wwPDB X-ray Structure Validation Summary Report (i)

Jun 11, 2024 – 06:13 PM EDT

|		
PDB ID	:	1VHJ
Title	:	Crystal structure of purine nucleoside phosphorylase
Authors	:	Structural GenomiX
Deposited on	:	2003-12-01
Resolution	:	2.23 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 2.23 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.
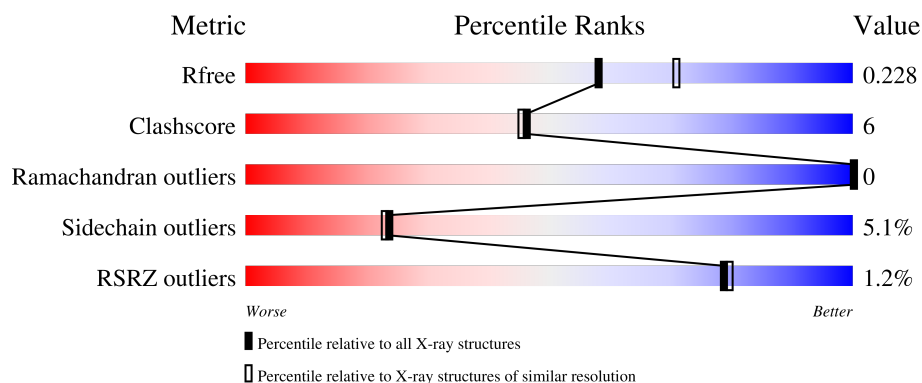
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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	253	77% 15% • 8%
1	B	253	79% 13% • 6%
1	C	253	74% 17% • 8%
1	D	253	76% 15% • 8%
1	E	253	75% 15% • 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	253	79% 14% 6%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 10930 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	234	Total	C	N	O	S	0	0	0
			1752	1107	296	333	16			
1	B	237	Total	C	N	O	S	0	1	0
			1778	1123	303	336	16			
1	C	232	Total	C	N	O	S	0	0	0
			1731	1094	292	329	16			
1	D	232	Total	C	N	O	S	0	0	0
			1727	1092	292	327	16			
1	E	232	Total	C	N	O	S	0	0	0
			1733	1097	292	328	16			
1	F	237	Total	C	N	O	S	0	0	0
			1774	1121	302	335	16			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	cloning artifact	UNP Q9KPM0
A	0	SER	-	cloning artifact	UNP Q9KPM0
A	1	LEU	-	cloning artifact	UNP Q9KPM0
A	57	ARG	LYS	variant	UNP Q9KPM0
A	242	GLU	-	cloning artifact	UNP Q9KPM0
A	243	GLY	-	cloning artifact	UNP Q9KPM0
A	244	GLY	-	cloning artifact	UNP Q9KPM0
A	245	SER	-	cloning artifact	UNP Q9KPM0
A	246	HIS	-	cloning artifact	UNP Q9KPM0
A	247	HIS	-	cloning artifact	UNP Q9KPM0
A	248	HIS	-	cloning artifact	UNP Q9KPM0
A	249	HIS	-	cloning artifact	UNP Q9KPM0
A	250	HIS	-	cloning artifact	UNP Q9KPM0
A	251	HIS	-	cloning artifact	UNP Q9KPM0
B	-1	MET	-	cloning artifact	UNP Q9KPM0
B	0	SER	-	cloning artifact	UNP Q9KPM0
B	1	LEU	-	cloning artifact	UNP Q9KPM0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	57	ARG	LYS	variant	UNP Q9KPM0
B	242	GLU	-	cloning artifact	UNP Q9KPM0
B	243	GLY	-	cloning artifact	UNP Q9KPM0
B	244	GLY	-	cloning artifact	UNP Q9KPM0
B	245	SER	-	cloning artifact	UNP Q9KPM0
B	246	HIS	-	cloning artifact	UNP Q9KPM0
B	247	HIS	-	cloning artifact	UNP Q9KPM0
B	248	HIS	-	cloning artifact	UNP Q9KPM0
B	249	HIS	-	cloning artifact	UNP Q9KPM0
B	250	HIS	-	cloning artifact	UNP Q9KPM0
B	251	HIS	-	cloning artifact	UNP Q9KPM0
C	-1	MET	-	cloning artifact	UNP Q9KPM0
C	0	SER	-	cloning artifact	UNP Q9KPM0
C	1	LEU	-	cloning artifact	UNP Q9KPM0
C	57	ARG	LYS	variant	UNP Q9KPM0
C	242	GLU	-	cloning artifact	UNP Q9KPM0
C	243	GLY	-	cloning artifact	UNP Q9KPM0
C	244	GLY	-	cloning artifact	UNP Q9KPM0
C	245	SER	-	cloning artifact	UNP Q9KPM0
C	246	HIS	-	cloning artifact	UNP Q9KPM0
C	247	HIS	-	cloning artifact	UNP Q9KPM0
C	248	HIS	-	cloning artifact	UNP Q9KPM0
C	249	HIS	-	cloning artifact	UNP Q9KPM0
C	250	HIS	-	cloning artifact	UNP Q9KPM0
C	251	HIS	-	cloning artifact	UNP Q9KPM0
D	-1	MET	-	cloning artifact	UNP Q9KPM0
D	0	SER	-	cloning artifact	UNP Q9KPM0
D	1	LEU	-	cloning artifact	UNP Q9KPM0
D	57	ARG	LYS	variant	UNP Q9KPM0
D	242	GLU	-	cloning artifact	UNP Q9KPM0
D	243	GLY	-	cloning artifact	UNP Q9KPM0
D	244	GLY	-	cloning artifact	UNP Q9KPM0
D	245	SER	-	cloning artifact	UNP Q9KPM0
D	246	HIS	-	cloning artifact	UNP Q9KPM0
D	247	HIS	-	cloning artifact	UNP Q9KPM0
D	248	HIS	-	cloning artifact	UNP Q9KPM0
D	249	HIS	-	cloning artifact	UNP Q9KPM0
D	250	HIS	-	cloning artifact	UNP Q9KPM0
D	251	HIS	-	cloning artifact	UNP Q9KPM0
E	-1	MET	-	cloning artifact	UNP Q9KPM0
E	0	SER	-	cloning artifact	UNP Q9KPM0
E	1	LEU	-	cloning artifact	UNP Q9KPM0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	57	ARG	LYS	variant	UNP Q9KPM0
E	242	GLU	-	cloning artifact	UNP Q9KPM0
E	243	GLY	-	cloning artifact	UNP Q9KPM0
E	244	GLY	-	cloning artifact	UNP Q9KPM0
E	245	SER	-	cloning artifact	UNP Q9KPM0
E	246	HIS	-	cloning artifact	UNP Q9KPM0
E	247	HIS	-	cloning artifact	UNP Q9KPM0
E	248	HIS	-	cloning artifact	UNP Q9KPM0
E	249	HIS	-	cloning artifact	UNP Q9KPM0
E	250	HIS	-	cloning artifact	UNP Q9KPM0
E	251	HIS	-	cloning artifact	UNP Q9KPM0
F	-1	MET	-	cloning artifact	UNP Q9KPM0
F	0	SER	-	cloning artifact	UNP Q9KPM0
F	1	LEU	-	cloning artifact	UNP Q9KPM0
F	57	ARG	LYS	variant	UNP Q9KPM0
F	242	GLU	-	cloning artifact	UNP Q9KPM0
F	243	GLY	-	cloning artifact	UNP Q9KPM0
F	244	GLY	-	cloning artifact	UNP Q9KPM0
F	245	SER	-	cloning artifact	UNP Q9KPM0
F	246	HIS	-	cloning artifact	UNP Q9KPM0
F	247	HIS	-	cloning artifact	UNP Q9KPM0
F	248	HIS	-	cloning artifact	UNP Q9KPM0
F	249	HIS	-	cloning artifact	UNP Q9KPM0
F	250	HIS	-	cloning artifact	UNP Q9KPM0
F	251	HIS	-	cloning artifact	UNP Q9KPM0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	84	Total 84	O 84	0	0
2	B	55	Total 55	O 55	0	0
2	C	76	Total 76	O 76	0	0
2	D	77	Total 77	O 77	0	0
2	E	85	Total 85	O 85	0	0
2	F	58	Total 58	O 58	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: purine nucleoside phosphorylase

Chain A:

- Molecule 1: purine nucleoside phosphorylase

Chain B:

- Molecule 1: purine nucleoside phosphorylase

Chain C:

- Molecule 1: purine nucleoside phosphorylase

Chain D:

MET SER LEU A2 N7 A8 Q9 L18 M19 P20 G21 D22 R25 A26 K27 E31 V38 V43 R44 N45 R56 R57 I58 I67 P68 S69 C70 V74 K79 D80 V89 R102 K136 E140 A141 T162 P163 V179 E180 I186 V189

H206 I207 K208 T209 E211 GLN THR THR SER GLU E217 T221 F222 N223 E227 V233 L234 I235 Q238 ALA GLY TYR GLU GLY GLY SER HIS HIS HIS HIS HIS HIS

• Molecule 1: purine nucleoside phosphorylase

Chain E:

%

75%　　15%　•　8%

MET SER LEU A2 T3 P4 V16 P20 G21 D22 P23 L24 R25 G48 T52 M61 G62 H63 C70 V74 T75 K79 D80 V89 V101 R102 I106 G107 M108 K115 K122 K136 E140 R145 G146 I147 M167 V179 E180 M181 E182

I186 V189 E192 L198 E211 GLN THR THR SER GLU E217 N220 T221 F222 N223 E224 I228 V233 Q238 ALA GLY TYR GLU GLY GLY SER HIS HIS HIS HIS HIS HIS

• Molecule 1: purine nucleoside phosphorylase

Chain F:

%

79%　　14%　•　6%

MET SER LEU A2 N7 V16 M19 R25 A26 K27 E31 D35 V43 R44 N45 T52 Y53 R57 I58 M61 G62 H63 I67 P68 S69 C70 V74 T75 K79 R88 V101 R102 M108 K122 K136 E140 I147 D148 V179

E180 M181 E182 G185 V189 E211 Q212 T213 Q219 I228 Q238 ALA GLY TYR GLU GLY GLY SER HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	46.70Å 153.25Å 96.76Å 90.00° 103.91° 90.00°	Depositor
Resolution (Å)	31.31 – 2.23 31.31 – 2.23	Depositor EDS
% Data completeness (in resolution range)	(Not available) (31.31-2.23) 96.9 (31.31-2.23)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.88 (at 2.22Å)	Xtriage
Refinement program	REFMAC 4.0	Depositor
R, R_{free}	0.196 , 0.249 0.183 , 0.228	Depositor DCC
R_{free} test set	3165 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	22.6	Xtriage
Anisotropy	0.599	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 24.8	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.478 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10930	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.07% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.55	0/1779	0.94	2/2400 (0.1%)
1	B	0.56	0/1812	1.08	5/2444 (0.2%)
1	C	0.55	0/1758	0.96	3/2374 (0.1%)
1	D	0.56	0/1754	1.04	3/2369 (0.1%)
1	E	0.56	0/1760	0.94	3/2375 (0.1%)
1	F	0.53	0/1802	0.87	0/2432
All	All	0.55	0/10665	0.97	16/14394 (0.1%)

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	218	ARG	CD-NE-CZ	20.64	152.50	123.60
1	D	25	ARG	CD-NE-CZ	19.65	151.11	123.60
1	B	218	ARG	NE-CZ-NH2	18.63	129.61	120.30
1	C	25	ARG	CD-NE-CZ	10.56	138.39	123.60
1	D	25	ARG	NE-CZ-NH2	9.58	125.09	120.30

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	1752	0	1726	25	0
1	B	1778	0	1765	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1731	0	1698	26	0
1	D	1727	0	1694	22	0
1	E	1733	0	1707	25	0
1	F	1774	0	1751	19	0
2	A	84	0	0	3	0
2	B	55	0	0	2	0
2	C	76	0	0	3	0
2	D	77	0	0	0	0
2	E	85	0	0	2	0
2	F	58	0	0	0	0
All	All	10930	0	10341	132	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 6.

The worst 5 of 132 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:213:THR:OG1	1:B:218:ARG:HD3	1.80	0.81
1:A:35:ASP:HB2	1:A:52:THR:HG22	1.61	0.80
1:B:213:THR:O	1:B:218:ARG:NH2	2.16	0.78
1:B:214:THR:O	1:B:218:ARG:HG3	1.84	0.77
1:B:52:THR:HG23	2:B:304:HOH:O	1.89	0.73

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	230/253 (91%)	225 (98%)	5 (2%)	0	100	100
1	B	236/253 (93%)	232 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	228/253 (90%)	222 (97%)	6 (3%)	0	100	100
1	D	228/253 (90%)	223 (98%)	5 (2%)	0	100	100
1	E	228/253 (90%)	221 (97%)	7 (3%)	0	100	100
1	F	235/253 (93%)	230 (98%)	5 (2%)	0	100	100
All	All	1385/1518 (91%)	1353 (98%)	32 (2%)	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	180/201 (90%)	174 (97%)	6 (3%)	38	43
1	B	184/201 (92%)	170 (92%)	14 (8%)	13	9
1	C	177/201 (88%)	167 (94%)	10 (6%)	21	19
1	D	176/201 (88%)	168 (96%)	8 (4%)	27	29
1	E	177/201 (88%)	171 (97%)	6 (3%)	37	42
1	F	181/201 (90%)	170 (94%)	11 (6%)	18	16
All	All	1075/1206 (89%)	1020 (95%)	55 (5%)	24	23

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	179	VAL
1	D	180	GLU
1	F	219	GLN
1	F	88	ARG
1	C	180	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	39	GLN
1	F	220	ASN
1	D	220	ASN
1	F	223	ASN
1	F	9	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(Å²)	Q<0.9
1	A	234/253 (92%)	-0.38	2 (0%)	84	84	15, 25, 48, 74	0
1	B	237/253 (93%)	-0.26	4 (1%)	70	71	17, 32, 53, 79	0
1	C	232/253 (91%)	-0.42	1 (0%)	92	93	17, 25, 48, 93	0
1	D	232/253 (91%)	-0.45	4 (1%)	70	71	16, 25, 49, 94	0
1	E	232/253 (91%)	-0.44	3 (1%)	77	78	16, 24, 47, 77	0
1	F	237/253 (93%)	-0.25	3 (1%)	77	78	18, 32, 56, 81	0
All	All	1404/1518 (92%)	-0.37	17 (1%)	79	80	15, 27, 53, 94	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2	ALA	5.2
1	F	2	ALA	3.8
1	D	2	ALA	3.7
1	A	217	GLU	3.2
1	F	148	ASP	3.1

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.