Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 12:18 am BST

PDB ID	:	4NFX
Title	:	Structure and atypical hydrolysis mechanism of the Nudix hydrolase Orf153 (YmfB) from Escherichia coli
Authors	:	Hong, M.K.; Kim, J.K.; Kang, L.W.
Deposited on	:	2013-11-01
Resolution	:	2.69 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/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:

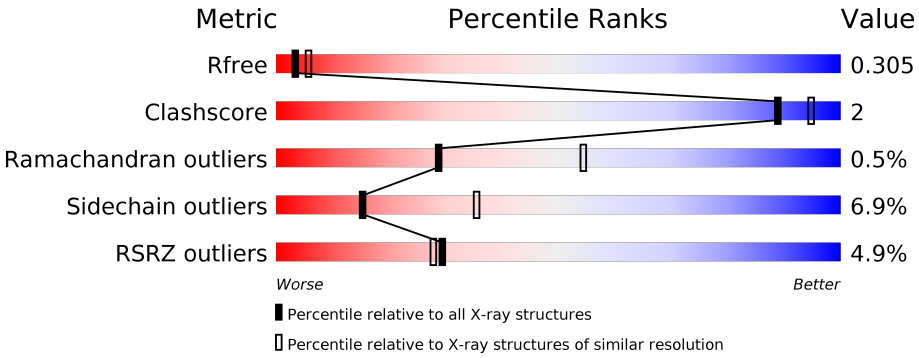
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.69 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.



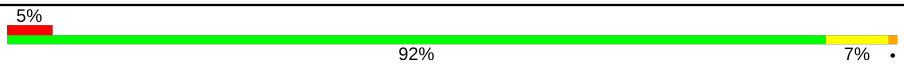
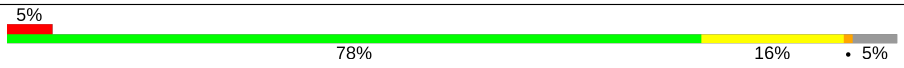
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 >=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	153	4% 90% 10%
1	B	153	5% 88% 10%
1	C	153	3% 92% 8%
1	D	153	5% 92% 5%
1	E	153	5% 88% 12%
1	F	153	8% 90% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	153	
1	H	153	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9905 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 Putative Nudix hydrolase ymfB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	153	Total	C	N	O	S	0	0	0
			1227	786	209	224	8			
1	B	149	Total	C	N	O	S	0	0	0
			1199	767	204	220	8			
1	C	153	Total	C	N	O	S	0	0	0
			1227	786	209	224	8			
1	D	150	Total	C	N	O	S	0	0	0
			1208	773	206	221	8			
1	E	153	Total	C	N	O	S	0	0	0
			1227	786	209	224	8			
1	F	153	Total	C	N	O	S	0	0	0
			1227	786	209	224	8			
1	G	153	Total	C	N	O	S	0	0	0
			1228	786	209	225	8			
1	H	145	Total	C	N	O	S	0	0	0
			1160	738	199	215	8			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	34	Total	O	0	0
			34	34		
2	B	20	Total	O	0	0
			20	20		
2	C	17	Total	O	0	0
			17	17		
2	D	23	Total	O	0	0
			23	23		
2	E	31	Total	O	0	0
			31	31		
2	F	31	Total	O	0	0
			31	31		

Continued on next page...

Continued from previous page...

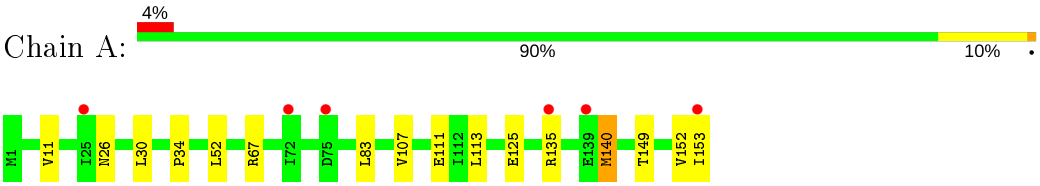
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	18	Total O 18 18	0	0
2	H	28	Total O 28 28	0	0

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 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.

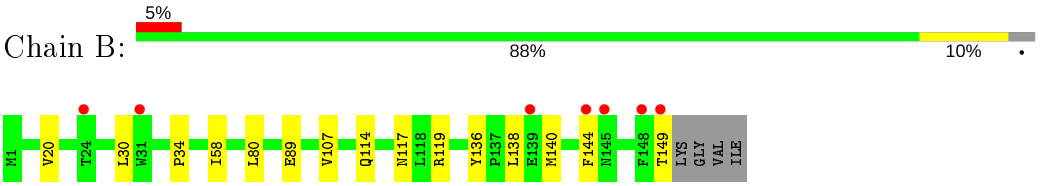
- Molecule 1: Putative Nudix hydrolase ymfB

Chain A:



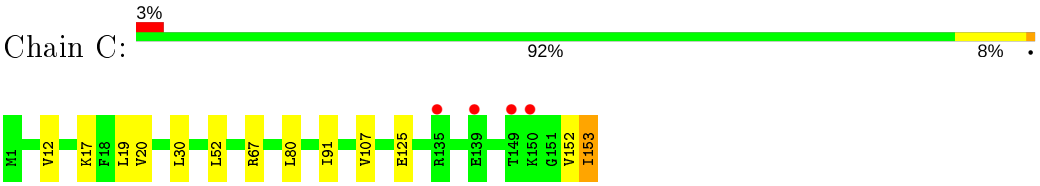
- Molecule 1: Putative Nudix hydrolase ymfB

Chain B:



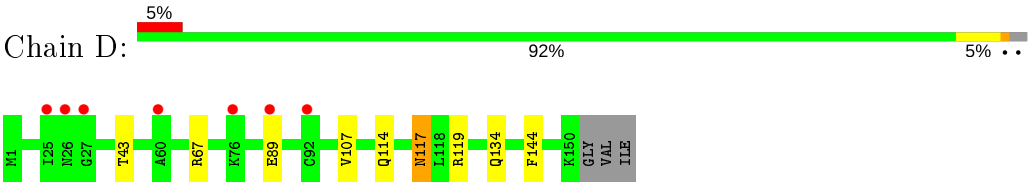
- Molecule 1: Putative Nudix hydrolase ymfB

Chain C:



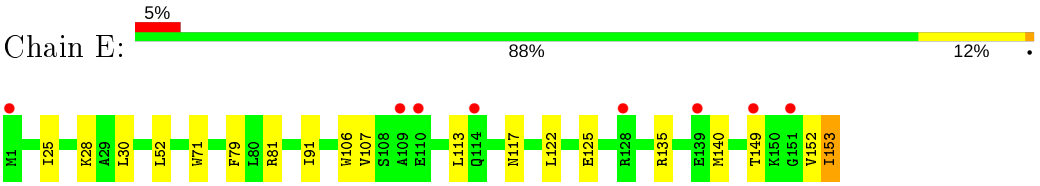
- Molecule 1: Putative Nudix hydrolase ymfB

Chain D:



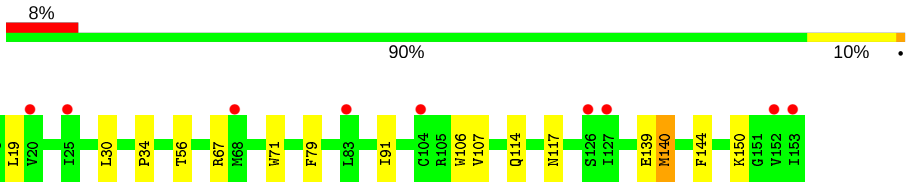
- Molecule 1: Putative Nudix hydrolase ymfB

Chain E:



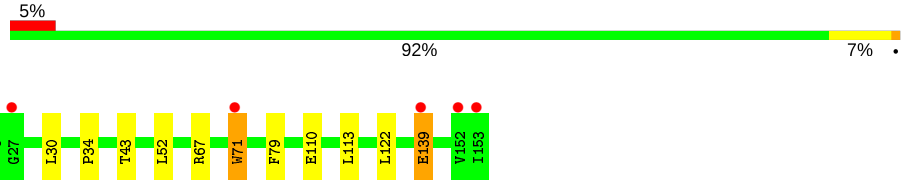
- Molecule 1: Putative Nudix hydrolase ymfB

Chain F:



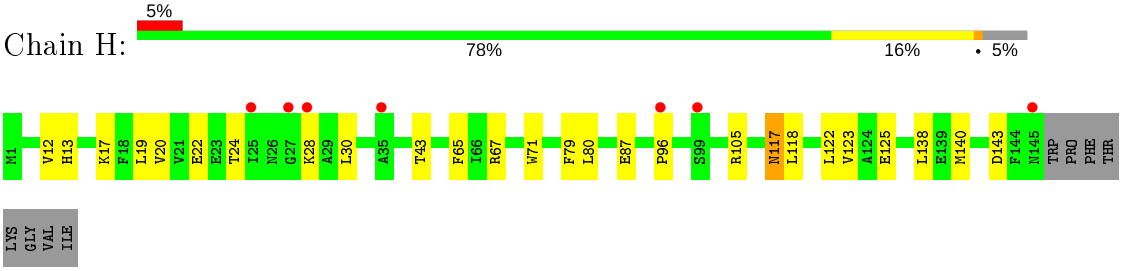
• Molecule 1: Putative Nudix hydrolase ymfB

Chain G:



• Molecule 1: Putative Nudix hydrolase ymfB

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.96Å 70.52Å 145.03Å 90.00° 103.30° 90.00°	Depositor
Resolution (Å)	33.56 – 2.69 33.56 – 2.69	Depositor EDS
% Data completeness (in resolution range)	98.2 (33.56-2.69) 98.3 (33.56-2.69)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.60 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.233 , 0.288 0.272 , 0.305	Depositor DCC
R_{free} test set	1865 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	33.6	Xtriage
Anisotropy	0.134	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 19.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.019 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.86	EDS
Total number of atoms	9905	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.40% 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	0.27	0/1262	0.46	0/1719
1	B	0.27	0/1234	0.47	0/1682
1	C	0.27	0/1262	0.48	0/1719
1	D	0.27	0/1243	0.46	0/1693
1	E	0.27	0/1262	0.47	0/1719
1	F	0.27	0/1262	0.46	0/1719
1	G	0.27	0/1263	0.47	0/1719
1	H	0.28	0/1191	0.48	0/1621
All	All	0.27	0/9979	0.47	0/13591

There are no bond length outliers.

There are no bond angle outliers.

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	1227	0	1190	5	0
1	B	1199	0	1154	4	0
1	C	1227	0	1190	4	0
1	D	1208	0	1167	2	0
1	E	1227	0	1190	9	0
1	F	1227	0	1190	6	0
1	G	1228	0	1190	3	0
1	H	1160	0	1121	8	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	34	0	0	0	0
2	B	20	0	0	1	0
2	C	17	0	0	0	0
2	D	23	0	0	0	0
2	E	31	0	0	0	0
2	F	31	0	0	0	0
2	G	18	0	0	0	0
2	H	28	0	0	0	0
All	All	9905	0	9392	36	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 2.

All (36) 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:E:152:VAL:HG13	1:E:153:ILE:HD13	1.74	0.69
1:A:107:VAL:HG22	1:A:111:GLU:HB2	1.83	0.60
1:A:152:VAL:HG12	1:A:153:ILE:HG13	1.89	0.55
1:A:125:GLU:HG3	1:B:144:PHE:CE1	2.45	0.52
1:A:140:MET:HG2	1:A:140:MET:O	2.11	0.51
1:E:81:ARG:HD3	1:E:122:LEU:HD13	1.91	0.51
1:E:125:GLU:HG3	1:F:144:PHE:CE1	2.46	0.51
1:G:139:GLU:HA	1:G:139:GLU:OE1	2.11	0.51
1:C:17:LYS:HG3	1:C:91:ILE:HD11	1.94	0.49
1:D:117:ASN:C	1:D:117:ASN:HD22	2.16	0.49
1:G:71:TRP:CD1	1:G:79:PHE:HB2	2.48	0.48
1:H:13:HIS:ND1	1:H:87:GLU:OE2	2.43	0.48
1:H:71:TRP:CD1	1:H:79:PHE:HB2	2.50	0.47
1:F:19:LEU:HD21	1:F:56:THR:HG21	1.97	0.46
1:H:117:ASN:C	1:H:117:ASN:HD22	2.19	0.46
1:C:152:VAL:HG12	1:C:153:ILE:H	1.81	0.45
1:E:71:TRP:CD1	1:E:79:PHE:HB2	2.51	0.45
1:C:125:GLU:HG3	1:D:144:PHE:CE1	2.54	0.43
1:E:135:ARG:HD3	1:E:135:ARG:HA	1.75	0.42
1:H:12:VAL:HB	1:H:19:LEU:HB3	2.02	0.42
1:E:71:TRP:CD1	1:E:122:LEU:HD11	2.55	0.42
1:H:118:LEU:HD22	1:H:123:VAL:HG11	2.00	0.42
1:A:11:VAL:HG23	1:A:83:LEU:HD11	2.01	0.42
1:H:65:PHE:HB2	1:H:138:LEU:HD23	2.02	0.42
1:B:138:LEU:HG	2:B:206:HOH:O	2.19	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:12:VAL:HB	1:C:19:LEU:HB3	2.01	0.41
1:B:136:TYR:HB3	1:B:140:MET:HE3	2.03	0.41
1:E:91:ILE:HG23	1:E:106:TRP:CE2	2.56	0.41
1:H:122:LEU:HA	1:H:125:GLU:HB3	2.03	0.41
1:E:125:GLU:HG3	1:F:144:PHE:CZ	2.56	0.41
1:H:117:ASN:C	1:H:117:ASN:ND2	2.75	0.41
1:E:140:MET:HE1	1:F:140:MET:SD	2.62	0.40
1:F:91:ILE:HG23	1:F:106:TRP:CE2	2.55	0.40
1:G:71:TRP:CD1	1:G:122:LEU:HD11	2.57	0.40
1:F:71:TRP:CD1	1:F:79:PHE:HB2	2.57	0.40
1:B:136:TYR:HB3	1:B:140:MET:CE	2.51	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	151/153 (99%)	144 (95%)	5 (3%)	2 (1%)	12	30
1	B	147/153 (96%)	140 (95%)	6 (4%)	1 (1%)	22	46
1	C	151/153 (99%)	145 (96%)	6 (4%)	0	100	100
1	D	148/153 (97%)	144 (97%)	4 (3%)	0	100	100
1	E	151/153 (99%)	148 (98%)	3 (2%)	0	100	100
1	F	151/153 (99%)	144 (95%)	6 (4%)	1 (1%)	22	46
1	G	151/153 (99%)	144 (95%)	6 (4%)	1 (1%)	22	46
1	H	143/153 (94%)	138 (96%)	4 (3%)	1 (1%)	22	46
All	All	1193/1224 (98%)	1147 (96%)	40 (3%)	6 (0%)	29	54

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	ASN
1	G	34	PRO
1	H	96	PRO
1	B	34	PRO
1	F	34	PRO
1	A	34	PRO

5.3.2 Protein sidechains (i)

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	132/132 (100%)	125 (95%)	7 (5%)	22 48
1	B	129/132 (98%)	119 (92%)	10 (8%)	12 29
1	C	132/132 (100%)	125 (95%)	7 (5%)	22 48
1	D	130/132 (98%)	122 (94%)	8 (6%)	18 40
1	E	132/132 (100%)	123 (93%)	9 (7%)	16 36
1	F	132/132 (100%)	124 (94%)	8 (6%)	18 41
1	G	132/132 (100%)	122 (92%)	10 (8%)	13 30
1	H	125/132 (95%)	112 (90%)	13 (10%)	7 16
All	All	1044/1056 (99%)	972 (93%)	72 (7%)	15 35

All (72) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	LEU
1	A	52	LEU
1	A	67	ARG
1	A	113	LEU
1	A	135	ARG
1	A	140	MET
1	A	149	THR
1	B	20	VAL
1	B	30	LEU
1	B	58	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	80	LEU
1	B	89	GLU
1	B	107	VAL
1	B	114	GLN
1	B	117	ASN
1	B	119	ARG
1	B	149	THR
1	C	20	VAL
1	C	30	LEU
1	C	52	LEU
1	C	67	ARG
1	C	80	LEU
1	C	107	VAL
1	C	153	ILE
1	D	43	THR
1	D	67	ARG
1	D	89	GLU
1	D	107	VAL
1	D	114	GLN
1	D	117	ASN
1	D	119	ARG
1	D	134	GLN
1	E	25	ILE
1	E	28	LYS
1	E	30	LEU
1	E	52	LEU
1	E	107	VAL
1	E	113	LEU
1	E	117	ASN
1	E	149	THR
1	E	153	ILE
1	F	30	LEU
1	F	67	ARG
1	F	107	VAL
1	F	114	GLN
1	F	117	ASN
1	F	139	GLU
1	F	140	MET
1	F	150	LYS
1	G	24	THR
1	G	25	ILE
1	G	30	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	43	THR
1	G	52	LEU
1	G	67	ARG
1	G	71	TRP
1	G	110	GLU
1	G	113	LEU
1	G	139	GLU
1	H	17	LYS
1	H	20	VAL
1	H	22	GLU
1	H	24	THR
1	H	28	LYS
1	H	30	LEU
1	H	43	THR
1	H	67	ARG
1	H	80	LEU
1	H	105	ARG
1	H	117	ASN
1	H	140	MET
1	H	143	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	HIS
1	A	95	GLN
1	A	134	GLN
1	B	117	ASN
1	C	26	ASN
1	C	63	GLN
1	C	95	GLN
1	D	37	HIS
1	D	117	ASN
1	E	117	ASN
1	F	117	ASN
1	F	134	GLN
1	G	95	GLN
1	H	117	ASN

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 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^{th} 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	153/153 (100%)	0.34	6 (3%)	39	38	44, 51, 62, 67	2 (1%)
1	B	149/153 (97%)	0.33	7 (4%)	31	30	39, 49, 65, 82	0
1	C	153/153 (100%)	0.39	4 (2%)	56	57	46, 57, 74, 96	2 (1%)
1	D	150/153 (98%)	0.29	7 (4%)	31	30	29, 44, 54, 65	0
1	E	153/153 (100%)	0.32	8 (5%)	27	25	36, 44, 54, 61	2 (1%)
1	F	153/153 (100%)	0.85	13 (8%)	10	9	33, 66, 87, 105	0
1	G	153/153 (100%)	0.88	7 (4%)	32	31	58, 79, 103, 119	2 (1%)
1	H	145/153 (94%)	0.59	7 (4%)	30	28	48, 62, 84, 106	0
All	All	1209/1224 (98%)	0.50	59 (4%)	29	28	29, 54, 85, 119	8 (0%)

All (59) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	149	THR	5.2
1	G	153	ILE	4.3
1	D	26	ASN	3.7
1	G	152	VAL	3.7
1	A	153	ILE	3.6
1	H	145	ASN	3.6
1	E	1	MET	3.5
1	E	149	THR	3.5
1	E	114	GLN	3.4
1	A	139	GLU	3.4
1	H	28	LYS	3.4
1	F	25	ILE	3.2
1	H	27	GLY	3.2
1	H	96	PRO	3.2
1	G	139	GLU	3.2
1	A	135	ARG	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	25	ILE	3.0
1	H	25	ILE	3.0
1	D	89	GLU	3.0
1	E	109	ALA	3.0
1	C	139	GLU	2.9
1	E	151	GLY	2.9
1	A	25	ILE	2.9
1	F	152	VAL	2.9
1	C	150	LYS	2.8
1	F	83	LEU	2.8
1	F	127	ILE	2.8
1	C	149	THR	2.8
1	G	27	GLY	2.7
1	D	76	LYS	2.7
1	D	27	GLY	2.7
1	A	72	ILE	2.6
1	F	68	MET	2.6
1	G	71	TRP	2.6
1	F	20	VAL	2.6
1	C	135	ARG	2.5
1	F	12	VAL	2.5
1	F	9	ALA	2.4
1	B	24	THR	2.4
1	F	18	PHE	2.4
1	G	1	MET	2.3
1	D	92	CYS	2.3
1	F	153	ILE	2.3
1	B	31	TRP	2.3
1	G	9	ALA	2.3
1	F	1	MET	2.2
1	H	35	ALA	2.2
1	E	110	GLU	2.1
1	B	148	PHE	2.1
1	A	75	ASP	2.1
1	E	128	ARG	2.1
1	F	104	CYS	2.1
1	B	144	PHE	2.1
1	B	139	GLU	2.1
1	H	99	SER	2.1
1	B	145	ASN	2.0
1	F	126	SER	2.0
1	D	60	ALA	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	139	GLU	2.0

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.