



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 09:15 am BST

PDB ID : 3Q6C  
Title : X-ray crystal structure of duf2500 (pf10694) from klebsiella pneumoniae, northeast structural genomics consortium target kpr96  
Authors : Seetharaman, J.; Su, M.; Wang, D.; Ciccocanti, C.; Sahdev, S.; Nair, R.; Rost, B.; Acton, T.B.; Xiao, R.; Everett, J.K.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)  
Deposited on : 2010-12-31  
Resolution : 2.60 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](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

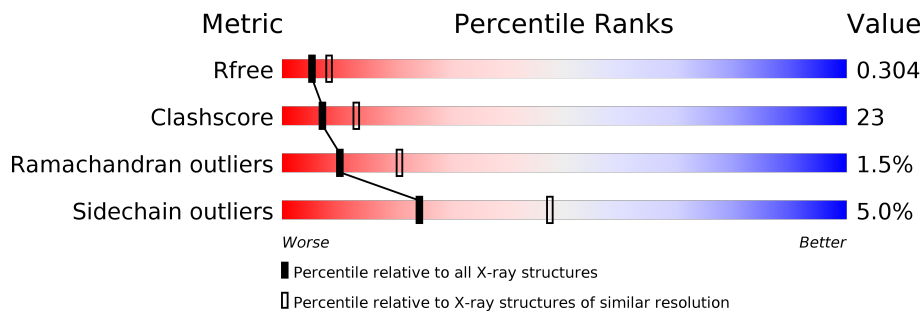
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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.





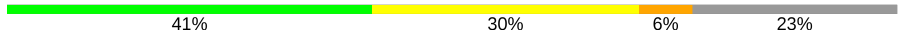
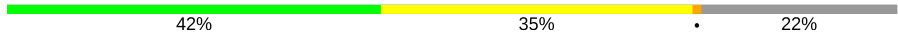

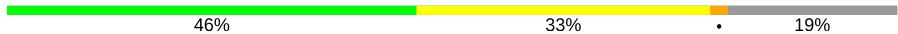
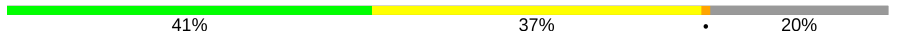
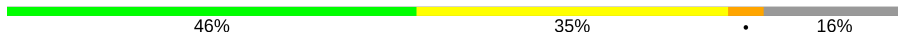
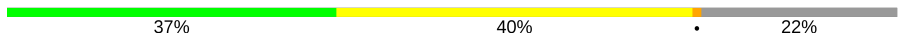
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	83	46% 31% • 20%
1	B	83	51% 27% • 19%
1	C	83	52% 25% • 19%
1	D	83	42% 37% • 19%
1	E	83	45% 29% • 23%
1	F	83	48% 31% • 19%
1	G	83	48% 30% • 19%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	H	83	 39% 40% .. 19%
1	I	83	 45% 35% 6% 14%
1	J	83	 41% 30% 6% 23%
1	K	83	 42% 35% . 22%
1	L	83	 55% 20% 5% 19%
1	M	83	 46% 33% . 19%
1	N	83	 41% 37% . 20%
1	O	83	 46% 35% . 16%
1	P	83	 37% 40% . 22%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8847 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 probable receptor YhhM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	66	Total 527	C 336	N 94	O 95	Se 2	0	0	0
1	B	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	C	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	D	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	E	64	Total 515	C 330	N 91	O 92	Se 2	0	0	0
1	F	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	G	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	H	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	I	71	Total 576	C 363	N 106	O 105	Se 2	0	0	0
1	J	64	Total 510	C 327	N 91	O 90	Se 2	0	0	0
1	K	65	Total 519	C 332	N 92	O 93	Se 2	0	0	0
1	L	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	M	67	Total 531	C 338	N 95	O 96	Se 2	0	0	0
1	N	66	Total 526	C 335	N 94	O 95	Se 2	0	0	0
1	O	70	Total 558	C 352	N 102	O 102	Se 2	0	0	0
1	P	65	Total 485	C 310	N 85	O 88	Se 2	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	120	LEU	-	EXPRESSION TAG	UNP D3RD65
B	120	LEU	-	EXPRESSION TAG	UNP D3RD65
C	120	LEU	-	EXPRESSION TAG	UNP D3RD65
D	120	LEU	-	EXPRESSION TAG	UNP D3RD65
E	120	LEU	-	EXPRESSION TAG	UNP D3RD65
F	120	LEU	-	EXPRESSION TAG	UNP D3RD65
G	120	LEU	-	EXPRESSION TAG	UNP D3RD65
H	120	LEU	-	EXPRESSION TAG	UNP D3RD65
I	120	LEU	-	EXPRESSION TAG	UNP D3RD65
J	120	LEU	-	EXPRESSION TAG	UNP D3RD65
K	120	LEU	-	EXPRESSION TAG	UNP D3RD65
L	120	LEU	-	EXPRESSION TAG	UNP D3RD65
M	120	LEU	-	EXPRESSION TAG	UNP D3RD65
N	120	LEU	-	EXPRESSION TAG	UNP D3RD65
O	120	LEU	-	EXPRESSION TAG	UNP D3RD65
P	120	LEU	-	EXPRESSION TAG	UNP D3RD65

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	29	Total O 29 29	0	0
2	B	27	Total O 27 27	0	0
2	C	19	Total O 19 19	0	0
2	D	23	Total O 23 23	0	0
2	E	22	Total O 22 22	0	0
2	F	21	Total O 21 21	0	0
2	G	19	Total O 19 19	0	0
2	H	23	Total O 23 23	0	0
2	I	29	Total O 29 29	0	0
2	J	28	Total O 28 28	0	0
2	K	25	Total O 25 25	0	0

*Continued on next page...*

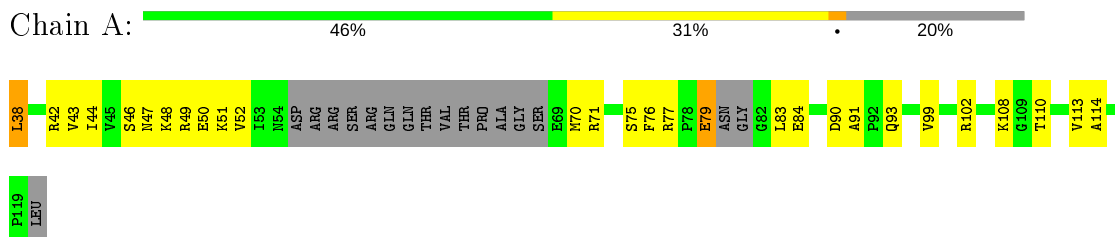
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
2	L	24	Total O 24 24	0	0
2	M	22	Total O 22 22	0	0
2	N	27	Total O 27 27	0	0
2	O	32	Total O 32 32	0	0
2	P	13	Total O 13 13	0	0

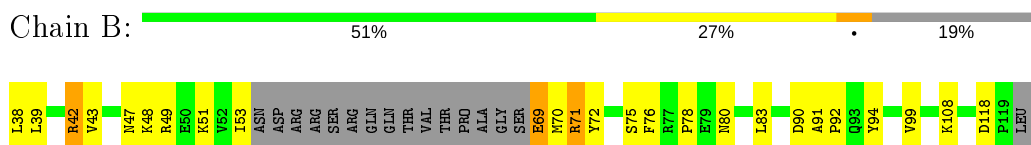
### 3 Residue-property plots [i](#)

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. 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. 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.

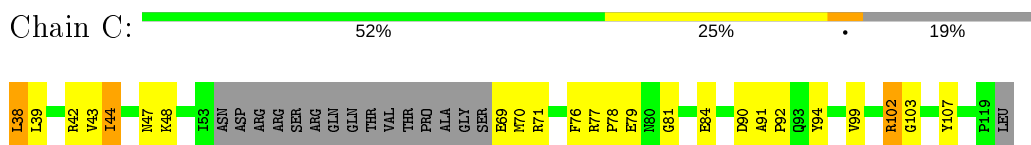
- Molecule 1: probable receptor YhhM



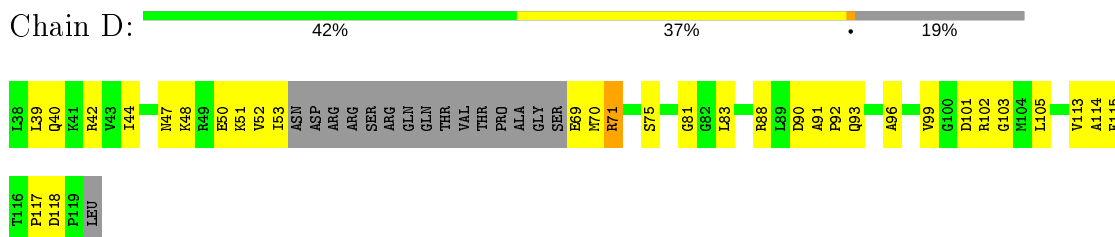
- Molecule 1: probable receptor YhhM



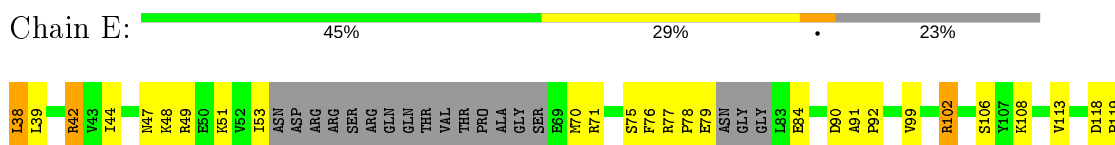
- Molecule 1: probable receptor YhhM



- Molecule 1: probable receptor YhhM

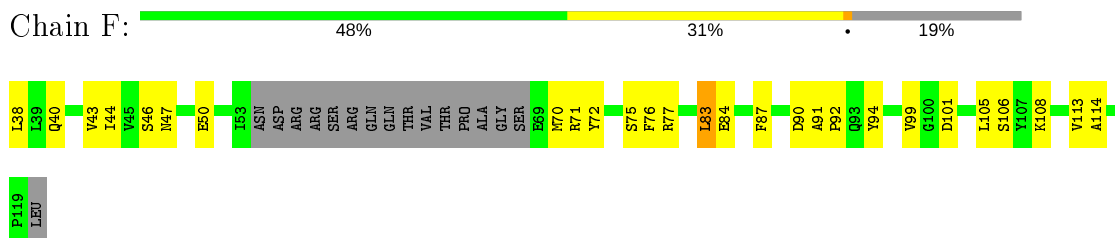


- Molecule 1: probable receptor YhhM

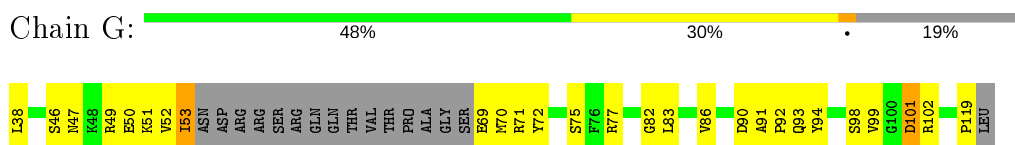


LEU

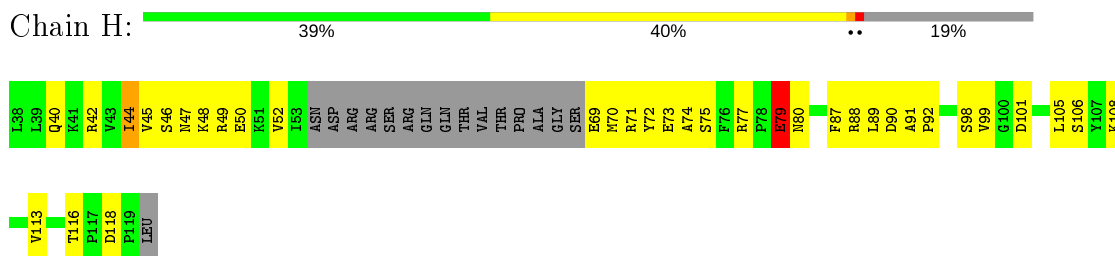
- Molecule 1: probable receptor YhhM



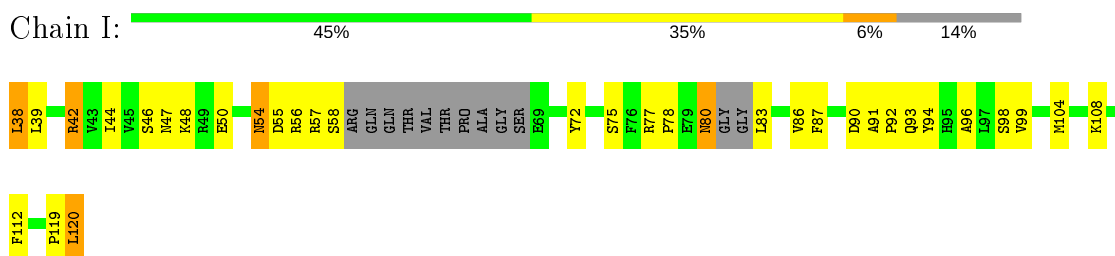
- Molecule 1: probable receptor YhhM



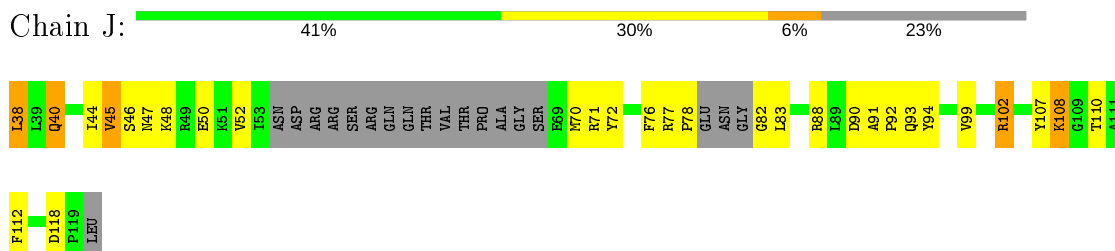
- Molecule 1: probable receptor YhhM



- Molecule 1: probable receptor YhhM

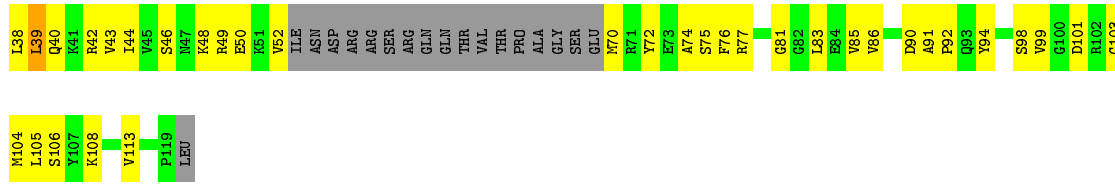


- Molecule 1: probable receptor YhhM









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	58.71Å 93.61Å 137.34Å 90.00° 97.31° 90.00°	Depositor
Resolution (Å)	49.45 – 2.60 49.45 – 2.60	Depositor EDS
% Data completeness (in resolution range)	83.9 (49.45-2.60) 96.6 (49.45-2.60)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.92 (at 2.61Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.225 , 0.284 0.294 , 0.304	Depositor DCC
$R_{free}$ test set	3651 reflections (4.33%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.8	Xtrriage
Anisotropy	0.178	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	8847	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.61% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.41	0/534	0.65	0/712
1	B	0.40	0/539	0.66	0/720
1	C	0.38	0/539	0.65	0/720
1	D	0.40	0/539	0.66	0/720
1	E	0.39	0/522	0.64	0/696
1	F	0.40	0/539	0.66	0/720
1	G	0.39	0/539	0.62	0/720
1	H	0.37	0/539	0.60	0/720
1	I	0.41	0/583	0.66	0/776
1	J	0.42	0/517	0.64	0/689
1	K	0.38	0/526	0.66	0/701
1	L	0.39	0/539	0.66	0/720
1	M	0.39	0/539	0.65	0/720
1	N	0.39	0/533	0.63	0/711
1	O	0.39	0/566	0.64	0/756
1	P	0.33	0/493	0.55	0/664
All	All	0.39	0/8586	0.64	0/11465

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 [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	527	0	525	31	0
1	B	531	0	529	23	0
1	C	531	0	529	25	0
1	D	531	0	529	33	0
1	E	515	0	516	21	0
1	F	531	0	529	25	0
1	G	531	0	529	23	0
1	H	531	0	529	35	1
1	I	576	0	574	37	0
1	J	510	0	513	35	0
1	K	519	0	519	28	0
1	L	531	0	529	20	0
1	M	531	0	529	31	0
1	N	526	0	523	30	1
1	O	558	0	552	21	0
1	P	485	0	465	25	0
2	A	29	0	0	0	0
2	B	27	0	0	2	0
2	C	19	0	0	1	0
2	D	23	0	0	4	0
2	E	22	0	0	0	0
2	F	21	0	0	1	0
2	G	19	0	0	1	0
2	H	23	0	0	2	0
2	I	29	0	0	2	0
2	J	28	0	0	2	0
2	K	25	0	0	3	0
2	L	24	0	0	2	0
2	M	22	0	0	0	0
2	N	27	0	0	1	0
2	O	32	0	0	3	0
2	P	13	0	0	1	0
All	All	8847	0	8419	383	1

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 23.

The worst 5 of 383 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:C:70:MSE:HE1	1:F:99:VAL:HB	1.32	1.10
1:I:42:ARG:HB3	1:I:42:ARG:HH11	1.18	1.05

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:M:51:LYS:HE3	1:M:71:ARG:NH1	1.82	0.93
1:K:50:GLU:HG2	1:N:99:VAL:HG21	1.52	0.92
1:D:99:VAL:HB	1:G:70:MSE:HE1	1.53	0.91

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
1:H:116:THR:OG1	1:N:113:VAL:O[2_647]	1.98	0.22

## 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	60/83 (72%)	55 (92%)	4 (7%)	1 (2%)	9 18
1	B	63/83 (76%)	59 (94%)	3 (5%)	1 (2%)	9 19
1	C	63/83 (76%)	61 (97%)	2 (3%)	0	100 100
1	D	63/83 (76%)	58 (92%)	4 (6%)	1 (2%)	9 19
1	E	58/83 (70%)	56 (97%)	1 (2%)	1 (2%)	9 18
1	F	63/83 (76%)	57 (90%)	5 (8%)	1 (2%)	9 19
1	G	63/83 (76%)	57 (90%)	5 (8%)	1 (2%)	9 19
1	H	63/83 (76%)	59 (94%)	3 (5%)	1 (2%)	9 19
1	I	65/83 (78%)	59 (91%)	6 (9%)	0	100 100
1	J	58/83 (70%)	54 (93%)	2 (3%)	2 (3%)	3 5
1	K	59/83 (71%)	54 (92%)	4 (7%)	1 (2%)	9 18
1	L	63/83 (76%)	58 (92%)	3 (5%)	2 (3%)	4 6
1	M	63/83 (76%)	61 (97%)	2 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	N	60/83 (72%)	56 (93%)	4 (7%)	0	100	100
1	O	66/83 (80%)	60 (91%)	5 (8%)	1 (2%)	10	21
1	P	61/83 (74%)	52 (85%)	7 (12%)	2 (3%)	4	6
All	All	991/1328 (75%)	916 (92%)	60 (6%)	15 (2%)	10	21

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	LEU
1	B	83	LEU
1	F	83	LEU
1	G	83	LEU
1	K	83	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	56/69 (81%)	53 (95%)	3 (5%)	22	44
1	B	56/69 (81%)	53 (95%)	3 (5%)	22	44
1	C	56/69 (81%)	53 (95%)	3 (5%)	22	44
1	D	56/69 (81%)	55 (98%)	1 (2%)	59	80
1	E	55/69 (80%)	51 (93%)	4 (7%)	14	28
1	F	56/69 (81%)	56 (100%)	0	100	100
1	G	56/69 (81%)	53 (95%)	3 (5%)	22	44
1	H	56/69 (81%)	54 (96%)	2 (4%)	35	61
1	I	62/69 (90%)	56 (90%)	6 (10%)	8	15
1	J	54/69 (78%)	50 (93%)	4 (7%)	13	28
1	K	55/69 (80%)	53 (96%)	2 (4%)	35	61
1	L	56/69 (81%)	53 (95%)	3 (5%)	22	44
1	M	56/69 (81%)	52 (93%)	4 (7%)	14	29

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	N	56/69 (81%)	54 (96%)	2 (4%)	35	61
1	O	59/69 (86%)	55 (93%)	4 (7%)	16	32
1	P	48/69 (70%)	47 (98%)	1 (2%)	53	77
All	All	893/1104 (81%)	848 (95%)	45 (5%)	24	47

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	42	ARG
1	J	38	LEU
1	O	71	ARG
1	I	80	ASN
1	J	40	GLN

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

Mol	Chain	Res	Type
1	I	93	GLN
1	J	40	GLN
1	L	47	ASN
1	I	80	ASN
1	K	95	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 carbohydrates 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

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.