WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 04:33 PM EST

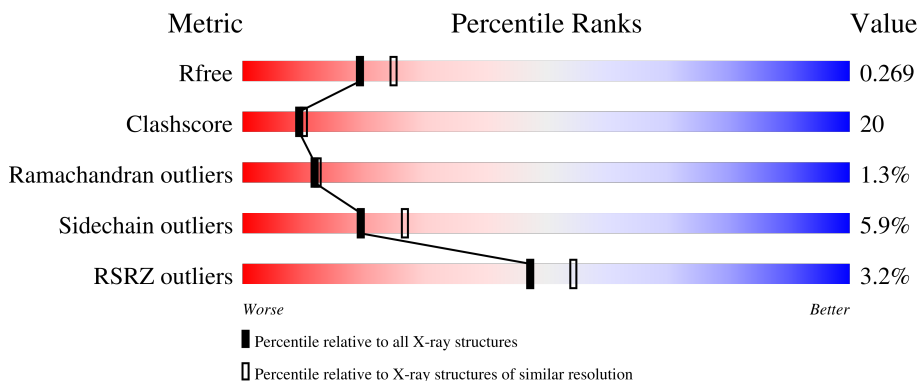|		
PDB ID	:	3EXA
Title	:	Crystal structure of the full-length tRNA isopentenylpyrophosphate transferase (BH2366) from Bacillus halodurans, Northeast Structural Genomics Consortium target BhR41.
Authors	:	Forouhar, F.; Abashidze, M.; Neely, H.; Seetharaman, J.; Shastry, R.; Janjua, H.; Cunningham, K.; Ma, L.-C.; Xiao, R.; Liu, J.; Baran, M.C.; Acton, T.B.; Rost, B.; Montelione, G.T.; Tong, L.; Hunt, J.F.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2008-10-16
Resolution	:	2.30 Å(reported)

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 2.30 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		0.269
Clashscore		20
Ramachandran outliers		1.3%
Sidechain outliers		5.9%
RSRZ outliers		3.2%

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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	322	3% / 54% 36% • 6%
1	B	322	2% / 59% 30% • 6%
1	C	322	2% / 54% 37% • 6%

Continued on next page...

Ideal geometry (DNA, RNA)   :   Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)   :   2.36

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	D	322	4% 61% 30% · 6%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 10249 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 tRNA delta(2)-isopentenylpyrophosphate transferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	Se	0	0	0
			2426	1530	428	457	1	10			
1	B	302	Total	C	N	O	S	Se	0	0	0
			2416	1524	425	456	1	10			
1	C	303	Total	C	N	O	S	Se	0	0	0
			2426	1530	428	457	1	10			
1	D	302	Total	C	N	O	S	Se	0	0	0
			2416	1524	425	456	1	10			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	315	LEU	-	expression tag	UNP Q9KAC3
A	316	GLU	-	expression tag	UNP Q9KAC3
A	317	HIS	-	expression tag	UNP Q9KAC3
A	318	HIS	-	expression tag	UNP Q9KAC3
A	319	HIS	-	expression tag	UNP Q9KAC3
A	320	HIS	-	expression tag	UNP Q9KAC3
A	321	HIS	-	expression tag	UNP Q9KAC3
A	322	HIS	-	expression tag	UNP Q9KAC3
B	315	LEU	-	expression tag	UNP Q9KAC3
B	316	GLU	-	expression tag	UNP Q9KAC3
B	317	HIS	-	expression tag	UNP Q9KAC3
B	318	HIS	-	expression tag	UNP Q9KAC3
B	319	HIS	-	expression tag	UNP Q9KAC3
B	320	HIS	-	expression tag	UNP Q9KAC3
B	321	HIS	-	expression tag	UNP Q9KAC3
B	322	HIS	-	expression tag	UNP Q9KAC3
C	315	LEU	-	expression tag	UNP Q9KAC3
C	316	GLU	-	expression tag	UNP Q9KAC3
C	317	HIS	-	expression tag	UNP Q9KAC3
C	318	HIS	-	expression tag	UNP Q9KAC3
C	319	HIS	-	expression tag	UNP Q9KAC3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	320	HIS	-	expression tag	UNP Q9KAC3
C	321	HIS	-	expression tag	UNP Q9KAC3
C	322	HIS	-	expression tag	UNP Q9KAC3
D	315	LEU	-	expression tag	UNP Q9KAC3
D	316	GLU	-	expression tag	UNP Q9KAC3
D	317	HIS	-	expression tag	UNP Q9KAC3
D	318	HIS	-	expression tag	UNP Q9KAC3
D	319	HIS	-	expression tag	UNP Q9KAC3
D	320	HIS	-	expression tag	UNP Q9KAC3
D	321	HIS	-	expression tag	UNP Q9KAC3
D	322	HIS	-	expression tag	UNP Q9KAC3

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	194	Total 194	O 194	0	0
2	B	132	Total 132	O 132	0	0
2	C	134	Total 134	O 134	0	0
2	D	105	Total 105	O 105	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: tRNA delta(2)-isopentenylpyrophosphate transferase

- Molecule 1: tRNA delta(2)-isopentenylpyrophosphate transferase

- Molecule 1: tRNA delta(2)-isopentenylpyrophosphate transferase

MSE K2 E3 A7 I8 V9 G10 P11 T12 A13 V14 G15 K16 T17 K18 T19 S20 V21 M22 L23 A24 K25 R26 L27 N28 G29 E30 S36 M37 Q38 V39 Y40 M43 D44 I45 K49 I50 T51 E54 H60 H61 L62 I63 K66 D67 P68 V74 A75 T82 P83 L84

I85 T86 E87 I88 H89 E90 R91 G92 R93 L94 P95 G99 G100 T101 G102 L103 Y104 V105 N106 I109 H110 Q111 L114 G115 D116 I117 R118 A119 R124 H125 E128 V131 N132 S133 Q137 A138 L139 L143 S144 K145 I146 A150 A153 I154 H155 P156 I163 R164 A165

L166 E167 I168 I169 T172 G173 LYS THR VAL THR GLU GLN ALA ARG HIS GLU GLU THR PRO S188 N191 L192 V193 M194 I195 T198 M199 E200 L204 R207 I208 R211 M215 K225 K226 L227 Y228 D229 R230 G231 I232 R233 D234 C235 Q236 S237 V238 I241 G242

Y247 L250 N266 R269 Y270 R273 Q274 L275 T276 W277 N280 K281 A282 N283 M289 T290 D291 V292 D293 K296 K297 I298 H302 I305 K308 H318 HIS HIS HIS

• Molecule 1: tRNA delta(2)-isopentenylpyrophosphate transferase

Chain D:

4%

61%   30%   • 6%

MSE K2 E3 K4 P11 T12 A13 V14 G15 K16 T17 K18 T19 S20 V21 M22 L23 A24 K25 R26 L27 E30 S33 M37 M43 I50 E54 V58 P59 H60 H61 L62 I63 D67 P68 V74 L80 A81 T82 P83 T86 E90 R93 T101 G102

L103 Y104 V105 N106 I109 H110 N113 L114 G115 D116 I117 R118 A119 R124 H125 E128 V131 Y134 L139 H140 D141 K142 L143 S144 K145 I146 D147 P148 K149 A153 I154 H155 P156 Y159 R160 R161 V162 I163 R164 A165 L166 I169 T172 G173 LYS THR VAL THR GLU

GLN ALA ARG HIS GLU GLU THR PRO S188 N191 L192 V193 M194 I195 T198 M199 R210 R211 Q214 K226 R230 G231 I232 C235 Q236 S237 V238 Q239 A240 I241 G242 Y247 L263 N266 R269 R273 Q274 L275 F278 A282 N283 V284 M289 T290

D291 V292 D293 F294 I298 H302 I305 H317 HIS HIS HIS HIS HIS

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	51.80Å 73.82Å 95.56Å 89.99° 93.73° 90.02°	Depositor
Resolution (Å)	19.96 – 2.30 28.97 – 2.29	Depositor EDS
% Data completeness (in resolution range)	68.4 (19.96-2.30) 80.9 (28.97-2.29)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	3.90 (at 2.29Å)	Xtriage
Refinement program	CNS 1.2, XTALVIEW	Depositor
R, R_{free}	0.220 , 0.263 0.232 , 0.269	Depositor DCC
R_{free} test set	4065 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	20.2	Xtriage
Anisotropy	0.805	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 23.1	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.458 for -h,k,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10249	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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 38.05 % 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 3.9084e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality (i)

5.1   Standard geometry (i)

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	0.36	0/2457	0.57	1/3294 (0.0%)
1	B	0.37	0/2446	0.83	9/3279 (0.3%)
1	C	0.36	0/2457	0.56	1/3294 (0.0%)
1	D	0.36	0/2446	0.84	9/3279 (0.3%)
All	All	0.36	0/9806	0.71	20/13146 (0.2%)

There are no bond length outliers.

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	164	ARG	NE-CZ-NH1	-14.60	113.00	120.30
1	D	164	ARG	NE-CZ-NH2	-13.97	113.31	120.30
1	D	164	ARG	NE-CZ-NH1	13.61	127.11	120.30
1	D	161	ARG	NE-CZ-NH1	13.61	127.10	120.30
1	D	161	ARG	NE-CZ-NH2	-13.46	113.57	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2426	0	2431	105	0
1	B	2416	0	2424	94	0
1	C	2426	0	2431	106	0
1	D	2416	0	2424	88	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	194	0	0	19	0
2	B	132	0	0	11	0
2	C	134	0	0	20	0
2	D	105	0	0	10	0
All	All	10249	0	9710	384	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 20.

The worst 5 of 384 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:D:273:ARG:HB2	1:D:273:ARG:HH21	1.32	0.95
1:B:273:ARG:HH21	1:B:273:ARG:HB2	1.32	0.94
1:C:102:GLY:HA3	1:C:277:TRP:CD1	2.17	0.80
1:A:102:GLY:HA3	1:A:277:TRP:CD1	2.17	0.79
1:C:51:THR:OG1	1:C:54:GLU:HG3	1.85	0.76

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	299/322 (93%)	280 (94%)	16 (5%)	3 (1%)	15	17
1	B	298/322 (92%)	279 (94%)	14 (5%)	5 (2%)	9	8
1	C	299/322 (93%)	279 (93%)	17 (6%)	3 (1%)	15	17
1	D	298/322 (92%)	280 (94%)	13 (4%)	5 (2%)	9	8
All	All	1194/1288 (93%)	1118 (94%)	60 (5%)	16 (1%)	12	12

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	241	ILE
1	B	241	ILE
1	C	241	ILE
1	D	241	ILE
1	A	291	ASP

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	261/268 (97%)	244 (94%)	17 (6%)	17	23
1	B	260/268 (97%)	248 (95%)	12 (5%)	27	38
1	C	261/268 (97%)	244 (94%)	17 (6%)	17	23
1	D	260/268 (97%)	245 (94%)	15 (6%)	20	27
All	All	1042/1072 (97%)	981 (94%)	61 (6%)	19	27

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	298	ILE
1	D	211	ARG
1	C	137	GLN
1	D	198	THR
1	D	298	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	125	HIS
1	D	60	HIS
1	C	132	ASN
1	C	253	ASN
1	D	78	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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	293/322 (90%)	0.05	11 (3%)	40	47	9, 27, 52, 80	0
1	B	292/322 (90%)	0.09	7 (2%)	59	66	8, 26, 62, 75	0
1	C	293/322 (90%)	0.00	7 (2%)	59	66	10, 27, 53, 74	0
1	D	292/322 (90%)	0.16	13 (4%)	33	40	10, 27, 61, 74	0
All	All	1170/1288 (90%)	0.08	38 (3%)	47	54	8, 27, 58, 80	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	292	VAL	6.3
1	D	173	GLY	4.5
1	A	173	GLY	4.5
1	D	154	ILE	4.3
1	B	134	TYR	4.1

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5   Other polymers ⓘ

There are no such residues in this entry.