W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 08:01 AM GMT

PDB ID : 3D2X
Title : Structure of the thiamine pyrophosphate-specific riboswitch bound to oxythi-
amine pyrophosphate
Authors : Thore, S.; Frick, C.; Ban, N.
Deposited on : 2008-05-09
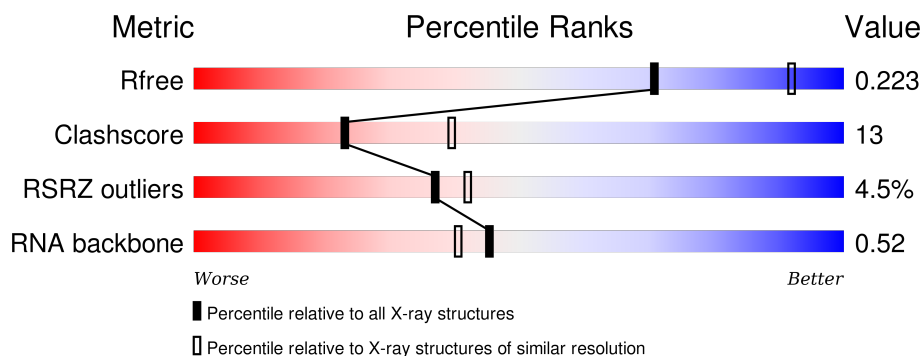Resolution : 2.50 Å(reported)

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 2.50 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		0.223
Clashscore		13
RSRZ outliers		4.5%
RNA backbone		0.52

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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)
RNA backbone	2183	1172 (3.00-2.00)

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	77	51% 32% 16% •
1	B	77	9% / 44% 42% 14%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3503 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 RNA chain called TPP-specific riboswitch.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	77	Total	C	N	O	P	0	0	0
			1656	737	303	539	77			
1	B	77	Total	C	N	O	P	0	0	0
			1656	737	303	539	77			

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	6	Total	Mg	0	0
			6	6		
2	A	7	Total	Mg	0	0
			7	7		

- Molecule 3 is 3-[(4-HYDROXY-2-METHYLPYRIMIDIN-5-YL)METHYL]-5-(2-{[(R)-HYDROXY(PHOSPHONOOXY)PHOSPHORYL]OXY}ETHYL)-4-METHYL-1,3-THIAZOL-3-IUM (three-letter code: D2X) (formula: $C_{12}H_{18}N_3O_8P_2S$).
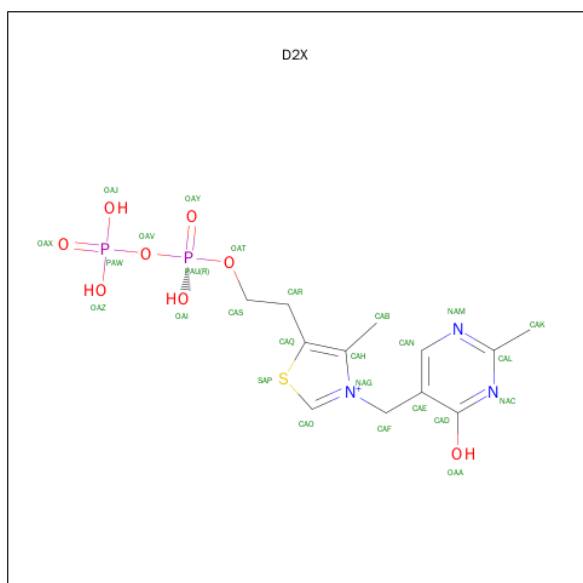
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total 26	C 12	N 3	O 8	P 2	S 1	0	0
3	B	1	Total 26	C 12	N 3	O 8	P 2	S 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	68	Total 68	O 68	0	0
4	B	58	Total 58	O 58	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: TPP-specific riboswitch
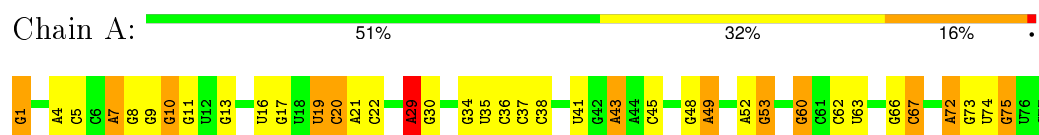
- Molecule 1: TPP-specific riboswitch
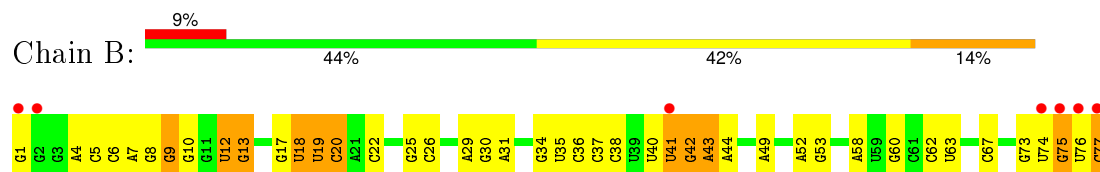
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	75.50Å 111.17Å 55.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.50 14.96 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.8 (15.00-2.50) 97.8 (14.96-2.50)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.27 (at 2.51Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.233 , 0.267 0.236 , 0.223	Depositor DCC
R_{free} test set	825 reflections (5.07%)	DCC
Wilson B-factor (Å^2)	38.1	Xtriage
Anisotropy	0.664	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 27.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.28$	Xtriage
Outliers	0 of 16268 reflections	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3503	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.11% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: D2X, MG

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	0.51	1/1853 (0.1%)	0.76	1/2888 (0.0%)
1	B	0.48	1/1853 (0.1%)	0.70	0/2888
All	All	0.50	2/3706 (0.1%)	0.73	1/5776 (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	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	G	OP3-P	-7.35	1.52	1.61
1	A	1	G	OP3-P	-6.96	1.52	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	29	A	C5'-C4'-C3'	-5.52	107.17	116.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	25	G	Sidechain

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	1656	0	832	29	0
1	B	1656	0	832	34	0
2	A	7	0	0	0	0
2	B	6	0	0	0	0
3	A	26	0	14	3	0
3	B	26	0	14	5	0
4	A	68	0	0	3	0
4	B	58	0	0	2	0
All	All	3503	0	1692	67	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 13.

The worst 5 of 67 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:9:G:H2'	1:A:10:G:H5"	1.29	1.07
1:A:9:G:C2'	1:A:10:G:H5"	1.88	1.02
1:A:41:U:H5"	4:A:143:HOH:O	1.73	0.88
1:A:9:G:H2'	1:A:10:G:C5'	2.08	0.82
1:A:66:G:H4'	1:A:67:C:H5'	1.64	0.80

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

There are no protein molecules in this entry.

5.3.2 Protein sidechains ⓘ

There are no protein molecules in this entry.

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	A	76/77 (98%)	14 (18%)	2 (2%)
1	B	76/77 (98%)	16 (21%)	2 (2%)
All	All	152/154 (98%)	30 (19%)	4 (2%)

5 of 30 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	A	7	A
1	A	10	G
1	A	13	G
1	A	20	C
1	A	29	A

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	A	10	G
1	A	19	U
1	B	19	U
1	B	41	U

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 15 ligands modelled in this entry, 13 are monoatomic - leaving 2 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	#$|Z| > 2$
3	D2X	A	85	2	19,27,27	2.02	2 (10%)	27,40,40	1.85	8 (29%)
3	D2X	B	84	2	19,27,27	2.00	2 (10%)	27,40,40	1.91	7 (25%)

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	D2X	A	85	2	-	0/16/17/17	0/2/2/2
3	D2X	B	84	2	-	0/16/17/17	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	85	D2X	CAH-NAG	-6.68	1.33	1.39
3	B	84	D2X	CAH-NAG	-6.51	1.34	1.39
3	A	85	D2X	CAL-NAC	3.27	1.40	1.34
3	B	84	D2X	CAL-NAC	3.46	1.40	1.34

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	85	D2X	PAU-OAV-PAW	-3.71	120.22	132.67
3	B	84	D2X	PAU-OAV-PAW	-3.46	121.06	132.67
3	B	84	D2X	NAM-CAL-NAC	-3.16	119.76	125.60
3	B	84	D2X	CAB-CAH-CAQ	-3.03	122.10	128.90
3	A	85	D2X	NAM-CAL-NAC	-2.96	120.12	125.60

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	85	D2X	3	0
3	B	84	D2X	5	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	77/77 (100%)	-0.04	0	100	100	16, 30, 47, 54	0
1	B	77/77 (100%)	0.23	7 (9%)	11	12	18, 33, 69, 85	0
All	All	154/154 (100%)	0.09	7 (4%)	37	42	16, 32, 57, 85	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1	G	4.4
1	B	74	U	4.3
1	B	75	G	3.1
1	B	41	U	2.5
1	B	2	G	2.4

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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(Å²)	Q<0.9
3	D2X	A	85	26/26	0.95	0.15	-0.30	25,27,29,30	0
3	D2X	B	84	26/26	0.96	0.12	-1.85	22,28,29,30	0
2	MG	A	83	1/1	0.68	0.22	-	40,40,40,40	0
2	MG	B	80	1/1	0.98	0.09	-	20,20,20,20	0
2	MG	B	82	1/1	0.90	0.06	-	34,34,34,34	0
2	MG	A	82	1/1	0.97	0.12	-	19,19,19,19	0
2	MG	A	78	1/1	0.94	0.23	-	13,13,13,13	0
2	MG	A	81	1/1	0.94	0.18	-	33,33,33,33	0
2	MG	B	81	1/1	0.97	0.08	-	18,18,18,18	0
2	MG	B	83	1/1	0.74	0.15	-	45,45,45,45	0
2	MG	A	84	1/1	0.91	0.12	-	38,38,38,38	0
2	MG	A	79	1/1	0.91	0.10	-	21,21,21,21	0
2	MG	B	79	1/1	0.94	0.24	-	19,19,19,19	0
2	MG	A	80	1/1	0.83	0.27	-	24,24,24,24	0
2	MG	B	78	1/1	0.85	0.20	-	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.