



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 11:36 AM GMT

PDB ID : 3PHC  
Title : Crystal Structure of Plasmodium falciparum purine nucleoside phosphorylase  
in complex with DADMe-ImmG  
Authors : Ho, M.; Edwards, A.A.; Almo, S.C.; Schramm, V.L.  
Deposited on : 2010-11-03  
Resolution : 2.00 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 (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

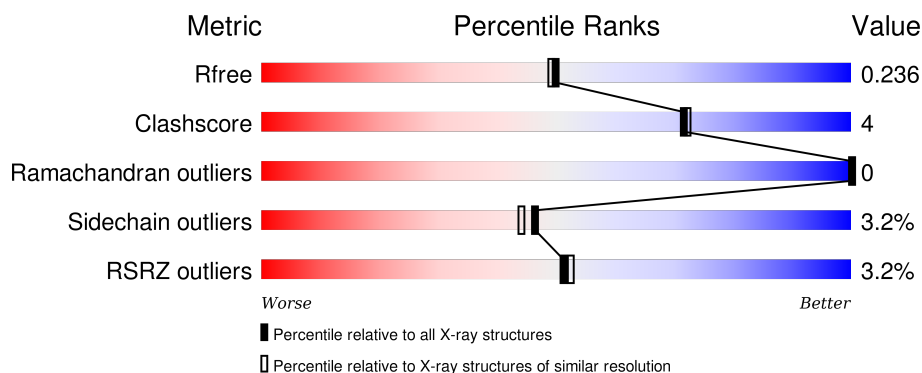
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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  $\geq 3$ , 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	275	<div> <div>3%</div> <div>79%</div> <div>7%</div> <div>12%</div> </div>
1	B	275	<div> <div>3%</div> <div>73%</div> <div>14%</div> <div>12%</div> </div>
1	C	275	<div> <div>3%</div> <div>79%</div> <div>10%</div> <div>12%</div> </div>
1	D	275	<div> <div>2%</div> <div>80%</div> <div>8%</div> <div>12%</div> </div>
1	E	275	<div> <div>0%</div> <div>79%</div> <div>8%</div> <div>12%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	275	

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	PO4	E	502	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11631 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 Purine nucleoside phosphorylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			
1	B	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			
1	C	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			
1	D	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			
1	E	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			
1	F	243	Total	C	N	O	S	0	0	0
			1861	1179	319	347	16			

There are 180 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
A	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
A	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
A	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
A	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
A	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
A	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
A	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
A	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
A	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
A	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7
A	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
A	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
A	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
A	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
A	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
A	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
A	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
A	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
A	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
A	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
A	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
A	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
A	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
A	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
A	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
A	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
A	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
A	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
B	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
B	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
B	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
B	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
B	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
B	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
B	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
B	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
B	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
B	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
B	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7
B	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
B	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
B	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
B	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
B	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
B	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
B	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
B	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
B	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
B	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
B	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
B	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
B	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
B	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
B	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
B	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
B	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
B	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
C	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
C	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
C	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
C	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
C	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
C	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
C	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
C	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
C	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
C	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7
C	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
C	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
C	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
C	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
C	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
C	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
C	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
C	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
C	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
C	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
C	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
C	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
C	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
C	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
C	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
D	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
D	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
D	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
D	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
D	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
D	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
D	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
D	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
D	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
D	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
D	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
D	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
D	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
D	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
D	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
D	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
D	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
D	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
D	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
D	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
D	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
D	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
D	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
D	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
E	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
E	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
E	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
E	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
E	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
E	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
E	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
E	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
E	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
E	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7
E	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
E	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
E	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
E	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
E	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
E	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
E	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
E	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
E	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
E	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
E	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
E	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
E	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
E	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	1	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
F	246	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
F	247	GLY	-	EXPRESSION TAG	UNP Q8T9Z7
F	248	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
F	249	PHE	-	EXPRESSION TAG	UNP Q8T9Z7
F	250	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
F	251	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
F	252	TYR	-	EXPRESSION TAG	UNP Q8T9Z7
F	253	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
F	254	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
F	255	GLN	-	EXPRESSION TAG	UNP Q8T9Z7
F	256	LYS	-	EXPRESSION TAG	UNP Q8T9Z7
F	257	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
F	258	ILE	-	EXPRESSION TAG	UNP Q8T9Z7
F	259	SER	-	EXPRESSION TAG	UNP Q8T9Z7
F	260	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
F	261	GLU	-	EXPRESSION TAG	UNP Q8T9Z7
F	262	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
F	263	LEU	-	EXPRESSION TAG	UNP Q8T9Z7
F	264	ASN	-	EXPRESSION TAG	UNP Q8T9Z7
F	265	SER	-	EXPRESSION TAG	UNP Q8T9Z7
F	266	ALA	-	EXPRESSION TAG	UNP Q8T9Z7
F	267	VAL	-	EXPRESSION TAG	UNP Q8T9Z7
F	268	ASP	-	EXPRESSION TAG	UNP Q8T9Z7
F	269	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	270	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	271	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	272	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	273	HIS	-	EXPRESSION TAG	UNP Q8T9Z7
F	274	HIS	-	EXPRESSION TAG	UNP Q8T9Z7

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

*Continued on next page...*

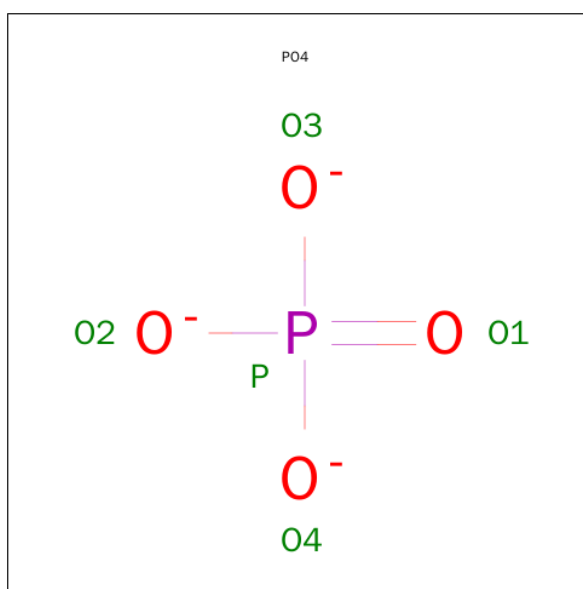


Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
-----	-------	----------	-------	---------	---------

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total K 1 1	0	0
2	C	1	Total K 1 1	0	0
2	E	1	Total K 1 1	0	0

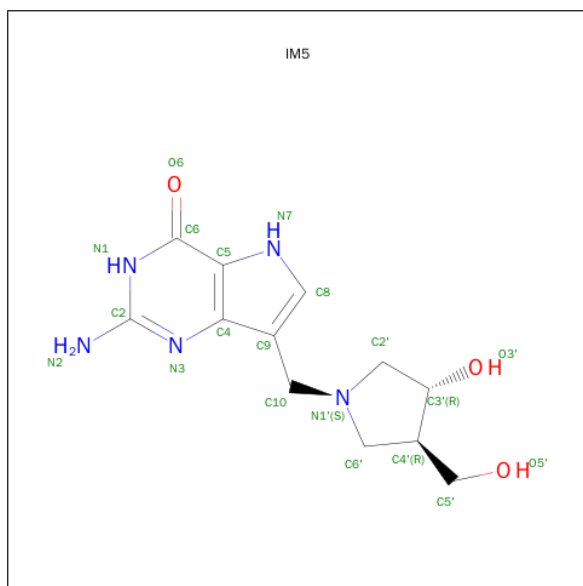
- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O P 5 4 1	0	0
3	B	1	Total O P 5 4 1	0	0
3	C	1	Total O P 5 4 1	0	0
3	D	1	Total O P 5 4 1	0	0
3	E	1	Total O P 5 4 1	0	0
3	F	1	Total O P 5 4 1	0	0

- Molecule 4 is 2-AMINO-7-([(3R,4R)-3-HYDROXY-4-(HYDROXYMETHYL)PYRROLIDINE-1-YL]METHYL)-3,5-DIHYDRO-4H-PYRROLO[3,2-D]PYRIMIDIN-4-ONE (three-letter

code: IM5) (formula: C<sub>12</sub>H<sub>17</sub>N<sub>5</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			20	12	5	3		
4	B	1	Total	C	N	O	0	0
			20	12	5	3		
4	C	1	Total	C	N	O	0	0
			20	12	5	3		
4	D	1	Total	C	N	O	0	0
			20	12	5	3		
4	E	1	Total	C	N	O	0	0
			20	12	5	3		
4	F	1	Total	C	N	O	0	0
			20	12	5	3		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	38	Total	O	0	0
			38	38		
5	B	48	Total	O	0	0
			48	48		
5	C	47	Total	O	0	0
			47	47		
5	D	59	Total	O	0	0
			59	59		
5	E	68	Total	O	0	0
			68	68		

*Continued on next page...*

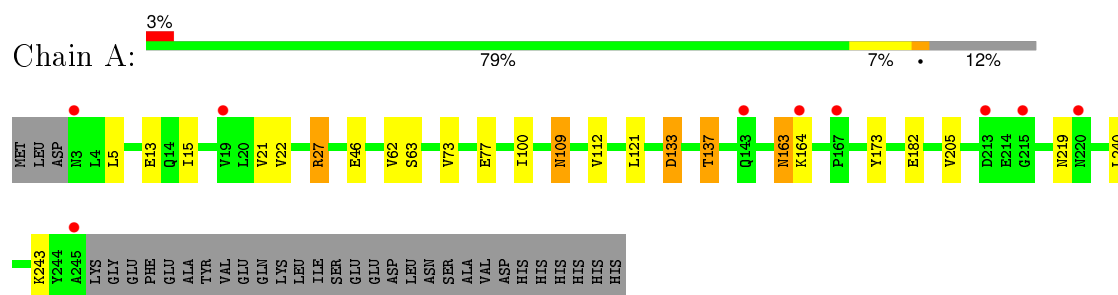
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	52	Total	O	0	0
			52	52		

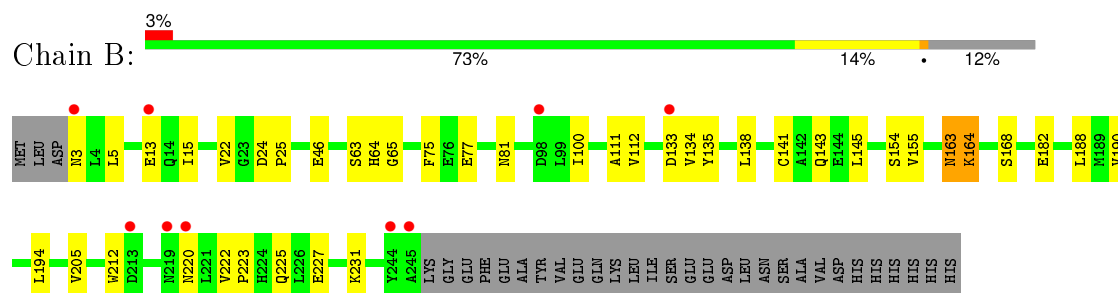
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

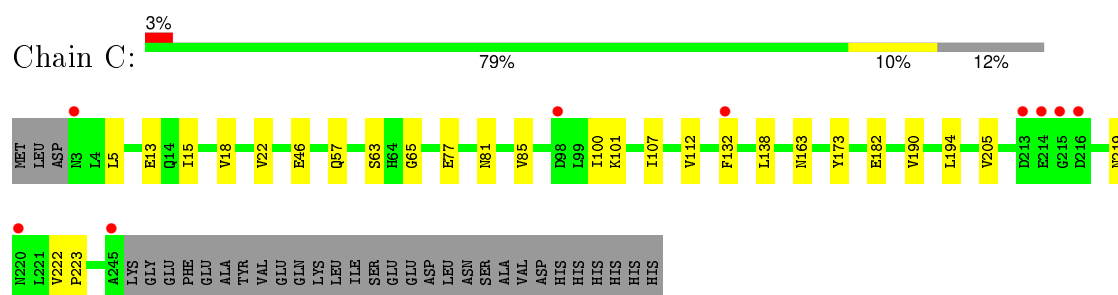
#### • Molecule 1: Purine nucleoside phosphorylase



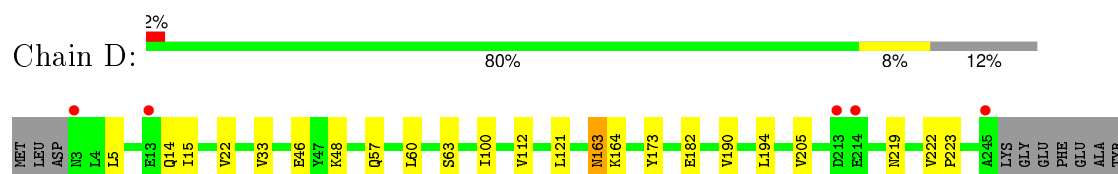
#### • Molecule 1: Purine nucleoside phosphorylase



#### • Molecule 1: Purine nucleoside phosphorylase

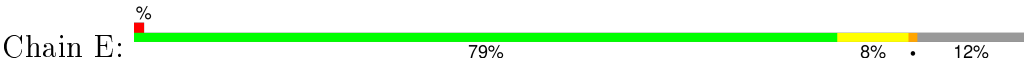


#### • Molecule 1: Purine nucleoside phosphorylase



VAL
GLU
GLN
LYS
LEU
ILE
SER
GLU
GLU
ASP
LEU
ASN
SER
ALA
VAL
ASP
HIS
HIS
HIS
HIS
HIS

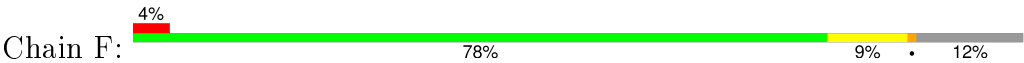
• Molecule 1: Purine nucleoside phosphorylase



MET	ASP	N3	L4	L5	I16	V22	R27	V33	E46	Y47	K48	Q57	R58	V62	S63	E77	N81	G90	R102	L121	M163	K164	P167	S168	E182	D216	V222	P223	A238	T242	A245	LYS	GLY	GLU	PHE	GLU	ALA	TYR
-----	-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----

VAL
GLU
GLN
LYS
LEU
ILE
SER
GLU
GLU
ASP
LEU
ASN
SER
ALA
VAL
ASP
HIS
HIS
HIS
HIS

• Molecule 1: Purine nucleoside phosphorylase



MET	ASP	N3	L4	L5	K12	E13	Q14	I15	V22	L41	E46	Q57	C61	V62	S63	H64	G65	D98	L99	I100	V112	V117	L121	F132	D133	S157	N163	S168	Y173	E182	M183	V205	C208	P209	D213	E214	G215	D216
-----	-----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

P220	L221	P222	P223	A241	A245	LYS	GLY	GLU	PHE	GLU	ALA	TYR	VAL	GLN	GLN	LYS	LEU	ILE	SER	GLU	GLU	ASP	LEU	ASN	SER	ALA	VAL	ASP	HIS	HIS	HIS	HIS	HIS	HIS
------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.20Å 77.37Å 92.20Å 67.71° 73.62° 86.03°	Depositor
Resolution (Å)	20.00 – 2.00 19.92 – 1.99	Depositor EDS
% Data completeness (in resolution range)	93.8 (20.00-2.00) 84.2 (19.92-1.99)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.46 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.198 , 0.235 0.198 , 0.236	Depositor DCC
$R_{free}$ test set	4769 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	27.8	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 38.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 95774 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	11631	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.36% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.375 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: K, IM5, PO4

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	0.49	0/1893	0.59	1/2561 (0.0%)
1	B	0.51	0/1893	0.59	0/2561
1	C	0.51	0/1893	0.59	0/2561
1	D	0.53	0/1893	0.61	0/2561
1	E	0.54	0/1893	0.62	0/2561
1	F	0.49	0/1893	0.62	0/2561
All	All	0.51	0/11358	0.60	1/15366 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	27	ARG	NE-CZ-NH2	-5.06	117.77	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	1861	0	1882	18	0
1	B	1861	0	1882	24	0
1	C	1861	0	1882	15	0
1	D	1861	0	1882	14	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	1861	0	1882	17	0
1	F	1861	0	1882	16	0
2	A	1	0	0	0	0
2	C	1	0	0	0	0
2	E	1	0	0	0	0
3	A	5	0	0	1	0
3	B	5	0	0	0	0
3	C	5	0	0	0	0
3	D	5	0	0	0	0
3	E	5	0	0	2	0
3	F	5	0	0	0	0
4	A	20	0	17	0	0
4	B	20	0	17	1	0
4	C	20	0	17	0	0
4	D	20	0	17	0	0
4	E	20	0	17	0	0
4	F	20	0	17	0	0
5	A	38	0	0	2	0
5	B	48	0	0	0	0
5	C	47	0	0	2	0
5	D	59	0	0	1	0
5	E	68	0	0	2	0
5	F	52	0	0	0	0
All	All	11631	0	11394	97	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.

The worst 5 of 97 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:163:ASN:HD22	1:A:164:LYS:H	1.19	0.88
1:D:163:ASN:HD22	1:D:164:LYS:H	1.24	0.85
1:A:46:GLU:HB3	1:B:46:GLU:HB3	1.62	0.79
1:C:46:GLU:HB3	1:D:46:GLU:HB3	1.64	0.77
1:A:109:ASN:HD22	1:A:109:ASN:H	1.35	0.74

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	241/275 (88%)	233 (97%)	8 (3%)	0	100	100
1	B	241/275 (88%)	232 (96%)	9 (4%)	0	100	100
1	C	241/275 (88%)	234 (97%)	7 (3%)	0	100	100
1	D	241/275 (88%)	236 (98%)	5 (2%)	0	100	100
1	E	241/275 (88%)	235 (98%)	6 (2%)	0	100	100
1	F	241/275 (88%)	234 (97%)	7 (3%)	0	100	100
All	All	1446/1650 (88%)	1404 (97%)	42 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	206/235 (88%)	199 (97%)	7 (3%)	44	41
1	B	206/235 (88%)	197 (96%)	9 (4%)	35	30
1	C	206/235 (88%)	202 (98%)	4 (2%)	65	67
1	D	206/235 (88%)	201 (98%)	5 (2%)	57	58
1	E	206/235 (88%)	201 (98%)	5 (2%)	57	58
1	F	206/235 (88%)	196 (95%)	10 (5%)	31	25
All	All	1236/1410 (88%)	1196 (97%)	40 (3%)	46	44

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	182	GLU
1	D	163	ASN
1	F	121	LEU
1	D	33	VAL
1	D	182	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	57	GLN
1	D	53	HIS
1	F	53	HIS
1	C	219	ASN
1	D	57	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](#)

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, 3 are monoatomic - leaving 12 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
4	IM5	A	501	-	20,22,22	0.99	1 (5%)	17,32,32	2.61	5 (29%)
3	PO4	A	502	-	4,4,4	0.33	0	6,6,6	0.28	0
4	IM5	B	501	-	20,22,22	0.94	1 (5%)	17,32,32	2.64	7 (41%)
3	PO4	B	502	-	4,4,4	0.42	0	6,6,6	0.27	0
4	IM5	C	501	-	20,22,22	1.00	1 (5%)	17,32,32	2.58	8 (47%)
3	PO4	C	502	-	4,4,4	0.40	0	6,6,6	0.27	0
4	IM5	D	501	-	20,22,22	0.96	1 (5%)	17,32,32	2.61	6 (35%)
3	PO4	D	502	-	4,4,4	0.32	0	6,6,6	0.29	0
4	IM5	E	501	-	20,22,22	0.92	1 (5%)	17,32,32	2.53	5 (29%)
3	PO4	E	502	-	4,4,4	0.39	0	6,6,6	0.28	0
4	IM5	F	501	-	20,22,22	1.03	1 (5%)	17,32,32	2.58	7 (41%)
3	PO4	F	502	-	4,4,4	0.31	0	6,6,6	0.27	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
4	IM5	A	501	-	-	0/5/18/18	0/3/3/3
3	PO4	A	502	-	-	0/0/0/0	0/0/0/0
4	IM5	B	501	-	-	0/5/18/18	0/3/3/3
3	PO4	B	502	-	-	0/0/0/0	0/0/0/0
4	IM5	C	501	-	-	0/5/18/18	0/3/3/3
3	PO4	C	502	-	-	0/0/0/0	0/0/0/0
4	IM5	D	501	-	-	0/5/18/18	0/3/3/3
3	PO4	D	502	-	-	0/0/0/0	0/0/0/0
4	IM5	E	501	-	-	0/5/18/18	0/3/3/3
3	PO4	E	502	-	-	0/0/0/0	0/0/0/0
4	IM5	F	501	-	-	0/5/18/18	0/3/3/3
3	PO4	F	502	-	-	0/0/0/0	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	501	IM5	C6-C5	3.12	1.47	1.41
4	B	501	IM5	C6-C5	3.24	1.47	1.41
4	D	501	IM5	C6-C5	3.28	1.47	1.41
4	C	501	IM5	C6-C5	3.31	1.47	1.41
4	A	501	IM5	C6-C5	3.42	1.48	1.41

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	501	IM5	C6-C5-C4	-4.79	115.18	120.90
4	F	501	IM5	C5-C6-N1	-4.78	117.06	123.59
4	D	501	IM5	C5-C6-N1	-4.58	117.33	123.59
4	C	501	IM5	C5-C6-N1	-4.56	117.35	123.59
4	A	501	IM5	C5-C6-N1	-4.48	117.47	123.59

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
3	A	502	PO4	1	0
4	B	501	IM5	1	0
3	E	502	PO4	2	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<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	243/275 (88%)	0.06	9 (3%) 45 47	20, 31, 44, 50	0
1	B	243/275 (88%)	0.11	9 (3%) 45 47	19, 32, 46, 53	0
1	C	243/275 (88%)	0.04	9 (3%) 45 47	20, 29, 44, 52	0
1	D	243/275 (88%)	-0.11	5 (2%) 67 67	18, 26, 38, 46	0
1	E	243/275 (88%)	-0.15	4 (1%) 74 75	18, 25, 35, 44	0
1	F	243/275 (88%)	0.00	10 (4%) 41 42	19, 28, 42, 47	0
All	All	1458/1650 (88%)	-0.01	46 (3%) 51 52	18, 28, 43, 53	0

The worst 5 of 46 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	245	ALA	7.1
1	A	3	ASN	6.3
1	B	245	ALA	5.1
1	F	245	ALA	4.8
1	E	245	ALA	4.6

### 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 [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, 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
4	IM5	E	501	20/20	0.94	0.12	0.92	20,22,25,28	0
4	IM5	F	501	20/20	0.92	0.12	0.59	25,28,29,30	0
4	IM5	C	501	20/20	0.92	0.12	0.42	26,28,31,32	0
4	IM5	D	501	20/20	0.93	0.12	0.36	25,28,31,32	0
4	IM5	A	501	20/20	0.92	0.10	0.14	27,31,32,32	0
4	IM5	B	501	20/20	0.92	0.10	-0.35	29,32,35,37	0
3	PO4	C	502	5/5	0.98	0.08	-0.64	30,31,33,33	0
3	PO4	A	502	5/5	0.97	0.08	-0.90	33,34,35,36	0
3	PO4	D	502	5/5	0.98	0.07	-1.83	27,28,29,29	0
3	PO4	E	502	5/5	0.98	0.07	-2.07	22,22,23,24	0
3	PO4	F	502	5/5	0.98	0.07	-2.30	28,29,31,32	0
2	K	E	275	1/1	0.99	0.04	-2.66	24,24,24,24	0
3	PO4	B	502	5/5	0.98	0.06	-2.90	37,38,39,39	0
2	K	A	275	1/1	0.99	0.03	-3.40	29,29,29,29	0
2	K	C	275	1/1	0.99	0.03	-3.81	28,28,28,28	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.