



# Full wwPDB X-ray Structure Validation Report

Mar 1, 2014 – 03:54 AM GMT

PDB ID : 2A0Y  
Title : Structure of human purine nucleoside phosphorylase H257D mutant  
Authors : Murkin, A.S.; Shi, W.; Schramm, V.L.  
Deposited on : 2005-06-17  
Resolution : 2.28 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

---

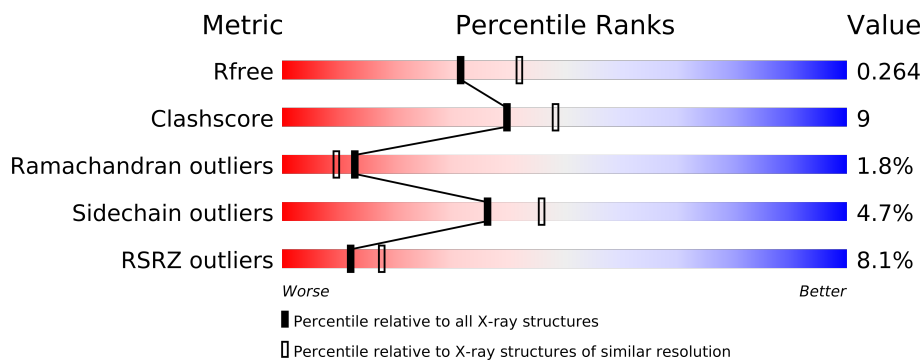
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.28 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))
$R_{free}$	66092	3861 (2.30-2.26)
Clashscore	79885	4801 (2.30-2.26)
Ramachandran outliers	78287	4729 (2.30-2.26)
Sidechain outliers	78261	4728 (2.30-2.26)
RSRZ outliers	66119	3864 (2.30-2.26)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	289	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 2279 atoms, of which 0 are hydrogen and 0 are deuterium.

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
			Total	C	N	O	S			
1	A	282	2204	1401	385	403	15	0	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	257	ASP	HIS	ENGINEERED	UNP P00491

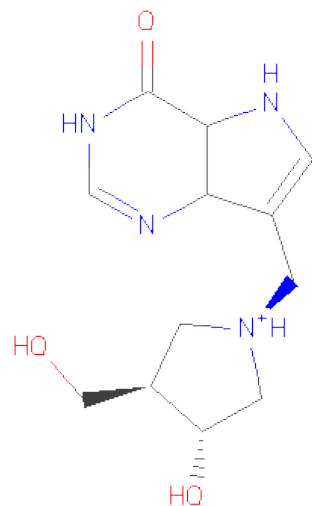
- Molecule 2 is SULFATE ION (three-letter code: SO<sub>4</sub>) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
2	A	1	5	4	1	0	0
2	A	1	5	4	1	0	0

- Molecule 3 is 3-HYDROXY-4-HYDROXYMETHYL-1-(4-OXO-4,4A,5,7A-TETRAHYDRO

-3H-PYRROLO[3,2-D]PYRIMIDIN-7-YLMETHYL)-PYRROLIDINIUM (three-letter code: DIH) (formula: C<sub>12</sub>H<sub>19</sub>N<sub>4</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	19	12	4	3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	46	46	46	0	0

i

- Molecule 1: Purine nucleoside phosphorylase

P283	D134	MT
L284	H135	GLU
PRO	I136	N3
ASP	N137	G4
LYS	D152	Y5
ALA	E153	E8
SER	R154	
	R171	G32
	A174	S33
	L175	G34
	K179	L35
	Q180	L38
	M181	L42
	G182	T43
	E183	Q44
	Q184	D49
	R185	E52
	L187	F56
	Q188	P57
	E201	R58
	C206	S59
	V222	T60
	C231	V61
G232	E250	P62
L233	S251	G63
	L252	H64
	E253	A65
	E259	L73
	A262	N74
	K265	G75
	A268	R76
	Q269	V79
	K270	R84
	L271	P92
	E272	K95
	Q273	F98
	F274	P99
	L278	N115
		G119
		L120
		I129
		M130
		L131
		I132
		R133

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	142.47 Å   142.47 Å   166.03 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	50.00 – 2.28 49.51 – 2.28	Depositor EDS
% Data completeness (in resolution range)	95.0 (50.00-2.28) 95.6 (49.51-2.28)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.25 (at 2.29 Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.243 , 0.266 0.242 , 0.264	Depositor DCC
$R_{free}$ test set	2804 reflections (9.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	50.6	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 40.6	EDS
Estimated twinning fraction	0.022 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.015 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.006 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 29445 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	2279	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DIH, SO4

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	0.39	0/2254	0.62	0/3050

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2204	0	2180	42	0
2	A	10	0	0	0	0
3	A	19	0	16	0	0
4	A	46	0	0	1	0
All	All	2279	0	2196	42	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

All (42) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:38:LEU:HD23	1:A:272:GLU:HG2	1.61	0.83
1:A:120:LEU:HD23	1:A:245:VAL:HG21	1.64	0.79
1:A:3:ASN:HD22	1:A:4:GLY:N	1.84	0.74
1:A:3:ASN:ND2	1:A:5:TYR:H	1.85	0.74
1:A:76:ARG:N	1:A:76:ARG:HD2	2.02	0.72
1:A:34:GLY:O	1:A:35:LEU:HD23	1.92	0.68
1:A:76:ARG:HG2	1:A:76:ARG:HH11	1.59	0.67
1:A:35:LEU:O	1:A:38:LEU:HG	1.99	0.63
1:A:49:ASP:O	1:A:52:GLU:HG2	1.99	0.63
1:A:119:GLY:O	1:A:245:VAL:HG22	2.00	0.61
1:A:35:LEU:HD22	1:A:271:LEU:HD22	1.85	0.58
1:A:269:GLN:O	1:A:273:GLN:HG3	2.06	0.55
1:A:76:ARG:CG	1:A:76:ARG:HH11	2.20	0.54
1:A:129:ILE:HD11	1:A:271:LEU:HD11	1.89	0.54
1:A:131:LEU:HD13	1:A:187:LEU:HD21	1.90	0.52
1:A:42:LEU:HD23	1:A:73:LEU:HB2	1.93	0.51
1:A:274:PHE:O	1:A:278:LEU:HB2	2.14	0.48
1:A:3:ASN:HD22	1:A:3:ASN:C	2.13	0.48
1:A:8:GLU:HG3	4:A:516:HOH:O	2.14	0.47
1:A:188:GLN:HE21	1:A:188:GLN:HA	1.79	0.47
1:A:42:LEU:CD2	1:A:73:LEU:HB2	2.46	0.46
1:A:79:VAL:HG13	1:A:79:VAL:O	2.15	0.46
1:A:262:ALA:O	1:A:265:LYS:HB3	2.15	0.46
1:A:231:CYS:HB3	1:A:233:LEU:HD23	1.97	0.46
1:A:76:ARG:NH1	1:A:76:ARG:CG	2.78	0.46
1:A:44:GLN:HE21	1:A:44:GLN:HB3	1.60	0.45
1:A:34:GLY:C	1:A:35:LEU:HD23	2.37	0.44
1:A:32:GLY:HA3	1:A:115:ASN:HA	1.99	0.44
1:A:174:ALA:HA	1:A:278:LEU:HD21	1.99	0.43
1:A:133:ARG:CZ	1:A:171:ARG:HH12	2.31	0.43
1:A:135:HIS:CD2	1:A:135:HIS:C	2.91	0.43
1:A:201:GLU:H	1:A:201:GLU:CD	2.22	0.43
1:A:92:PRO:HD2	1:A:95:LYS:HD2	1.99	0.43
1:A:137:ASN:HB2	1:A:222:VAL:HG11	2.01	0.43
1:A:61:VAL:HG12	1:A:64:HIS:CE1	2.54	0.42
1:A:43:THR:O	1:A:44:GLN:HB2	2.19	0.42
1:A:98:PHE:HB3	1:A:99:PRO:HD3	2.02	0.41
1:A:98:PHE:N	1:A:99:PRO:CD	2.84	0.41
1:A:75:GLY:C	1:A:76:ARG:HD2	2.40	0.41
1:A:152:ASP:OD1	1:A:154:ARG:HB2	2.20	0.41
1:A:269:GLN:OE1	1:A:270:LYS:N	2.54	0.40
1:A:206:CYS:SG	1:A:245:VAL:HG23	2.60	0.40

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	280/289 (97%)	265 (95%)	10 (4%)	5 (2%)	13 9

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	SER
1	A	60	THR
1	A	268	ALA
1	A	283	PRO
1	A	182	GLY

### 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	234/240 (98%)	223 (95%)	11 (5%)	36 46

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	ASN
1	A	56	PHE
1	A	58	ARG
1	A	76	ARG
1	A	84	ARG
1	A	135	HIS
1	A	179	LYS
1	A	233	LEU
1	A	259	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	269	GLN
1	A	278	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	3	ASN
1	A	26	GLN
1	A	44	GLN
1	A	55	ASN
1	A	64	HIS
1	A	188	GLN
1	A	273	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

3 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	DIH	A	300	-	21,21,21	4.98	13 (61%)	28,30,30	4.82	14 (50%)
2	SO4	A	401	-	4,4,4	3.41	2 (50%)	6,6,6	0.93	0
2	SO4	A	402	-	4,4,4	3.37	2 (50%)	6,6,6	0.92	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	DIH	A	300	-	2/2/7/11	0/6/41/41	0/1/3/3
2	SO4	A	401	-	-	0/0/0/0	0/0/0/0
2	SO4	A	402	-	-	0/0/0/0	0/0/0/0

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	300	DIH	C5-C4	-18.50	1.41	1.57
3	A	300	DIH	C6-N1	7.53	1.43	1.34
2	A	402	SO4	O3-S	-4.99	1.30	1.47
2	A	401	SO4	O3-S	-4.83	1.30	1.47
2	A	401	SO4	O1-S	4.72	1.62	1.47
2	A	402	SO4	O1-S	4.23	1.60	1.47
3	A	300	DIH	C5-C6	-3.81	1.46	1.53
3	A	300	DIH	C2'-N1'	-3.69	1.45	1.50
3	A	300	DIH	C2-N1	3.65	1.42	1.35
3	A	300	DIH	C6'-N1'	-3.63	1.45	1.50
3	A	300	DIH	C2-N3	3.50	1.38	1.29
3	A	300	DIH	C4-N3	-3.27	1.43	1.46
3	A	300	DIH	C5-N7	-3.05	1.38	1.45
3	A	300	DIH	C4-C9	-2.84	1.42	1.50
3	A	300	DIH	C10-N1'	-2.77	1.45	1.50
3	A	300	DIH	C4'-C3'	-2.34	1.49	1.53
3	A	300	DIH	C6'-C4'	-2.06	1.47	1.51

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	300	DIH	C4-C5-N7	14.86	110.16	100.65
3	A	300	DIH	C9-C10-N1'	13.43	128.23	113.94
3	A	300	DIH	O5'-C5'-C4'	7.88	127.96	111.44

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	300	DIH	C4-C9-C8	-5.74	104.18	110.46
3	A	300	DIH	O3'-C3'-C2'	5.48	120.44	109.81
3	A	300	DIH	C2'-C3'-C4'	-5.06	95.90	102.80
3	A	300	DIH	O6-C6-C5	4.46	126.68	119.47
3	A	300	DIH	C9-C8-N7	3.95	114.79	108.85
3	A	300	DIH	N1-C2-N3	-3.55	121.47	125.48
3	A	300	DIH	C2'-N1'-C6'	-2.72	98.50	105.04
3	A	300	DIH	C6'-C4'-C5'	-2.44	109.56	113.48
3	A	300	DIH	C10-C9-C4	2.33	126.61	122.58
3	A	300	DIH	O6-C6-N1	-2.06	119.05	122.17
3	A	300	DIH	C3'-C2'-N1'	2.04	108.11	104.17

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	300	DIH	C4
3	A	300	DIH	C5

There are no torsion outliers.

There are no ring outliers.

## 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, 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	282/289 (97%)	0.60	23 (8%) 12 17	28, 52, 82, 92	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	62	PRO	6.6
1	A	65	ALA	5.6
1	A	63	GLY	5.5
1	A	181	MET	4.7
1	A	60	THR	4.6
1	A	64	HIS	4.6
1	A	38	LEU	4.2
1	A	59	SER	4.1
1	A	252	LEU	3.3
1	A	274	PHE	3.0
1	A	184	GLN	3.0
1	A	253	GLU	3.0
1	A	265	LYS	2.9
1	A	250	GLU	2.7
1	A	58	ARG	2.7
1	A	179	LYS	2.4
1	A	269	GLN	2.4
1	A	180	GLN	2.3
1	A	175	LEU	2.3
1	A	185	ARG	2.2
1	A	61	VAL	2.2
1	A	236	PHE	2.1
1	A	271	LEU	2.1

## 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 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<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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	SO4	A	401	5/5	0.15	-0.03	57,57,58,58	0
3	DIH	A	300	19/19	0.14	-0.13	40,44,57,58	0
2	SO4	A	402	5/5	0.15	-0.57	42,45,47,48	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.