wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 05:48 PM EDT

PDB ID	:	3QOD
Title	:	Crystal Structure of Heterocyst Differentiation Protein, HetR from Fischerella mv11
Authors	:	Kim, Y.; Joachimiak, G.; Gornicki, P.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2011-02-09
Resolution	:	3.38 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

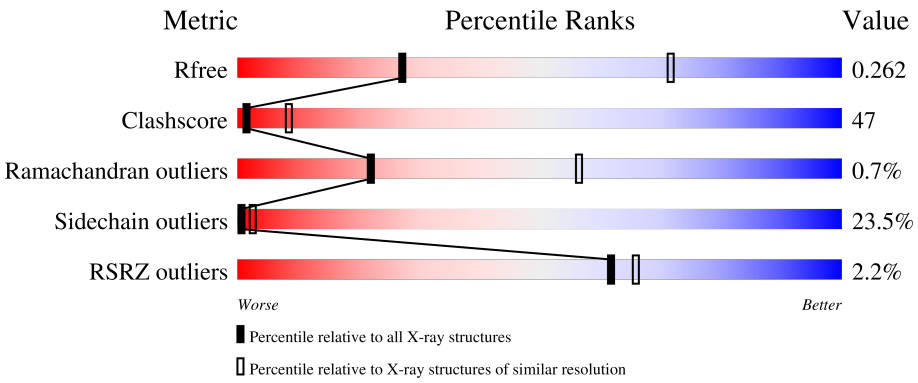
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.38 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	1691 (3.46-3.30)
Clashscore	141614	1762 (3.46-3.30)
Ramachandran outliers	138981	1732 (3.46-3.30)
Sidechain outliers	138945	1731 (3.46-3.30)
RSRZ outliers	127900	1635 (3.46-3.30)

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	302	2% 38% 45% 14% •
1	B	302	2% 30% 51% 14% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4796 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 Heterocyst differentiation protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	294	2415	1540	425	434	1	15	0	0	0
1	B	289	2377	1511	420	430	1	15	0	0	0

There are 110 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q2ACK9
A	-1	ASN	-	EXPRESSION TAG	UNP Q2ACK9
A	0	ALA	-	EXPRESSION TAG	UNP Q2ACK9
A	1	MSE	-	EXPRESSION TAG	UNP Q2ACK9
A	2	SER	-	EXPRESSION TAG	UNP Q2ACK9
A	3	ASN	-	EXPRESSION TAG	UNP Q2ACK9
A	4	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	5	VAL	-	EXPRESSION TAG	UNP Q2ACK9
A	6	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	7	LEU	-	EXPRESSION TAG	UNP Q2ACK9
A	8	ILE	-	EXPRESSION TAG	UNP Q2ACK9
A	9	LYS	-	EXPRESSION TAG	UNP Q2ACK9
A	10	ARG	-	EXPRESSION TAG	UNP Q2ACK9
A	11	LEU	-	EXPRESSION TAG	UNP Q2ACK9
A	12	GLY	-	EXPRESSION TAG	UNP Q2ACK9
A	13	PRO	-	EXPRESSION TAG	UNP Q2ACK9
A	14	SER	-	EXPRESSION TAG	UNP Q2ACK9
A	15	ALA	-	EXPRESSION TAG	UNP Q2ACK9
A	16	MSE	-	EXPRESSION TAG	UNP Q2ACK9
A	17	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	18	GLN	-	EXPRESSION TAG	UNP Q2ACK9
A	19	ILE	-	EXPRESSION TAG	UNP Q2ACK9
A	20	MSE	-	EXPRESSION TAG	UNP Q2ACK9
A	21	LEU	-	EXPRESSION TAG	UNP Q2ACK9
A	22	TYR	-	EXPRESSION TAG	UNP Q2ACK9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	LEU	-	EXPRESSION TAG	UNP Q2ACK9
A	24	ALA	-	EXPRESSION TAG	UNP Q2ACK9
A	25	PHE	-	EXPRESSION TAG	UNP Q2ACK9
A	26	SER	-	EXPRESSION TAG	UNP Q2ACK9
A	27	ALA	-	EXPRESSION TAG	UNP Q2ACK9
A	28	MSE	-	EXPRESSION TAG	UNP Q2ACK9
A	276	TRP	-	EXPRESSION TAG	UNP Q2ACK9
A	277	ALA	-	EXPRESSION TAG	UNP Q2ACK9
A	278	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	279	LYS	-	EXPRESSION TAG	UNP Q2ACK9
A	280	TYR	-	EXPRESSION TAG	UNP Q2ACK9
A	281	HIS	-	EXPRESSION TAG	UNP Q2ACK9
A	282	GLN	-	EXPRESSION TAG	UNP Q2ACK9
A	283	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	284	ASP	-	EXPRESSION TAG	UNP Q2ACK9
A	285	GLY	-	EXPRESSION TAG	UNP Q2ACK9
A	286	VAL	-	EXPRESSION TAG	UNP Q2ACK9
A	287	PRO	-	EXPRESSION TAG	UNP Q2ACK9
A	288	VAL	-	EXPRESSION TAG	UNP Q2ACK9
A	289	VAL	-	EXPRESSION TAG	UNP Q2ACK9
A	290	LEU	-	EXPRESSION TAG	UNP Q2ACK9
A	291	GLN	-	EXPRESSION TAG	UNP Q2ACK9
A	292	MSE	-	EXPRESSION TAG	UNP Q2ACK9
A	293	VAL	-	EXPRESSION TAG	UNP Q2ACK9
A	294	PHE	-	EXPRESSION TAG	UNP Q2ACK9
A	295	GLY	-	EXPRESSION TAG	UNP Q2ACK9
A	296	LYS	-	EXPRESSION TAG	UNP Q2ACK9
A	297	LYS	-	EXPRESSION TAG	UNP Q2ACK9
A	298	GLU	-	EXPRESSION TAG	UNP Q2ACK9
A	299	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	1	SER	-	EXPRESSION TAG	UNP Q2ACK9
B	2	ASN	-	EXPRESSION TAG	UNP Q2ACK9
B	3	ALA	-	EXPRESSION TAG	UNP Q2ACK9
B	4	MSE	-	EXPRESSION TAG	UNP Q2ACK9
B	5	SER	-	EXPRESSION TAG	UNP Q2ACK9
B	6	ASN	-	EXPRESSION TAG	UNP Q2ACK9
B	7	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	8	VAL	-	EXPRESSION TAG	UNP Q2ACK9
B	9	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	10	LEU	-	EXPRESSION TAG	UNP Q2ACK9
B	11	ILE	-	EXPRESSION TAG	UNP Q2ACK9
B	12	LYS	-	EXPRESSION TAG	UNP Q2ACK9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	13	ARG	-	EXPRESSION TAG	UNP Q2ACK9
B	14	LEU	-	EXPRESSION TAG	UNP Q2ACK9
B	15	GLY	-	EXPRESSION TAG	UNP Q2ACK9
B	16	PRO	-	EXPRESSION TAG	UNP Q2ACK9
B	17	SER	-	EXPRESSION TAG	UNP Q2ACK9
B	18	ALA	-	EXPRESSION TAG	UNP Q2ACK9
B	19	MSE	-	EXPRESSION TAG	UNP Q2ACK9
B	20	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	21	GLN	-	EXPRESSION TAG	UNP Q2ACK9
B	22	ILE	-	EXPRESSION TAG	UNP Q2ACK9
B	23	MSE	-	EXPRESSION TAG	UNP Q2ACK9
B	24	LEU	-	EXPRESSION TAG	UNP Q2ACK9
B	25	TYR	-	EXPRESSION TAG	UNP Q2ACK9
B	26	LEU	-	EXPRESSION TAG	UNP Q2ACK9
B	27	ALA	-	EXPRESSION TAG	UNP Q2ACK9
B	28	PHE	-	EXPRESSION TAG	UNP Q2ACK9
B	29	SER	-	EXPRESSION TAG	UNP Q2ACK9
B	30	ALA	-	EXPRESSION TAG	UNP Q2ACK9
B	31	MSE	-	EXPRESSION TAG	UNP Q2ACK9
B	279	TRP	-	EXPRESSION TAG	UNP Q2ACK9
B	280	ALA	-	EXPRESSION TAG	UNP Q2ACK9
B	281	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	282	LYS	-	EXPRESSION TAG	UNP Q2ACK9
B	283	TYR	-	EXPRESSION TAG	UNP Q2ACK9
B	284	HIS	-	EXPRESSION TAG	UNP Q2ACK9
B	285	GLN	-	EXPRESSION TAG	UNP Q2ACK9
B	286	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	287	ASP	-	EXPRESSION TAG	UNP Q2ACK9
B	288	GLY	-	EXPRESSION TAG	UNP Q2ACK9
B	289	VAL	-	EXPRESSION TAG	UNP Q2ACK9
B	290	PRO	-	EXPRESSION TAG	UNP Q2ACK9
B	291	VAL	-	EXPRESSION TAG	UNP Q2ACK9
B	292	VAL	-	EXPRESSION TAG	UNP Q2ACK9
B	293	LEU	-	EXPRESSION TAG	UNP Q2ACK9
B	294	GLN	-	EXPRESSION TAG	UNP Q2ACK9
B	295	MSE	-	EXPRESSION TAG	UNP Q2ACK9
B	296	VAL	-	EXPRESSION TAG	UNP Q2ACK9
B	297	PHE	-	EXPRESSION TAG	UNP Q2ACK9
B	298	GLY	-	EXPRESSION TAG	UNP Q2ACK9
B	299	LYS	-	EXPRESSION TAG	UNP Q2ACK9
B	300	LYS	-	EXPRESSION TAG	UNP Q2ACK9
B	301	GLU	-	EXPRESSION TAG	UNP Q2ACK9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	302	ASP	-	EXPRESSION TAG	UNP Q2ACK9

- Molecule 2 is water.

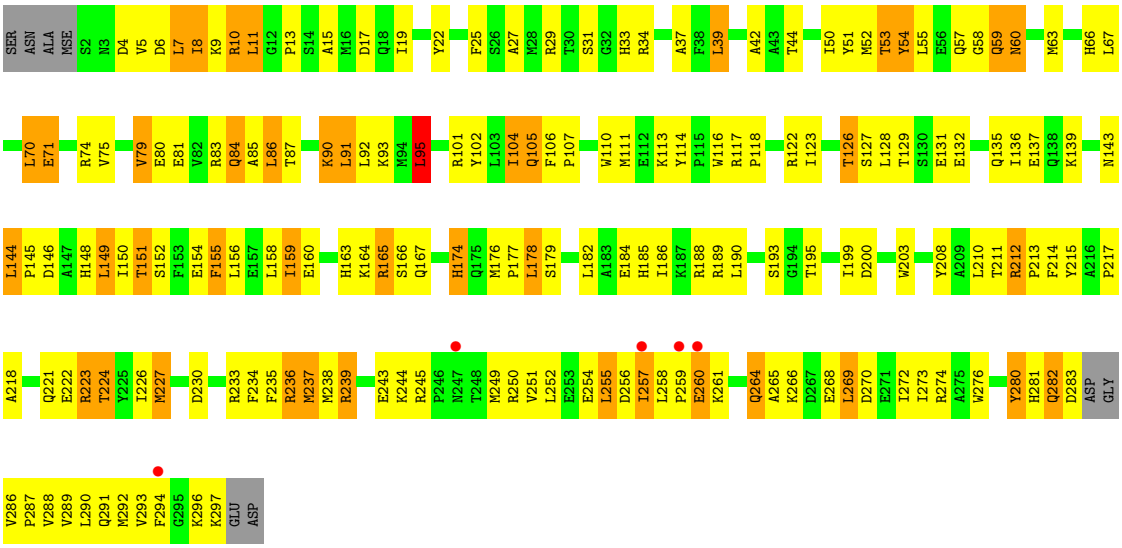
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	3	3	3	0	0
			Total	O		
2	B	1	1	1	0	0

3 Residue-property plots (i)

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.

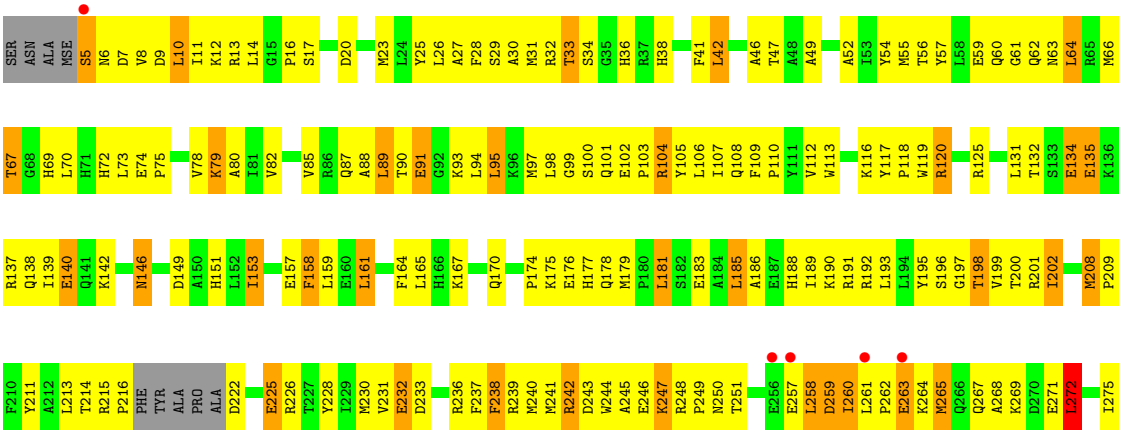
• Molecule 1: Heterocyst differentiation protein

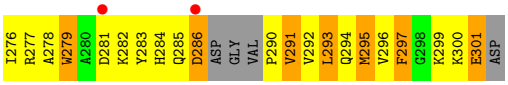
Chain A:



• Molecule 1: Heterocyst differentiation protein

Chain B:




I276
R277
A278
W279
A280
D281
K282
Y283
H284
Q285
D286
ASP
GLY
VAL
P290
V291
V292
L293
Q294
M295
V296
F297
G298
K299
K300
E301
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	123.38Å 123.38Å 109.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.39 – 3.38 40.39 – 3.38	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.39-3.38) 99.8 (40.39-3.38)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	3.50 (at 3.40Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7_650)	Depositor
R, R_{free}	0.204 , 0.267 0.193 , 0.262	Depositor DCC
R_{free} test set	681 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	94.2	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 73.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.077 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4796	wwPDB-VP
Average B, all atoms (Å^2)	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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 (i)

5.1 Standard geometry (i)

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	0.82	0/2458	0.91	2/3292 (0.1%)
1	B	0.82	3/2416 (0.1%)	0.92	2/3230 (0.1%)
All	All	0.82	3/4874 (0.1%)	0.91	4/6522 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	279	TRP	CB-CG	-7.73	1.36	1.50
1	B	140	GLU	CB-CG	-5.50	1.41	1.52
1	B	146	ASN	CB-CG	5.06	1.62	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	95	LEU	CA-CB-CG	6.11	129.36	115.30
1	B	213	LEU	CB-CG-CD2	-5.76	101.20	111.00
1	A	269	LEU	CA-CB-CG	-5.52	102.61	115.30
1	B	279	TRP	CA-CB-CG	-5.07	104.06	113.70

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	2415	0	2417	254	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2377	0	2379	279	0
2	A	3	0	0	0	0
2	B	1	0	0	0	0
All	All	4796	0	4796	447	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 47.

The worst 5 of 447 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:215:ARG:HH11	1:B:215:ARG:HB3	1.25	1.00
1:A:258:LEU:HD12	1:A:259:PRO:HD2	1.42	0.99
1:A:75:VAL:O	1:A:79:VAL:HG22	1.63	0.98
1:A:244:LYS:HG2	1:B:226:ARG:NH2	1.79	0.97
1:A:90:LYS:O	1:A:90:LYS:HG2	1.63	0.96

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	290/302 (96%)	264 (91%)	25 (9%)	1 (0%)	41 73
1	B	283/302 (94%)	245 (87%)	35 (12%)	3 (1%)	14 46
All	All	573/604 (95%)	509 (89%)	60 (10%)	4 (1%)	22 56

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	176	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	257	ILE
1	B	265	MSE
1	B	272	LEU

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	259/249 (104%)	200 (77%)	59 (23%)	1 3
1	B	256/249 (103%)	194 (76%)	62 (24%)	0 2
All	All	515/498 (103%)	394 (76%)	121 (24%)	1 2

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	280	TYR
1	B	263	GLU
1	B	73	LEU
1	B	261	LEU
1	B	295	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	282	GLN
1	B	21	GLN
1	B	170	GLN
1	B	69	HIS
1	B	72	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	279/302 (92%)	-0.02	5 (1%)	68	72	73, 99, 184, 197	0
1	B	274/302 (90%)	-0.07	7 (2%)	56	59	74, 104, 181, 219	0
All	All	553/604 (91%)	-0.04	12 (2%)	62	66	73, 102, 183, 219	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	259	PRO	3.1
1	B	263	GLU	3.1
1	B	256	GLU	3.0
1	A	294	PHE	2.6
1	B	261	LEU	2.6

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.