



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 08:54 PM GMT

PDB ID : 2Q4F  
Title : Ensemble refinement of the crystal structure of putative histidine-containing phosphotransfer protein from rice, Ak104879  
Authors : Levin, E.J.; Kondrashov, D.A.; Wesenberg, G.E.; Phillips Jr., G.N.; Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2007-05-31  
Resolution : 2.00 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

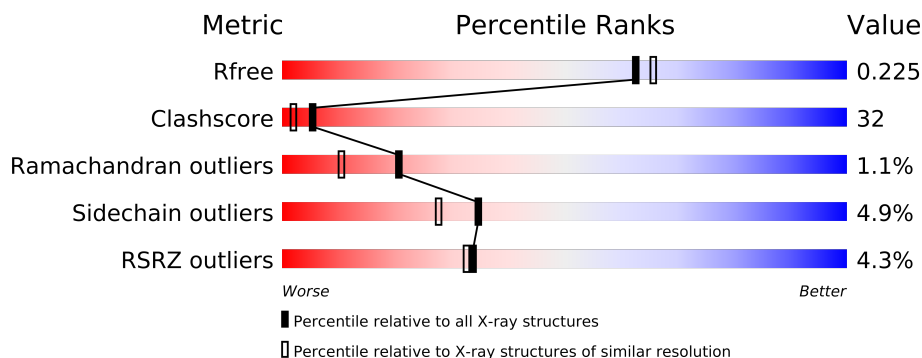
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.00 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.



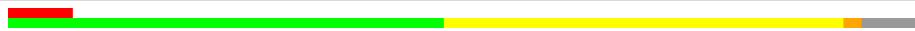
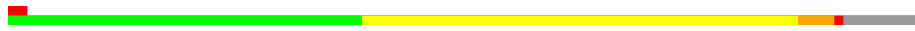
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. 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	1-A	149	
1	1-B	149	
1	2-A	149	
1	2-B	149	
1	3-A	149	
1	3-B	149	
1	4-A	149	
1	4-B	149	
1	5-A	149	
1	5-B	149	
1	6-A	149	
1	6-B	149	
1	7-A	149	
1	7-B	149	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	8-A	149	
1	8-B	149	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20000 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 Histidine-containing phosphotransfer protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	1-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	2-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	2-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	3-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	3-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	4-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	4-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	5-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	5-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	6-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	6-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	7-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	7-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			
1	8-A	140	Total	C	N	O	S	0	0	0
			1102	692	188	213	9			
1	8-B	136	Total	C	N	O	S	0	0	0
			1079	679	184	207	9			

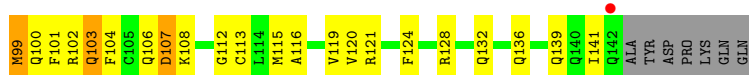
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	EXPRESSION TAG	UNP Q6VAK4
B	1	SER	-	EXPRESSION TAG	UNP Q6VAK4

- Molecule 2 is water.

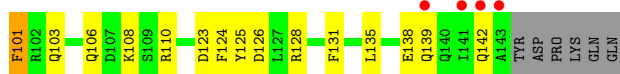
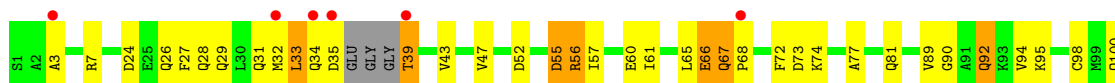
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	1-A	172	Total O 172 172	0	0
2	1-B	147	Total O 147 147	0	0
2	2-A	169	Total O 169 169	0	0
2	2-B	150	Total O 150 150	0	0
2	3-A	172	Total O 172 172	0	0
2	3-B	147	Total O 147 147	0	0
2	4-A	172	Total O 172 172	0	0
2	4-B	147	Total O 147 147	0	0
2	5-A	172	Total O 172 172	0	0
2	5-B	147	Total O 147 147	0	0
2	6-A	171	Total O 171 171	0	0
2	6-B	148	Total O 148 148	0	0
2	7-A	172	Total O 172 172	0	0
2	7-B	147	Total O 147 147	0	0
2	8-A	173	Total O 173 173	0	0
2	8-B	146	Total O 146 146	0	0





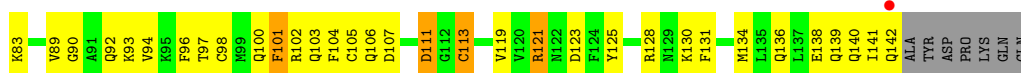
- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 3-A:



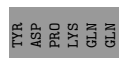
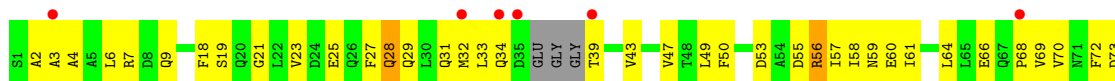
- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 3-B:



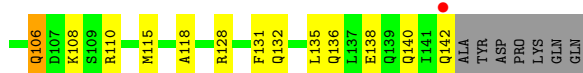
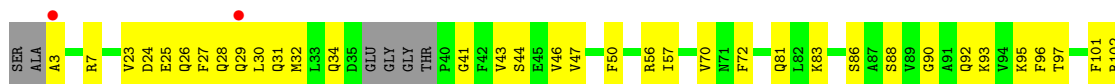
- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 4-A:



- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 4-B:



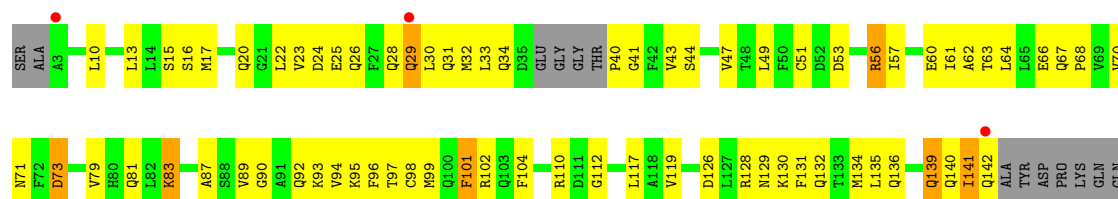
- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 5-A:



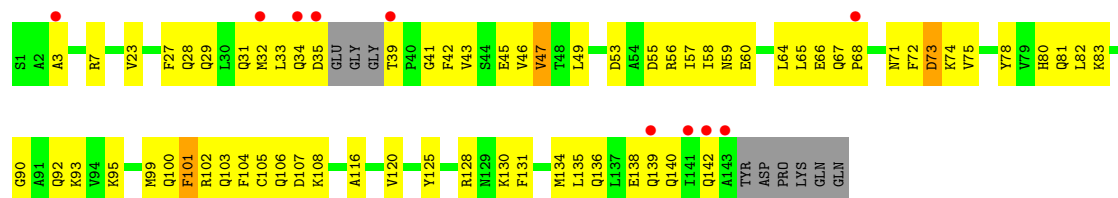


Chain 7-B: 



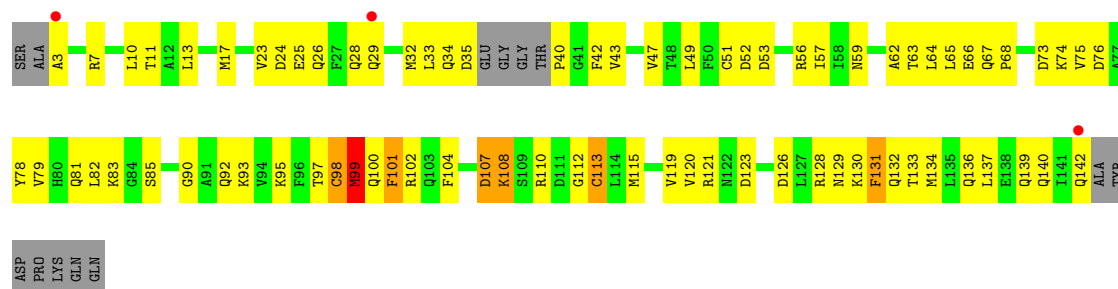
- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 8-A: 



- Molecule 1: Histidine-containing phosphotransfer protein 1

Chain 8-B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.56Å 100.56Å 69.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	37.77 – 2.00 37.77 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.0 (37.77-2.00) 99.0 (37.77-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.22 (at 2.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.145 , 0.222 0.157 , 0.225	Depositor DCC
$R_{free}$ test set	1249 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	19.5	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 57.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 24449 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	20000	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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 ⓘ

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	1-A	1.03	0/1115	0.83	1/1501 (0.1%)
1	1-B	0.99	0/1092	0.89	3/1468 (0.2%)
1	2-A	1.01	2/1115 (0.2%)	0.77	0/1501
1	2-B	0.93	0/1092	0.86	2/1468 (0.1%)
1	3-A	1.04	3/1115 (0.3%)	0.83	1/1501 (0.1%)
1	3-B	1.03	1/1092 (0.1%)	0.92	2/1468 (0.1%)
1	4-A	1.02	0/1115	0.93	5/1501 (0.3%)
1	4-B	0.95	0/1092	0.86	0/1468
1	5-A	1.12	2/1115 (0.2%)	0.92	0/1501
1	5-B	1.11	2/1092 (0.2%)	1.01	4/1468 (0.3%)
1	6-A	1.13	3/1115 (0.3%)	0.90	3/1501 (0.2%)
1	6-B	1.06	0/1092	1.00	6/1468 (0.4%)
1	7-A	1.12	3/1115 (0.3%)	0.90	1/1501 (0.1%)
1	7-B	1.07	2/1092 (0.2%)	0.94	3/1468 (0.2%)
1	8-A	1.14	3/1115 (0.3%)	0.89	2/1501 (0.1%)
1	8-B	1.17	3/1092 (0.3%)	0.96	2/1468 (0.1%)
All	All	1.06	24/17656 (0.1%)	0.90	35/23752 (0.1%)

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	8-B	113	CYS	CB-SG	-8.95	1.67	1.82
1	8-B	98	CYS	CB-SG	-8.49	1.67	1.82
1	5-B	51	CYS	CB-SG	7.64	1.95	1.82
1	3-B	113	CYS	CB-SG	-7.42	1.69	1.82
1	7-A	125	TYR	CD2-CE2	7.00	1.49	1.39

The worst 5 of 35 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	4-A	121	ARG	NE-CZ-NH1	9.77	125.19	120.30
1	6-B	52	ASP	CB-CG-OD1	8.35	125.81	118.30
1	5-B	123	ASP	CB-CG-OD2	-7.91	111.18	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	7-A	137	LEU	CA-CB-CG	7.76	133.15	115.30
1	5-B	121	ARG	NE-CZ-NH2	-7.68	116.46	120.30

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	1-A	1102	0	1092	87	1
1	1-B	1079	0	1068	85	1
1	2-A	1102	0	1092	36	1
1	2-B	1079	0	1068	68	0
1	3-A	1102	0	1092	50	1
1	3-B	1079	0	1068	83	0
1	4-A	1102	0	1092	72	0
1	4-B	1079	0	1068	48	0
1	5-A	1102	0	1092	78	1
1	5-B	1079	0	1068	107	1
1	6-A	1102	0	1092	74	0
1	6-B	1079	0	1068	73	0
1	7-A	1102	0	1092	58	0
1	7-B	1079	0	1068	100	0
1	8-A	1102	0	1092	69	0
1	8-B	1079	0	1068	73	0
2	1-A	172	0	0	23	0
2	1-B	147	0	0	17	0
2	2-A	169	0	0	12	1
2	2-B	150	0	0	14	0
2	3-A	172	0	0	17	1
2	3-B	147	0	0	15	0
2	4-A	172	0	0	18	0
2	4-B	147	0	0	15	0
2	5-A	172	0	0	17	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	5-B	147	0	0	28	0
2	6-A	171	0	0	15	0
2	6-B	148	0	0	17	0
2	7-A	172	0	0	18	0
2	7-B	147	0	0	21	0
2	8-A	173	0	0	17	0
2	8-B	146	0	0	18	0
All	All	20000	0	17280	1128	4

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 32.

The worst 5 of 1128 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:97:THR:HG23	2:B:185:HOH:O	1.34	1.23
1:B:57:ILE:HG21	1:B:82:LEU:HB2	1.34	1.10
1:B:97:THR:HG23	2:B:186:HOH:O	1.52	1.10
1:A:57:ILE:HD13	1:A:82:LEU:HD13	1.33	1.08
1:A:43:VAL:O	1:A:47:VAL:HG23	1.53	1.08

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:121:ARG:NH1	1:B:122:ASN:OD1[3_544]	2.03	0.17
1:A:136:GLN:OE1	1:B:67:GLN:OE1[3_544]	2.06	0.14
1:A:100:GLN:NE2	2:A:285:HOH:O[7_555]	2.19	0.01
1:A:100:GLN:NE2	2:A:288:HOH:O[7_555]	2.20	0.00

## 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	1-A	136/149 (91%)	129 (95%)	7 (5%)	0	100 100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	1-B	132/149 (89%)	123 (93%)	7 (5%)	2 (2%)	15	6
1	2-A	136/149 (91%)	133 (98%)	2 (2%)	1 (1%)	30	20
1	2-B	132/149 (89%)	125 (95%)	7 (5%)	0	100	100
1	3-A	136/149 (91%)	128 (94%)	5 (4%)	3 (2%)	10	3
1	3-B	132/149 (89%)	120 (91%)	8 (6%)	4 (3%)	7	2
1	4-A	136/149 (91%)	133 (98%)	2 (2%)	1 (1%)	30	20
1	4-B	132/149 (89%)	126 (96%)	6 (4%)	0	100	100
1	5-A	136/149 (91%)	122 (90%)	13 (10%)	1 (1%)	30	20
1	5-B	132/149 (89%)	117 (89%)	10 (8%)	5 (4%)	5	1
1	6-A	136/149 (91%)	128 (94%)	6 (4%)	2 (2%)	15	6
1	6-B	132/149 (89%)	124 (94%)	7 (5%)	1 (1%)	27	17
1	7-A	136/149 (91%)	132 (97%)	4 (3%)	0	100	100
1	7-B	132/149 (89%)	126 (96%)	5 (4%)	1 (1%)	27	17
1	8-A	136/149 (91%)	125 (92%)	10 (7%)	1 (1%)	30	20
1	8-B	132/149 (89%)	127 (96%)	4 (3%)	1 (1%)	27	17
All	All	2144/2384 (90%)	2018 (94%)	103 (5%)	23 (1%)	21	10

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	3-B	4	ALA
1	3-B	5	ALA
1	5-B	19	SER
1	6-A	65	LEU
1	1-B	4	ALA

### 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	1-A	122/129 (95%)	110 (90%)	12 (10%)	12	6
1	1-B	120/129 (93%)	117 (98%)	3 (2%)	60	59

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	2-A	122/129 (95%)	119 (98%)	3 (2%)	60	59
1	2-B	120/129 (93%)	114 (95%)	6 (5%)	34	27
1	3-A	122/129 (95%)	117 (96%)	5 (4%)	41	35
1	3-B	120/129 (93%)	116 (97%)	4 (3%)	50	46
1	4-A	122/129 (95%)	117 (96%)	5 (4%)	41	35
1	4-B	120/129 (93%)	118 (98%)	2 (2%)	73	75
1	5-A	122/129 (95%)	114 (93%)	8 (7%)	24	16
1	5-B	120/129 (93%)	110 (92%)	10 (8%)	16	10
1	6-A	122/129 (95%)	115 (94%)	7 (6%)	29	21
1	6-B	120/129 (93%)	110 (92%)	10 (8%)	16	10
1	7-A	122/129 (95%)	118 (97%)	4 (3%)	50	46
1	7-B	120/129 (93%)	114 (95%)	6 (5%)	34	27
1	8-A	122/129 (95%)	119 (98%)	3 (2%)	60	59
1	8-B	120/129 (93%)	114 (95%)	6 (5%)	34	27
All	All	1936/2064 (94%)	1842 (95%)	94 (5%)	35	28

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	5-A	32	MET
1	5-B	44	SER
1	8-A	35	ASP
1	5-A	39	THR
1	5-A	101	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 74 such sidechains are listed below:

Mol	Chain	Res	Type
1	4-B	106	GLN
1	5-A	81	GLN
1	8-A	31	GLN
1	4-B	136	GLN
1	5-A	26	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 ⓘ

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, 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	1-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	1-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	2-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	2-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	3-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	3-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	4-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	4-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	5-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	5-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	6-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	6-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	7-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	7-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
1	8-A	140/149 (93%)	-0.19	10 (7%) 16 15	7, 14, 37, 55	140 (100%)
1	8-B	136/149 (91%)	-0.24	3 (2%) 59 59	5, 15, 38, 51	136 (100%)
All	All	2208/2384 (92%)	-0.22	104 (4%) 34 29	5, 15, 39, 55	2208 (100%)

The worst 5 of 104 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	35	ASP	4.0
1	2-A	35	ASP	4.0
1	3-A	35	ASP	4.0
1	4-A	35	ASP	4.0
1	5-A	35	ASP	4.0

## 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 ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.