wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 09:48 PM GMT

PDB ID	:	1QPB
Title	:	PYRUVATE DECARBOYXLASE FROM YEAST (FORM B) COMPLEXED WITH PYRUVAMIDE
Authors	:	Lu, G.; Dobritzsch, D.; Schneider, G.
Deposited on	:	1999-11-26
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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

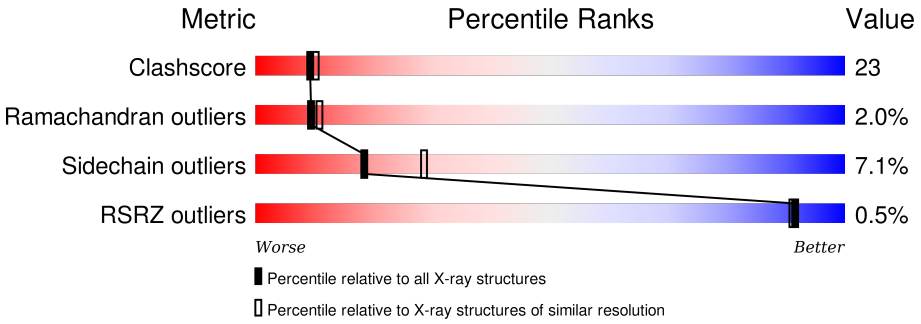
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.40 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))
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (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 >=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	563	55% 39%
1	B	563	% 59% 35%

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	MG	A	601	-	-	-	X
3	MG	B	601	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	PYM	A	602	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8757 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.

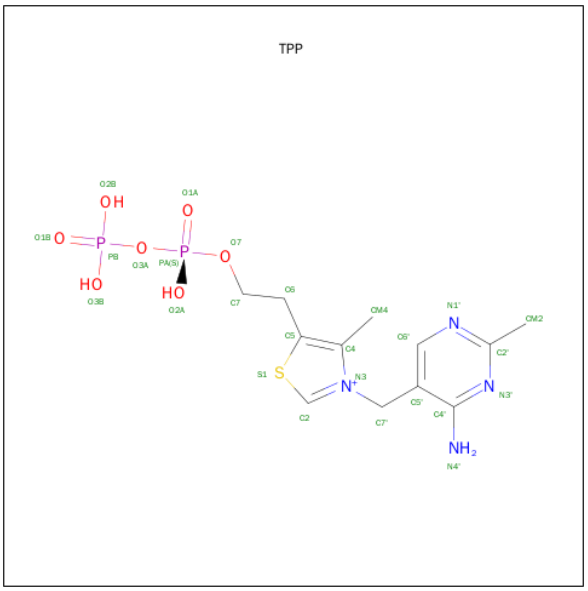
- Molecule 1 is a protein called PYRUVATE DECARBOXYLASE (FORM B).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	555	Total	C	N	O	S	75	0	0
			4275	2734	714	811	16			
1	B	555	Total	C	N	O	S	126	0	0
			4275	2734	714	811	16			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	206	VAL	ALA	VARIANT	UNP P06169
A	208	ILE	VAL	VARIANT	UNP P06169
A	336	ALA	THR	VARIANT	UNP P06169
A	538	VAL	ILE	VARIANT	UNP P06169
B	206	VAL	ALA	VARIANT	UNP P06169
B	208	ILE	VAL	VARIANT	UNP P06169
B	336	ALA	THR	VARIANT	UNP P06169
B	538	VAL	ILE	VARIANT	UNP P06169

- Molecule 2 is THIAMINE DIPHOSPHATE (three-letter code: TPP) (formula: $C_{12}H_{19}N_4O_7P_2S$).

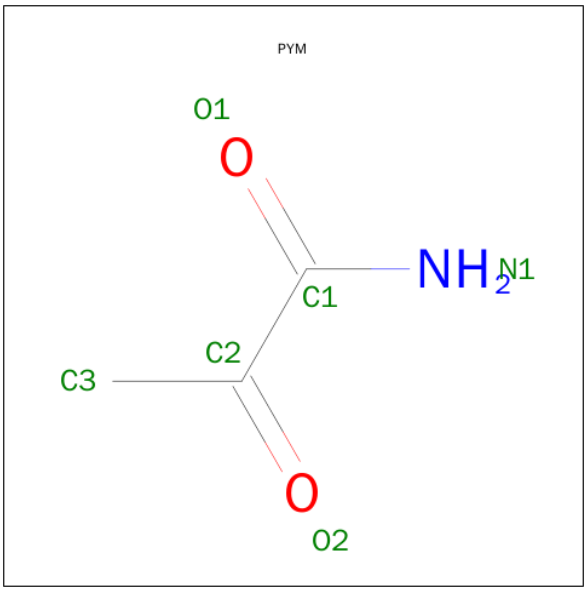


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	S	0	0
			26	12	4	7	2	1		
2	B	1	Total	C	N	O	P	S	0	0
			26	12	4	7	2	1		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Mg	0	0
			1	1		
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is PYRUVAMIDE (three-letter code: PYM) (formula: $C_3H_5NO_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 6 3 1 2	0	0
4	B	1	Total C N O 6 3 1 2	0	0

- Molecule 5 is water.

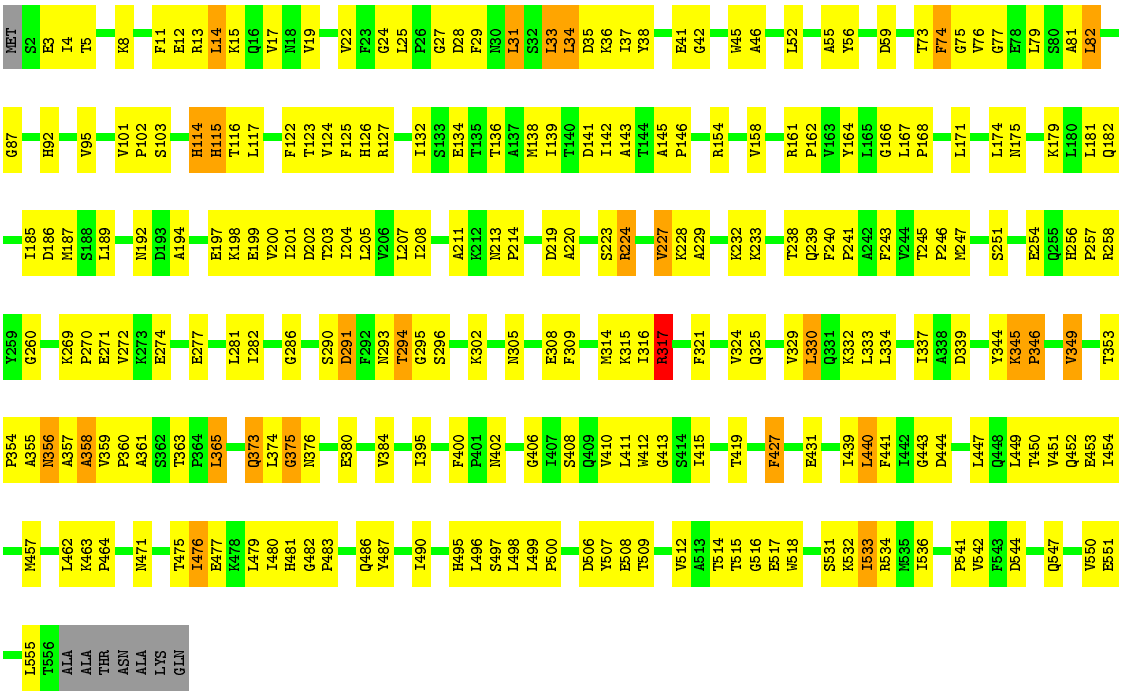
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	85	Total O 85 85	0	0
5	B	56	Total O 56 56	0	0

3 Residue-property plots (i)

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

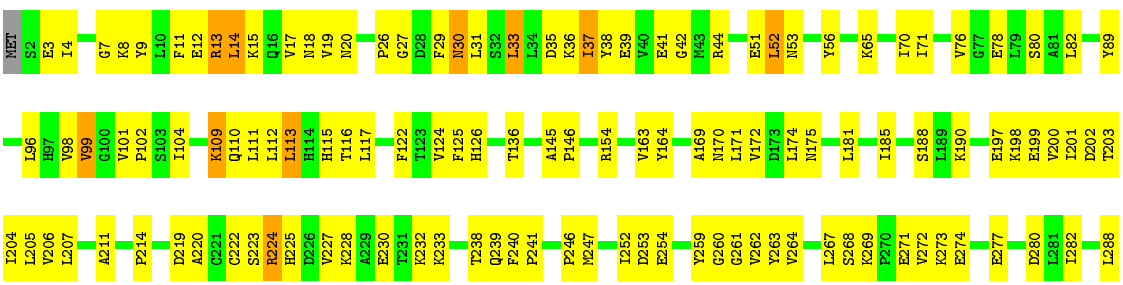
- Molecule 1: PYRUVATE DECARBOXYLASE (FORM B)

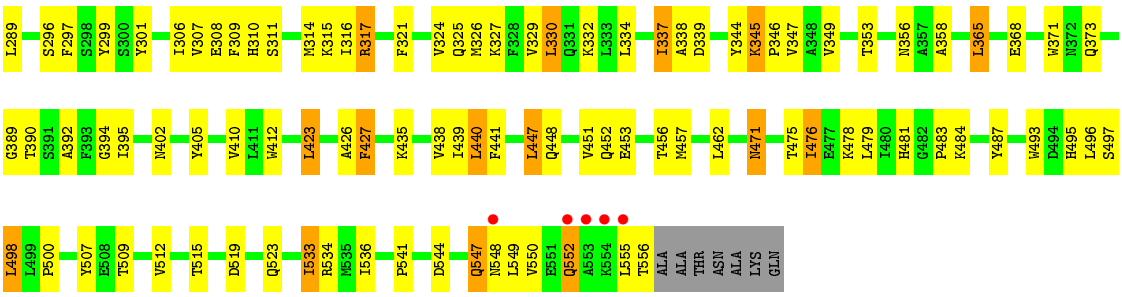
Chain A:



- Molecule 1: PYRUVATE DECARBOXYLASE (FORM B)

Chain B:




L289 S296 F297 S298 Y299 S300 Y301 I306 V307 E308 F309 H310 S311 M314 K315 I316 R317 F321 V324 Q325 M326 K327 F328 V329 L330 Q331 K332 I333 L334 I337 A338 D339 Y344 K345 P346 V347 A348 V349 T353 N356 A357 A358 L365 E368 W371 N372 Q373
G389 T390 S391 A392 F393 G394 I395 N402 Y405 V410 L411 W412 L423 A426 F427 K435 V438 I439 L440 F441 L447 Q448 V451 Q452 E453 T456 M457 L462 N471 T475 I476 E477 K478 L479 I480 H481 G482 P483 K484 Y487 W493 D494 H495 L496 S497
L498 L499 P500 Y507 E508 T509 V512 T515 D519 Q523 I533 R534 N535 I536 P541 D544 Q547 N548 L549 V550 E551 Q552 A553 K554 L555 T556 ALA ALA THR ASN ALA LYS GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	145.05Å 119.70Å 81.28Å 90.00° 120.21° 90.00°	Depositor
Resolution (Å)	15.00 – 2.40 24.46 – 2.39	Depositor EDS
% Data completeness (in resolution range)	89.0 (15.00-2.40) 89.4 (24.46-2.39)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.40 (at 2.39Å)	Xtriage
Refinement program	CNS 3.8	Depositor
R, R_{free}	0.233 , 0.287 0.242 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å^2)	29.3	Xtriage
Anisotropy	0.208	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 27.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Outliers	0 of 42568 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8757	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.40% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, PYM, TPP

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	0.41	0/4366	0.66	0/5937
1	B	0.38	0/4366	0.63	1/5937 (0.0%)
All	All	0.40	0/8732	0.64	1/11874 (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})
1	B	423	LEU	CA-CB-CG	5.55	128.07	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4275	0	4287	215	0
1	B	4275	0	4287	183	0
2	A	26	0	16	1	0
2	B	26	0	16	2	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	6	0	5	0	0
4	B	6	0	5	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	85	0	0	8	0
5	B	56	0	0	3	0
All	All	8757	0	8616	383	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 23.

The worst 5 of 383 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:B:533:ILE:H	1:B:533:ILE:HD12	1.22	1.04
1:B:267:LEU:HD21	1:B:552:GLN:HB3	1.37	1.04
1:A:533:ILE:H	1:A:533:ILE:HD12	1.24	1.01
1:B:496:LEU:HD22	1:B:509:THR:HB	1.51	0.92
1:B:224:ARG:HD2	1:B:405:TYR:OH	1.70	0.92

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	553/563 (98%)	493 (89%)	46 (8%)	14 (2%)	7 7
1	B	553/563 (98%)	489 (88%)	56 (10%)	8 (1%)	14 19
All	All	1106/1126 (98%)	982 (89%)	102 (9%)	22 (2%)	9 11

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	356	ASN
1	B	358	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	75	GLY
1	A	103	SER
1	A	346	PRO

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	462/467 (99%)	434 (94%)	28 (6%)	23	36
1	B	462/467 (99%)	424 (92%)	38 (8%)	14	21
All	All	924/934 (99%)	858 (93%)	66 (7%)	18	28

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	14	LEU
1	B	99	VAL
1	B	487	TYR
1	B	18	ASN
1	B	37	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	20	ASN
1	B	114	HIS
1	B	486	GLN
1	B	83	ASN
1	B	126	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	TPP	A	600	3	20,27,27	1.80	5 (25%)	31,40,40	1.77	7 (22%)
4	PYM	A	602	-	4,5,5	3.68	2 (50%)	3,6,6	4.34	1 (33%)
2	TPP	B	600	3	20,27,27	1.76	6 (30%)	31,40,40	1.69	8 (25%)
4	PYM	B	602	-	4,5,5	3.89	1 (25%)	3,6,6	4.36	1 (33%)

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	TPP	A	600	3	-	0/16/17/17	0/2/2/2
4	PYM	A	602	-	-	0/1/4/4	0/0/0/0
2	TPP	B	600	3	-	0/16/17/17	0/2/2/2
4	PYM	B	602	-	-	0/1/4/4	0/0/0/0

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	602	PYM	C3-C2	-7.38	1.34	1.50
4	A	602	PYM	C3-C2	-7.03	1.35	1.50

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	600	TPP	C4-N3	-4.91	1.35	1.39
2	A	600	TPP	PB-O3B	-2.48	1.45	1.54
2	B	600	TPP	C2'-N3'	-2.11	1.30	1.34

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	602	PYM	O2-C2-C3	-7.47	102.16	119.69
4	A	602	PYM	O2-C2-C3	-7.43	102.25	119.69
2	A	600	TPP	O3A-PA-O7	-4.85	90.08	102.94
2	B	600	TPP	O7-PA-O1A	-3.31	96.76	109.62
2	B	600	TPP	N1'-C2'-N3'	-2.73	120.54	125.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	TPP	1	0
2	B	600	TPP	2	0
4	B	602	PYM	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	A	545/563 (96%)	-0.83	0 100 100	10, 31, 53, 68	0
1	B	540/563 (95%)	-0.74	5 (0%) 85 85	10, 34, 64, 98	1 (0%)
All	All	1085/1126 (96%)	-0.78	5 (0%) 91 91	10, 32, 60, 98	1 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	555	LEU	3.9
1	B	552	GLN	2.9
1	B	548	ASN	2.2
1	B	553	ALA	2.1
1	B	554	LYS	2.0

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 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^{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	LLDF	B-factors(Å^2)	Q<0.9
3	MG	B	601	1/1	0.74	0.20	5.36	71,71,71,71	0
4	PYM	A	602	6/6	0.92	0.18	3.80	50,52,54,54	0
3	MG	A	601	1/1	0.91	0.14	2.60	26,26,26,26	0
4	PYM	B	602	6/6	0.95	0.11	0.24	44,46,46,46	0
2	TPP	B	600	26/26	0.99	0.08	-0.72	11,18,22,23	0
2	TPP	A	600	26/26	0.98	0.08	-0.94	18,24,27,28	0

6.5 Other polymers (i)

There are no such residues in this entry.