447 E448 T449 T466 S480 T483 M491 R501 E502 K505 K511

LEU

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	63.67Å 120.25Å 154.59Å 90.00° 89.96° 90.00°	Depositor
Resolution (Å)	20.00 – 2.45 19.85 – 2.45	Depositor EDS
% Data completeness (in resolution range)	84.8 (20.00-2.45) 84.7 (19.85-2.45)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.08 (at 2.44Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.196 , 0.260 0.201 , 0.261	Depositor DCC
R_{free} test set	3627 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	37.6	Xtriage
Anisotropy	0.180	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 24.3	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.480 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	16014	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.23% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: TPO, GOL, 4NP

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	0.50	0/4032	0.69	1/5456 (0.0%)
1	C	0.50	0/4032	0.68	0/5456
2	B	0.47	0/4024	0.68	0/5443
2	D	0.47	0/4024	0.67	0/5443
All	All	0.49	0/16112	0.68	1/21798 (0.0%)

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	C	0	1
2	B	0	1
All	All	0	2

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	24	ARG	NE-CZ-NH1	5.75	123.17	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	316	ALA	Peptide
1	C	345	ARG	Sidechain

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	A	3951	0	3963	47	0
1	C	3951	0	3963	22	0
2	B	3955	0	3962	39	0
2	D	3955	0	3962	33	0
3	A	14	0	4	0	0
3	C	14	0	4	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	A	47	0	0	1	0
5	B	24	0	0	0	0
5	C	49	0	0	0	0
5	D	30	0	0	0	0
All	All	16014	0	15890	139	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.

All (139) 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:224:ASN:HD22	1:A:302:GLN:H	1.26	0.84
2:D:364:ALA:HB1	2:D:365:PRO:CD	2.15	0.77
1:C:364:ALA:HB1	1:C:365:PRO:CD	2.15	0.77
2:B:364:ALA:HB1	2:B:365:PRO:CD	2.16	0.76
1:A:364:ALA:HB1	1:A:365:PRO:CD	2.19	0.72
1:A:341:GLU:OE2	1:A:345:ARG:NH2	2.24	0.71
1:A:364:ALA:HB1	1:A:365:PRO:HD3	1.73	0.70
1:A:274:TYR:HB3	1:A:423:LEU:HB2	1.72	0.70
2:B:324:VAL:HG21	2:B:423:LEU:HD21	1.74	0.70
2:B:364:ALA:HB1	2:B:365:PRO:HD3	1.72	0.70
2:B:419:VAL:HG21	2:B:430:MET:SD	2.32	0.69
2:D:269:GLU:OE1	2:D:418:ARG:NH2	2.26	0.67
2:B:224:ASN:HD22	2:B:302:GLN:H	1.43	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:374:THR:CG2	1:A:507:THR:HG22	2.26	0.65
2:D:364:ALA:HB1	2:D:365:PRO:HD3	1.78	0.64
1:A:336:HIS:CE1	1:A:338:THR:HG1	2.14	0.64
1:A:372:ARG:O	1:A:374:THR:HG22	1.97	0.64
2:D:15:ARG:NH2	2:D:448:GLU:OE2	2.31	0.63
2:D:224:ASN:HD22	2:D:302:GLN:H	1.47	0.63
2:D:419:VAL:HG21	2:D:430:MET:SD	2.39	0.61
2:B:372:ARG:O	2:B:374:THR:HG22	2.00	0.61
1:A:272:ASN:HB2	1:A:280:LEU:HD12	1.82	0.60
1:A:374:THR:HG21	1:A:507:THR:HG22	1.83	0.60
1:C:364:ALA:HB1	1:C:365:PRO:HD3	1.85	0.59
1:A:15:ARG:HD3	1:A:30:GLN:NE2	2.18	0.59
2:D:15:ARG:HH22	2:D:448:GLU:CD	2.05	0.59
2:B:337:ILE:HD13	2:B:337:ILE:H	1.67	0.59
1:C:224:ASN:HD22	1:C:302:GLN:H	1.50	0.58
2:B:364:ALA:CB	2:B:365:PRO:CD	2.80	0.58
2:D:414:LEU:HD23	2:D:439:VAL:HG21	1.84	0.58
2:D:364:ALA:CB	2:D:365:PRO:CD	2.81	0.57
1:C:364:ALA:CB	1:C:365:PRO:CD	2.83	0.57
2:B:374:THR:CG2	2:B:507:THR:HG22	2.34	0.57
1:A:233:THR:O	1:A:239:ALA:HB2	2.05	0.57
1:C:274:TYR:HB3	1:C:423:LEU:HB2	1.86	0.57
1:A:17:ILE:HD13	1:A:448:GLU:HG2	1.87	0.56
1:A:364:ALA:CB	1:A:365:PRO:CD	2.84	0.56
2:B:374:THR:HG21	2:B:507:THR:HA	1.87	0.56
2:B:17:ILE:HD13	2:B:28:VAL:HG23	1.87	0.55
2:B:62:ILE:HG21	2:B:241:ALA:HB1	1.87	0.55
1:A:226:GLU:O	1:A:249:MET:HA	2.06	0.55
1:A:90:TRP:HA	1:A:98:LEU:HD22	1.87	0.55
2:D:17:ILE:HD13	2:D:28:VAL:HG23	1.88	0.55
2:D:502:GLU:HA	2:D:505:LYS:HD3	1.89	0.55
1:C:17:ILE:HG12	1:C:28:VAL:HG23	1.89	0.55
1:C:65:LEU:HD13	1:C:72:PHE:CG	2.42	0.54
2:D:91:ASP:OD1	2:D:94:THR:HG23	2.07	0.54
1:A:96:GLU:HG3	1:A:97:PRO:HD2	1.90	0.54
1:C:364:ALA:HB1	1:C:365:PRO:HD2	1.90	0.54
1:A:23:GLN:HE22	1:A:475:SER:HB2	1.73	0.54
1:C:233:THR:O	1:C:239:ALA:HB2	2.08	0.53
1:C:341:GLU:OE2	1:C:345:ARG:NH2	2.41	0.53
1:C:141:ALA:HB3	1:C:193:THR:HA	1.91	0.52
2:B:1:MET:HE3	2:B:1:MET:HA	1.91	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:370:SER:HA	1:A:510:ALA:HB3	1.92	0.52
2:D:446:MET:SD	2:D:448:GLU:HB3	2.50	0.52
1:A:171:THR:HG21	1:A:219:PRO:HG3	1.92	0.52
2:D:433:GLN:HG2	2:D:441:ILE:HD11	1.92	0.52
2:D:16:PHE:CD2	2:D:58:MET:HA	2.45	0.52
1:C:110:TYR:O	1:C:113:THR:HB	2.10	0.51
1:A:141:ALA:HB3	1:A:193:THR:HA	1.93	0.51
1:A:224:ASN:ND2	1:A:302:GLN:H	2.03	0.51
2:B:1:MET:HE1	2:B:21:GLU:CG	2.41	0.51
2:B:358:ALA:HB2	2:B:362:LEU:HD13	1.92	0.51
1:A:374:THR:HG21	1:A:507:THR:HA	1.93	0.50
2:D:358:ALA:HB2	2:D:362:LEU:HD13	1.94	0.50
2:B:374:THR:HG23	2:B:507:THR:HG22	1.93	0.50
1:A:266:GLU:O	1:A:269:GLU:HG3	2.11	0.49
1:C:278:CYS:HB2	1:C:317:ILE:HB	1.94	0.49
1:A:15:ARG:NH1	5:A:701:HOH:O	2.44	0.49
2:D:280:LEU:HD11	2:D:417:LEU:HD11	1.95	0.49
2:B:1:MET:SD	2:B:20:ASP:HB2	2.54	0.48
1:A:15:ARG:HD3	1:A:30:GLN:HE21	1.79	0.48
2:D:372:ARG:O	2:D:374:THR:OG1	2.30	0.48
2:B:112:ILE:O	2:B:116:VAL:HG23	2.14	0.47
2:D:273:THR:HG23	2:D:420:ASP:OD2	2.14	0.47
2:B:278:CYS:SG	2:B:320:ALA:HB3	2.53	0.47
1:C:376:VAL:HG12	2:D:374:THR:HG23	1.96	0.47
2:D:83:GLN:HG3	2:D:83:GLN:O	2.14	0.47
1:A:287:GLU:OE2	1:A:289:ARG:NE	2.48	0.47
1:A:44:GLU:HG2	1:A:100:TYR:HB3	1.96	0.47
1:C:90:TRP:HA	1:C:98:LEU:HD22	1.95	0.46
1:C:171:THR:HG21	1:C:219:PRO:HG3	1.96	0.46
2:B:71:SER:O	2:B:72:PHE:C	2.54	0.46
2:B:233:THR:O	2:B:239:ALA:HB2	2.16	0.46
1:A:272:ASN:HD21	1:A:278:CYS:HB3	1.81	0.45
2:B:429:LEU:O	2:B:433:GLN:HB2	2.17	0.45
2:B:367:TRP:CG	2:B:367:TRP:O	2.70	0.44
2:B:81:THR:HA	2:B:252:ILE:O	2.17	0.44
1:A:201:ARG:HD3	1:A:309:CYS:SG	2.57	0.44
2:D:99:CYS:SG	2:D:148:MET:HE3	2.58	0.44
2:B:183:PHE:HZ	2:B:217:THR:O	2.00	0.44
1:A:262:ASN:HB3	1:A:418:ARG:HG3	1.99	0.44
1:A:289:ARG:HD3	1:A:410:ALA:HA	2.00	0.43
1:A:108:ARG:HD3	1:A:147:TRP:CE2	2.54	0.43

Continued on next page...

Full wwPDB X-ray Structure Validation Report

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:233:THR:OG1	1:A:234:ASP:N	2.52	0.43
1:A:271:LYS:C	1:A:271:LYS:HD2	2.38	0.43
2:B:1:MET:CE	2:B:1:MET:HA	2.48	0.43
2:D:90:TRP:HA	2:D:98:LEU:HD22	1.99	0.43
1:A:225:SER:O	1:A:226:GLU:HB3	2.17	0.43
2:D:284:VAL:HG11	2:D:313:LEU:HD12	1.99	0.43
2:D:447:HIS:O	2:D:449:THR:N	2.46	0.43
1:C:108:ARG:HD3	1:C:147:TRP:CZ2	2.54	0.43
1:A:419:VAL:HG21	1:A:430:MET:SD	2.59	0.43
1:A:89:ALA:HB2	1:A:148:MET:CE	2.49	0.43
2:B:16:PHE:CD2	2:B:58:MET:HA	2.54	0.43
1:A:336:HIS:CE1	1:A:338:THR:OG1	2.72	0.42
2:D:324:VAL:HG21	2:D:423:LEU:HD21	2.00	0.42
1:C:171:THR:HG21	1:C:219:PRO:CG	2.50	0.42
2:B:142:ALA:O	2:B:145:MET:HB2	2.19	0.42
2:D:183:PHE:HZ	2:D:217:THR:O	2.03	0.42
1:A:27:SER:HB2	1:A:65:LEU:HG	2.01	0.42
2:B:318:ALA:HB2	2:B:363:LEU:CD1	2.50	0.42
1:C:15:ARG:CG	1:C:30:GLN:HE21	2.33	0.42
1:C:91:ASP:OD2	1:C:161:ARG:NH2	2.53	0.42
1:A:235:GLU:O	1:A:236:CYS:SG	2.75	0.42
2:B:157:ASP:OD1	2:B:161:ARG:NH1	2.52	0.42
2:B:1:MET:HE2	2:B:2:LYS:N	2.35	0.41
1:A:252:ILE:HG23	1:A:257:SER:HB2	2.02	0.41
2:D:186:ASP:OD2	2:D:188:THR:OG1	2.24	0.41
2:B:212:LYS:HG3	2:B:212:LYS:O	2.20	0.41
2:D:429:LEU:O	2:D:433:GLN:HB2	2.21	0.41
2:B:17:ILE:CD1	2:B:28:VAL:HG23	2.50	0.41
2:B:81:THR:OG1	2:B:254:ASP:HA	2.21	0.41
2:D:197:ASP:O	2:D:201:ARG:N	2.53	0.41
2:D:142:ALA:O	2:D:145:MET:HB2	2.19	0.41
2:D:201:ARG:HH11	2:D:302:GLN:NE2	2.19	0.41
1:A:445:SER:OG	1:A:480:SER:O	2.38	0.41
2:B:198:LEU:HD22	2:B:311:TYR:HE1	1.86	0.41
1:A:65:LEU:HD13	1:A:72:PHE:CG	2.55	0.41
2:B:273:THR:HG23	2:B:420:ASP:OD1	2.20	0.41
1:A:102:PRO:HB2	1:A:144:LYS:HD3	2.03	0.40
2:B:48:MET:O	2:B:52:ARG:HG3	2.22	0.40
1:C:509:TRP:CD1	2:D:377:GLY:HA2	2.55	0.40
1:A:374:THR:HG23	1:A:507:THR:HG22	2.02	0.40
1:A:433:GLN:HG3	1:A:441:ILE:HD11	2.03	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:342:LYS:HA	2:B:342:LYS:HE3	2.03	0.40
2:B:348:PRO:HG2	2:B:351:GLN:HE21	1.86	0.40
1:C:326:TRP:CH2	1:C:331:MET:HG3	2.56	0.40

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	511/518 (99%)	480 (94%)	30 (6%)	1 (0%)	52	64
1	C	511/518 (99%)	482 (94%)	28 (6%)	1 (0%)	52	64
2	B	510/518 (98%)	482 (94%)	24 (5%)	4 (1%)	24	28
2	D	510/518 (98%)	476 (93%)	31 (6%)	3 (1%)	30	35
All	All	2042/2072 (99%)	1920 (94%)	113 (6%)	9 (0%)	39	49

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	448	GLU
2	D	448	GLU
1	A	364	ALA
2	B	364	ALA
1	C	364	ALA
2	D	364	ALA
2	B	367	TRP
2	D	348	PRO
2	B	208	CYS

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	A	420/425 (99%)	389 (93%)	31 (7%)	17	22
1	C	420/425 (99%)	391 (93%)	29 (7%)	19	25
2	B	419/424 (99%)	386 (92%)	33 (8%)	15	20
2	D	419/424 (99%)	377 (90%)	42 (10%)	9	10
All	All	1678/1698 (99%)	1543 (92%)	135 (8%)	15	19

All (135) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	9	GLN
1	A	24	ARG
1	A	65	LEU
1	A	71	SER
1	A	73	ARG
1	A	113	THR
1	A	146	ARG
1	A	160	ARG
1	A	173	LEU
1	A	177	LEU
1	A	204	SER
1	A	233	THR
1	A	244	GLU
1	A	271	LYS
1	A	280	LEU
1	A	317	ILE
1	A	336	HIS
1	A	337	ILE
1	A	342	LYS
1	A	362	LEU
1	A	366	TYR
1	A	374	THR
1	A	415	SER
1	A	418	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	442	LEU
1	A	445	SER
1	A	447	HIS
1	A	483	THR
1	A	491	MET
1	A	501	ARG
1	A	511	LYS
2	B	1	MET
2	B	9	GLN
2	B	65	LEU
2	B	66	ARG
2	B	73	ARG
2	B	94	THR
2	B	96	GLU
2	B	98	LEU
2	B	113	THR
2	B	129	SER
2	B	173	LEU
2	B	174	MET
2	B	177	LEU
2	B	198	LEU
2	B	199	ARG
2	B	271	LYS
2	B	305	ARG
2	B	317	ILE
2	B	329	ARG
2	B	335	SER
2	B	337	ILE
2	B	338	THR
2	B	342	LYS
2	B	347	VAL
2	B	374	THR
2	B	399	LEU
2	B	416	SER
2	B	418	ARG
2	B	445	SER
2	B	447	HIS
2	B	489	SER
2	B	491	MET
2	B	501	ARG
1	C	9	GLN
1	C	22	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	65	LEU
1	C	66	ARG
1	C	73	ARG
1	C	93	VAL
1	C	96	GLU
1	C	137	SER
1	C	146	ARG
1	C	157	ASP
1	C	160	ARG
1	C	173	LEU
1	C	177	LEU
1	C	178	SER
1	C	199	ARG
1	C	271	LYS
1	C	276	THR
1	C	305	ARG
1	C	329	ARG
1	C	336	HIS
1	C	355	PHE
1	C	360	SER
1	C	362	LEU
1	C	366	TYR
1	C	412	LEU
1	C	476	ARG
1	C	483	THR
1	C	489	SER
1	C	491	MET
2	D	9	GLN
2	D	56	LYS
2	D	65	LEU
2	D	66	ARG
2	D	73	ARG
2	D	74	LYS
2	D	92	ARG
2	D	94	THR
2	D	96	GLU
2	D	98	LEU
2	D	113	THR
2	D	114	LYS
2	D	127	PHE
2	D	130	LYS
2	D	160	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	173	LEU
2	D	177	LEU
2	D	198	LEU
2	D	216	GLU
2	D	234	ASP
2	D	244	GLU
2	D	245	ARG
2	D	271	LYS
2	D	317	ILE
2	D	336	HIS
2	D	337	ILE
2	D	342	LYS
2	D	345	ARG
2	D	346	SER
2	D	362	LEU
2	D	366	TYR
2	D	399	LEU
2	D	412	LEU
2	D	413	ASN
2	D	433	GLN
2	D	442	LEU
2	D	447	HIS
2	D	466	THR
2	D	480	SER
2	D	483	THR
2	D	491	MET
2	D	501	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	23	GLN
1	A	30	GLN
1	A	36	HIS
1	A	224	ASN
1	A	272	ASN
1	A	302	GLN
2	B	23	GLN
2	B	30	GLN
2	B	224	ASN
2	B	302	GLN
2	B	351	GLN

Continued from previous page...

Mol	Chain	Res	Type
1	C	23	GLN
1	C	30	GLN
1	C	36	HIS
1	C	224	ASN
1	C	272	ASN
1	C	336	HIS
2	D	23	GLN
2	D	151	ASN
2	D	224	ASN
2	D	302	GLN
2	D	351	GLN
2	D	413	ASN
2	D	433	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues 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	TPO	B	276	2	7,10,11	0.64	0	10,14,16	1.55	1 (10%)
2	TPO	D	276	2	7,10,11	0.75	0	10,14,16	1.47	2 (20%)

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	TPO	B	276	2	-	0/8/11/13	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TPO	D	276	2	-	0/8/11/13	0/0/0/0

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	276	TPO	CG2-CB-CA	-3.43	106.30	113.15
2	D	276	TPO	CG2-CB-CA	-3.14	106.87	113.15
2	D	276	TPO	P-OG1-CB	2.02	130.28	121.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	#$|Z| > 2$
3	4NP	A	601	-	11,14,14	0.67	0	17,20,20	1.10	1 (5%)
4	GOL	A	602	-	5,5,5	0.44	0	5,5,5	0.29	0
4	GOL	B	601	-	5,5,5	0.32	0	5,5,5	0.54	0
3	4NP	C	601	-	11,14,14	0.72	0	17,20,20	0.87	0
4	GOL	C	602	-	5,5,5	0.45	0	5,5,5	0.36	0
4	GOL	D	601	-	5,5,5	0.30	0	5,5,5	0.65	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	4NP	A	601	-	-	0/9/9/9	0/1/1/1
4	GOL	A	602	-	-	0/4/4/4	0/0/0/0
4	GOL	B	601	-	-	0/4/4/4	0/0/0/0
3	4NP	C	601	-	-	0/9/9/9	0/1/1/1
4	GOL	C	602	-	-	0/4/4/4	0/0/0/0
4	GOL	D	601	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	601	4NP	O6-N-C4	2.02	120.12	118.67

There are no chirality outliers.

There are no torsion outliers.

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 (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, 95th 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	A	513/518 (99%)	-0.09	4 (0%)	87	89	19, 37, 63, 79	1 (0%)
1	C	513/518 (99%)	-0.11	4 (0%)	87	89	18, 37, 61, 80	1 (0%)
2	B	512/518 (98%)	0.09	13 (2%)	61	63	19, 48, 73, 100	1 (0%)
2	D	512/518 (98%)	0.07	20 (3%)	43	47	21, 48, 75, 96	1 (0%)
All	All	2050/2072 (98%)	-0.01	41 (2%)	68	71	18, 41, 70, 100	4 (0%)

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	318	ALA	6.9
2	D	318	ALA	5.5
2	B	488	GLY	5.2
2	B	319	CYS	4.7
1	A	367	TRP	4.0
2	B	118	ALA	3.6
2	B	423	LEU	3.5
2	D	364	ALA	3.5
2	D	491	MET	3.5
2	B	179	GLY	3.2
2	B	491	MET	3.0
2	D	118	ALA	3.0
2	D	179	GLY	3.0
2	D	423	LEU	3.0
2	B	421	GLY	2.9
1	C	367	TRP	2.7
1	A	158	ALA	2.7
2	B	447	HIS	2.7
2	B	367	TRP	2.7
2	B	413	ASN	2.6
2	D	122	GLY	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	274	TYR	2.5
2	D	319	CYS	2.5
2	D	338	THR	2.5
1	C	481	TRP	2.4
2	D	275	GLY	2.4
2	D	421	GLY	2.3
2	B	490	ALA	2.3
2	D	367	TRP	2.3
2	D	422	GLY	2.3
1	C	479	ASN	2.2
2	D	420	ASP	2.2
2	D	294	GLY	2.2
1	C	305	ARG	2.2
1	A	111	ASP	2.1
1	A	479	ASN	2.1
2	D	158	ALA	2.1
2	D	440	ASP	2.1
2	B	122	GLY	2.1
2	D	342	LYS	2.1
2	D	209	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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. 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, 95th 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^2)	Q<0.9
2	TPO	D	276	11/12	0.91	0.12	-	45,50,51,51	0
2	TPO	B	276	11/12	0.91	0.12	-	44,51,55,57	0

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. 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, 95th 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^2)	Q<0.9
3	4NP	A	601	14/14	0.86	0.21	5.75	73,88,99,100	0
4	GOL	C	602	6/6	0.89	0.26	4.87	44,45,46,46	0
3	4NP	C	601	14/14	0.78	0.24	4.62	90,95,101,101	0
4	GOL	A	602	6/6	0.96	0.16	0.81	40,42,45,48	0
4	GOL	B	601	6/6	0.97	0.13	-0.02	24,26,28,30	0
4	GOL	D	601	6/6	0.99	0.12	-0.89	24,26,27,27	0

6.5 Other polymers (i)

There are no such residues in this entry.