



# wwPDB X-ray Structure Validation Summary Report i

Feb 27, 2014 – 02:14 PM GMT

PDB ID : 4B8C  
Title : nuclease module of the yeast Ccr4-Not complex  
Authors : Basquin, J.; Conti, E.  
Deposited on : 2012-08-26  
Resolution : 3.41 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>

---

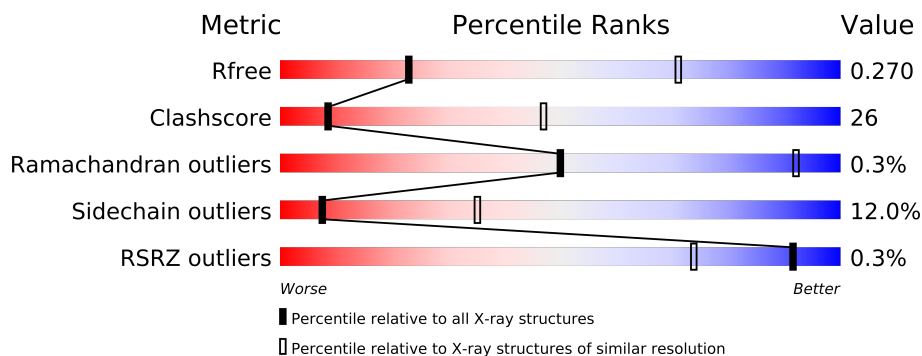
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.41 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	1013 (3.54-3.30)
Clashscore	79885	1270 (3.54-3.30)
Ramachandran outliers	78287	1232 (3.54-3.30)
Sidechain outliers	78261	1232 (3.54-3.30)
RSRZ outliers	66119	1013 (3.54-3.30)

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 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	A	288	
1	C	288	
1	E	288	
1	F	288	
2	B	249	
2	G	249	
2	H	249	
2	I	249	
3	D	727	
3	J	727	
3	K	727	
3	L	727	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 24298 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 POLY(A) RIBONUCLEASE POP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	267	Total	C	N	O	S	0	0	0
			2170	1409	346	404	11			
1	C	267	Total	C	N	O	S	0	0	0
			2174	1412	347	404	11			
1	E	267	Total	C	N	O	S	0	0	0
			2174	1411	347	405	11			
1	F	267	Total	C	N	O	S	0	0	0
			2170	1409	346	404	11			

- Molecule 2 is a protein called GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	233	Total	C	N	O	S	0	0	0
			1869	1221	309	333	6			
2	G	233	Total	C	N	O	S	0	0	0
			1869	1221	309	333	6			
2	H	233	Total	C	N	O	S	0	0	0
			1869	1221	309	333	6			
2	I	233	Total	C	N	O	S	0	0	0
			1869	1221	309	333	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	752	ARG	-	EXPRESSION TAG	UNP P25655
B	753	SER	-	EXPRESSION TAG	UNP P25655
B	754	MET	-	EXPRESSION TAG	UNP P25655
G	752	ARG	-	EXPRESSION TAG	UNP P25655
G	753	SER	-	EXPRESSION TAG	UNP P25655
G	754	MET	-	EXPRESSION TAG	UNP P25655
H	752	ARG	-	EXPRESSION TAG	UNP P25655
H	753	SER	-	EXPRESSION TAG	UNP P25655

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	754	MET	-	EXPRESSION TAG	UNP P25655
I	752	ARG	-	EXPRESSION TAG	UNP P25655
I	753	SER	-	EXPRESSION TAG	UNP P25655
I	754	MET	-	EXPRESSION TAG	UNP P25655

- Molecule 3 is a protein called GLUCOSE-REPRESSIBLE ALCOHOL DEHYDROGENASE TRANSCRIPTIONAL EFFECTOR.

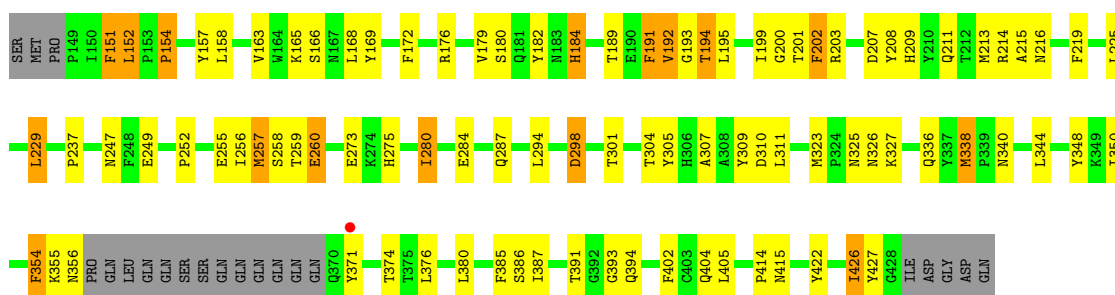
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	318	Total	C	N	O	S	0	0	0
			2456	1599	387	459	11			
3	J	210	Total	C	N	O	S	0	0	0
			1604	1034	266	298	6			
3	K	318	Total	C	N	O	S	0	0	0
			2470	1608	392	459	11			
3	L	210	Total	C	N	O	S	0	0	0
			1604	1034	266	298	6			

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

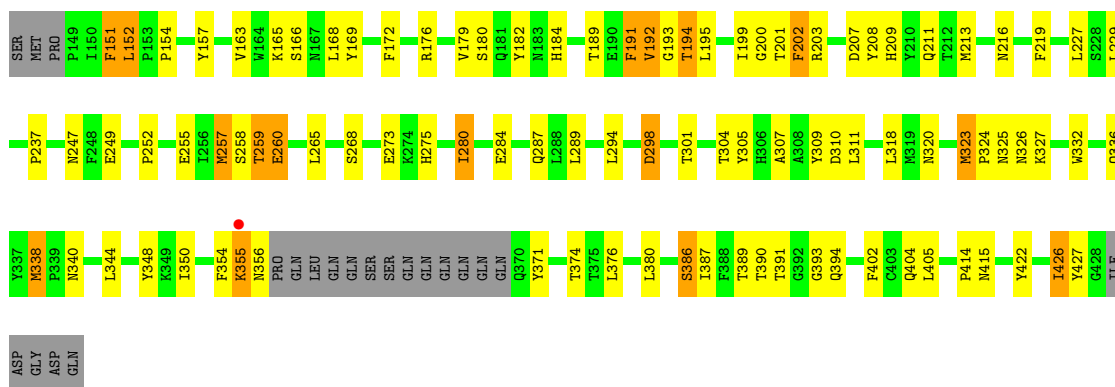
#### • Molecule 1: POLY(A) RIBONUCLEASE POP2

Chain A:



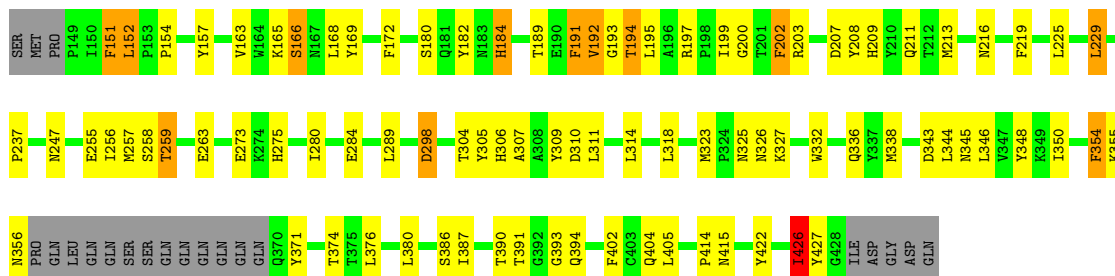
#### • Molecule 1: POLY(A) RIBONUCLEASE POP2

Chain C:



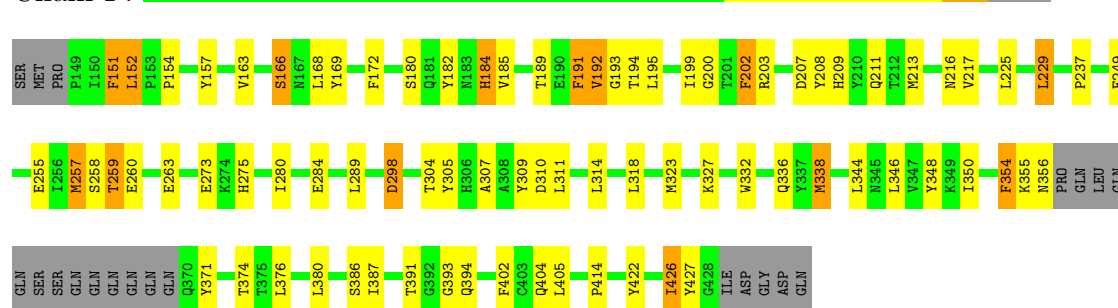
#### • Molecule 1: POLY(A) RIBONUCLEASE POP2

Chain E:



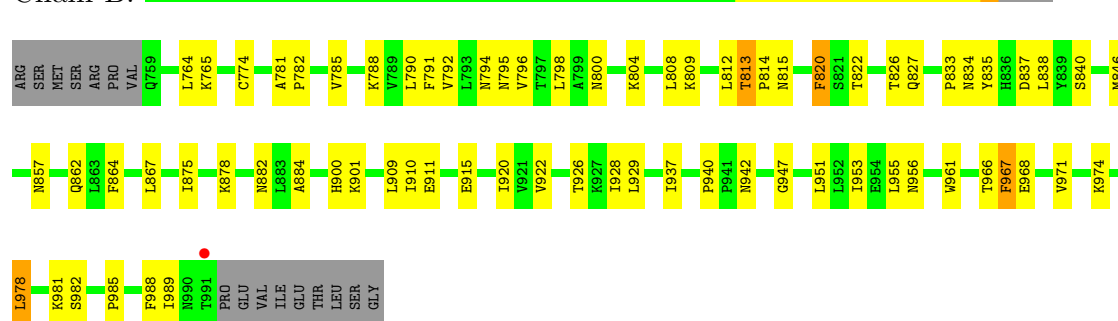
- Molecule 1: POLY(A) RIBONUCLEASE POP2

Chain F:



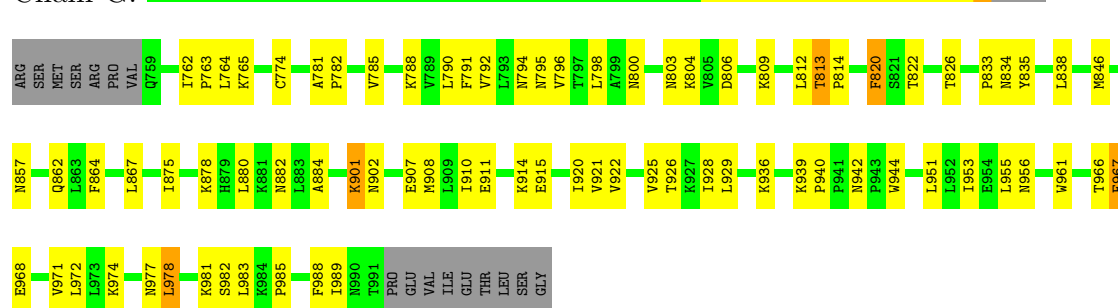
- Molecule 2: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

Chain B:



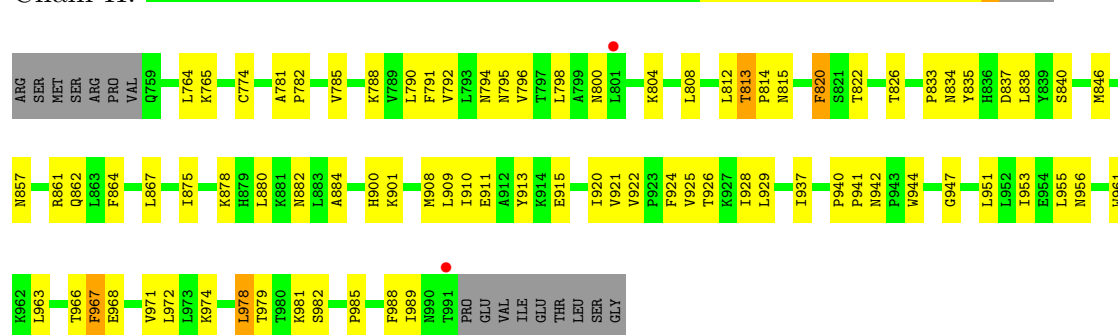
- Molecule 2: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

Chain G:



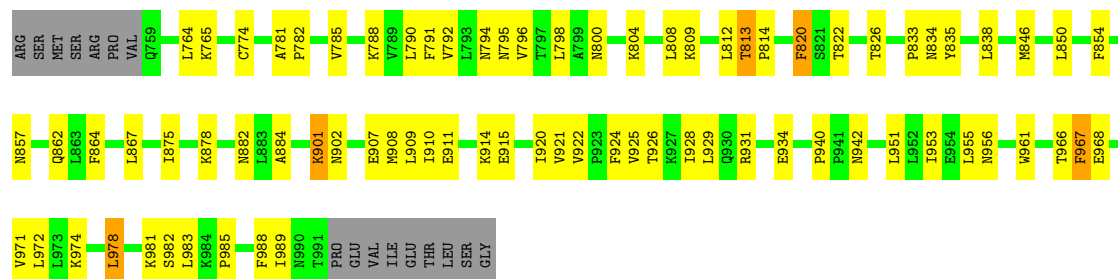
- Molecule 2: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

Chain H:



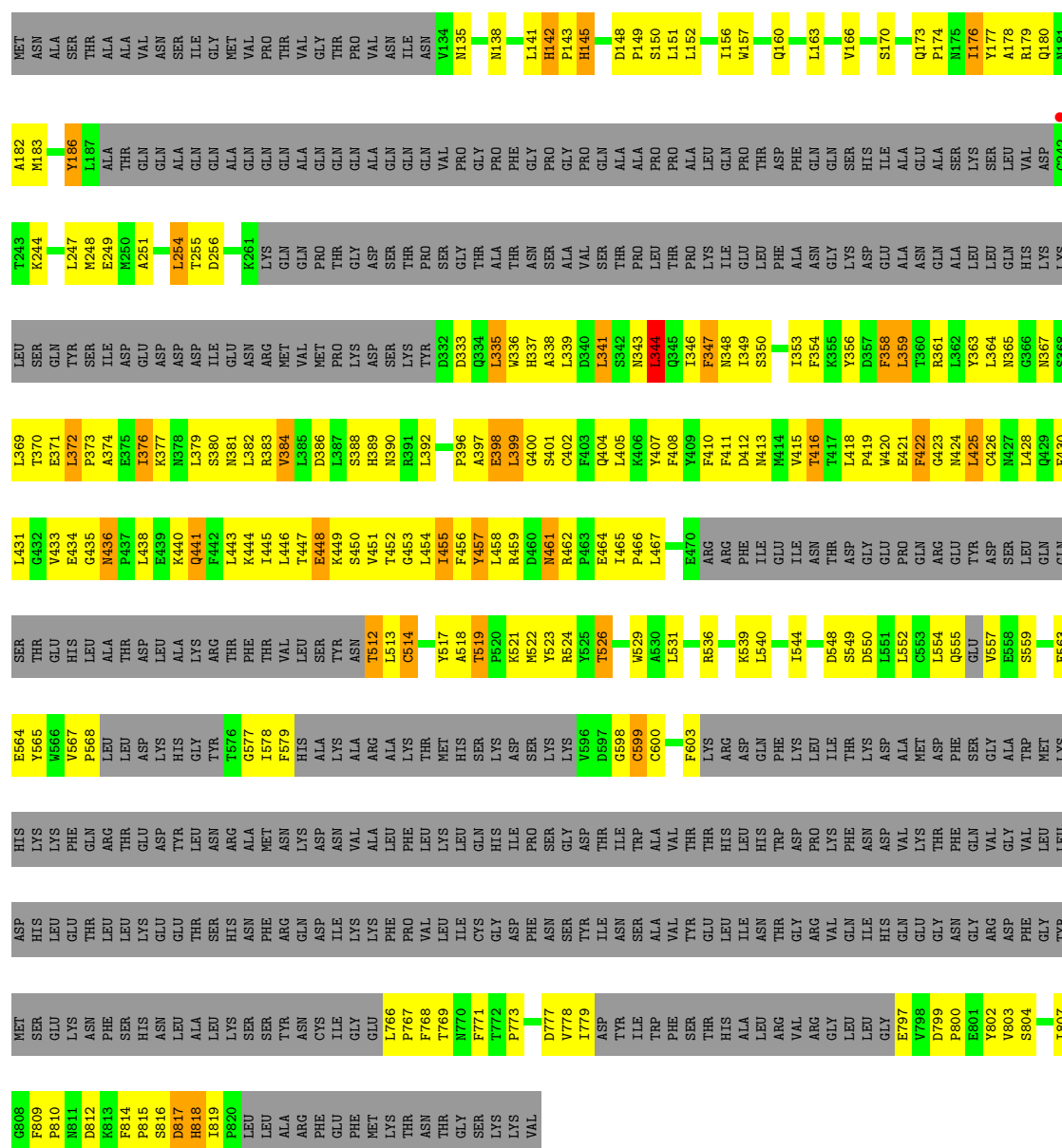
- Molecule 2: GENERAL NEGATIVE REGULATOR OF TRANSCRIPTION SUBUNIT 1

## Chain I:



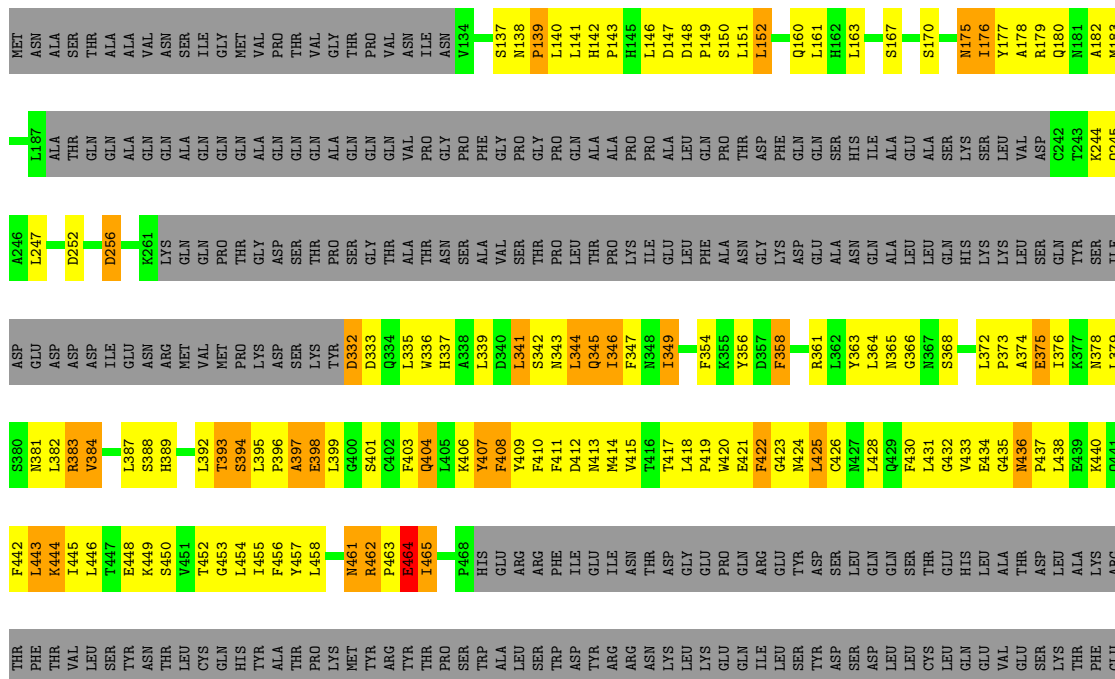
● Molecule 3: GLUCOSE-REPRESSIBLE ALCOHOL DEHYDROGENASE TRANSCRIPTIONAL EFFECTOR

## Chain D:









SER	TYP	LEU	LYS	GLY
LYS	MET	ASP	HIS	TRP
PHE	SER	HIS	LYS	VAL
ILE	GLU	LEU	PHE	PRO
PHE	ASN	THR	GLN	LEU
PRO	PHE	LEU	ARG	LEU
ASN	SER	LEU	THR	ASP
ASP	HIS	LYS	GLU	LYS
LYS	ASN	GLU	ASP	HIS
PHE	LEU	GLU	TYR	GLY
PRO	ALA	THR	LEU	TYR
SER	LEU	SER	ASN	THR
ASP	LYS	HIS	ARG	GLY
HIS	SER	ASN	ALA	ILE
ILE	SER	PHE	MET	PHE
PRO	TYR	ARG	ASN	HIS
LEU	ASN	GLN	LYS	ALA
LEU	CYS	ASP	ASP	LYS
ALA	ILE	ILE	ASN	ALA
ARG	GLY	LYS	VAL	ARG
PHE	GLU	LYS	ALA	ALA
GLU	LEU	PHE	LEU	LYS
PHE	PRO	PRO	PHE	THR
MET	PHE	VAL	LEU	MET
LYS	THR	ILE	LYS	SER
ASN	PHE	CYS	GLN	LYS
THR	THR	GLY	HIS	ASP
GLY	PRO	ASP	ILE	SER
SER	SER	PHE	PRO	LYS
LYS	PHE	ASN	SER	LYS
VAL	THR	SER	GLY	VAL
	ASP	TYR	ASP	ASP
	VAL	ILE	THR	CYS
	ILE	ASN	ILE	GLY
	ASP	SER	TRP	CYS
	TYR	ALA	ALA	ILE
	ILE	VAL	VAL	PHE
	TRP	TYR	THR	PHE
	PHE	GLU	THR	LYS
	SER	LEU	HIS	ARG
	THR	ILE	LEU	ASP
	HIS	ASN	HIS	GLN
	ALA	THR	TRP	PHE
	LEU	GLY	ASP	LYS
	ARG	ARG	PRO	LEU
	VAL	VAL	LYS	ILE
	ARG	GLN	PHE	THR
	GLY	ILE	ASN	LYS
	LEU	HIS	ASP	ALA
	LEU	GLN	VAL	MET
	GLY	GLY	LYS	ASP
	GLU	GLY	THR	PHE
	VAL	ASN	PHE	SER
	ASP	GLY	GLN	THR
	PRO	ARG	VAL	GLY
	GLU	ASP	GLY	ALA
	TYR	PHE	VAL	TRP
	VAL	GLY	LEU	MET

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	122.65Å 122.91Å 126.42Å 89.47° 89.74° 64.22°	Depositor
Resolution (Å)	47.95 – 3.41 48.23 – 3.28	Depositor EDS
% Data completeness (in resolution range)	97.1 (47.95-3.41) 97.4 (48.23-3.28)	Depositor EDS
$R_{merge}$	0.01	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 3.25Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE: 1.7.1_743)	Depositor
R, $R_{free}$	0.233 , 0.270 0.233 , 0.270	Depositor DCC
$R_{free}$ test set	4940 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	77.9	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 51.2	EDS
Estimated twinning fraction	0.026 for -h,-k,l 0.018 for k,h,-l 0.022 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 98830 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	24298	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.27% 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	A	0.56	0/2231	0.67	0/3028
1	C	0.54	0/2235	0.67	0/3032
1	E	0.50	0/2235	0.66	0/3033
1	F	0.52	0/2231	0.66	0/3028
2	B	0.43	0/1908	0.58	0/2588
2	G	0.42	0/1908	0.57	0/2588
2	H	0.43	0/1908	0.58	0/2588
2	I	0.44	0/1908	0.58	0/2588
3	D	0.53	0/2520	0.76	2/3446 (0.1%)
3	J	0.62	1/1640 (0.1%)	0.94	4/2240 (0.2%)
3	K	0.54	1/2535 (0.0%)	0.76	1/3463 (0.0%)
3	L	0.63	2/1640 (0.1%)	0.94	3/2240 (0.1%)
All	All	0.52	4/24899 (0.0%)	0.71	10/33862 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
1	E	0	1
3	D	0	1
3	J	0	1
3	K	0	1
3	L	0	2
All	All	0	8

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	462	ARG	C-N	8.80	1.50	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	J	462	ARG	C-N	8.53	1.50	1.34
3	L	332	ASP	CA-CB	5.73	1.66	1.53
3	K	426	CYS	CB-SG	5.22	1.91	1.82

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	L	462	ARG	O-C-N	13.09	145.96	121.10
3	J	462	ARG	O-C-N	12.95	145.70	121.10
3	J	462	ARG	CA-C-N	-9.27	91.15	117.10
3	L	462	ARG	CA-C-N	-9.18	91.41	117.10
3	J	397	ALA	N-CA-C	-6.26	94.10	111.00

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	194	THR	Peptide
1	C	194	THR	Peptide
3	D	344	LEU	Peptide
1	E	194	THR	Peptide
3	J	332	ASP	Peptide

## 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	A	2170	0	2083	87	0
1	C	2174	0	2094	94	0
1	E	2174	0	2089	86	1
1	F	2170	0	2083	74	1
2	B	1869	0	1936	49	0
2	G	1869	0	1936	51	1
2	H	1869	0	1936	62	0
2	I	1869	0	1936	50	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	2456	0	2268	226	0
3	J	1604	0	1505	145	1
3	K	2470	0	2293	204	0
3	L	1604	0	1505	157	1
All	All	24298	0	23664	1240	3

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 26.

The worst 5 of 1240 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:D:176:ILE:HD11	3:D:343:ASN:HD22	1.20	1.04
3:J:422:PHE:HD1	3:J:422:PHE:O	1.42	1.00
3:K:176:ILE:HD11	3:K:343:ASN:HD22	1.21	1.00
3:K:567:VAL:HG13	3:K:568:PRO:HD3	1.41	0.99
3:D:424:ASN:HD21	3:D:807:ILE:HG21	1.29	0.98

All (3) 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:F:166:SER:OG	3:L:146:LEU:O[1_554]	2.08	0.12
2:G:835:TYR:OH	2:I:800:ASN:CA[1_465]	2.10	0.10
1:E:166:SER:OG	3:J:146:LEU:O[1_455]	2.14	0.06

## 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	263/288 (91%)	250 (95%)	12 (5%)	1 (0%)	43	90
1	C	263/288 (91%)	250 (95%)	12 (5%)	1 (0%)	43	90
1	E	263/288 (91%)	249 (95%)	13 (5%)	1 (0%)	43	90
1	F	263/288 (91%)	250 (95%)	13 (5%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	231/249 (93%)	222 (96%)	9 (4%)	0	100	100
2	G	231/249 (93%)	222 (96%)	9 (4%)	0	100	100
2	H	231/249 (93%)	221 (96%)	10 (4%)	0	100	100
2	I	231/249 (93%)	222 (96%)	9 (4%)	0	100	100
3	D	300/727 (41%)	253 (84%)	46 (15%)	1 (0%)	50	92
3	J	204/727 (28%)	168 (82%)	34 (17%)	2 (1%)	22	77
3	K	300/727 (41%)	253 (84%)	46 (15%)	1 (0%)	50	92
3	L	204/727 (28%)	171 (84%)	31 (15%)	2 (1%)	22	77
All	All	2984/5056 (59%)	2731 (92%)	244 (8%)	9 (0%)	50	92

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	380	SER
3	K	380	SER
3	J	464	GLU
3	J	465	ILE
3	L	465	ILE

### 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	241/264 (91%)	216 (90%)	25 (10%)	10	44
1	C	242/264 (92%)	218 (90%)	24 (10%)	11	47
1	E	242/264 (92%)	220 (91%)	22 (9%)	14	52
1	F	241/264 (91%)	218 (90%)	23 (10%)	12	49
2	B	211/231 (91%)	199 (94%)	12 (6%)	29	75
2	G	211/231 (91%)	198 (94%)	13 (6%)	26	72
2	H	211/231 (91%)	199 (94%)	12 (6%)	29	75
2	I	211/231 (91%)	198 (94%)	13 (6%)	26	72
3	D	258/648 (40%)	206 (80%)	52 (20%)	2	9

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	J	165/648 (26%)	129 (78%)	36 (22%)	1	7
3	K	261/648 (40%)	208 (80%)	53 (20%)	2	9
3	L	165/648 (26%)	130 (79%)	35 (21%)	1	8
All	All	2659/4572 (58%)	2339 (88%)	320 (12%)	7	35

5 of 320 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	209	HIS
2	H	961	TRP
3	L	345	GLN
1	F	298	ASP
2	G	822	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	424	ASN
3	L	175	ASN
3	K	413	ASN
3	D	343	ASN
3	K	352	ASN

### 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	A	267/288 (92%)	-0.18	1 (0%) 90 68	36, 58, 98, 121	0
1	C	267/288 (92%)	-0.17	1 (0%) 90 68	36, 58, 99, 121	0
1	E	267/288 (92%)	-0.22	0 100 100	35, 58, 100, 121	0
1	F	267/288 (92%)	-0.20	0 100 100	36, 59, 98, 124	0
2	B	233/249 (93%)	-0.08	1 (0%) 90 68	48, 79, 115, 146	0
2	G	233/249 (93%)	-0.14	0 100 100	49, 81, 112, 146	0
2	H	233/249 (93%)	-0.03	2 (0%) 81 47	48, 80, 115, 145	0
2	I	233/249 (93%)	-0.11	0 100 100	50, 81, 114, 145	0
3	D	318/727 (43%)	-0.02	1 (0%) 91 74	50, 87, 136, 160	0
3	J	210/727 (28%)	-0.26	0 100 100	35, 61, 101, 127	0
3	K	318/727 (43%)	0.03	4 (1%) 74 37	50, 87, 136, 160	0
3	L	210/727 (28%)	-0.22	0 100 100	35, 62, 102, 131	0
All	All	3056/5056 (60%)	-0.13	10 (0%) 91 74	35, 72, 115, 160	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	991	THR	2.8
3	K	597	ASP	2.8
2	H	991	THR	2.6
2	H	801	LEU	2.4
1	C	355	LYS	2.3

### 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.