



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 29, 2017 – 08:33 PM EST

PDB ID : 4D28  
Title : Crystal structure of the kinase domain of CIPK24/SOS2  
Authors : Gonzalez-Rubio, J.M.; Chaves-Sanjuan, A.; Sanchez-Barrena, M.J.; Albert, A.  
Deposited on : unknown  
Resolution : 3.30 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

<http://wwpdb.org/validation/2016/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) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

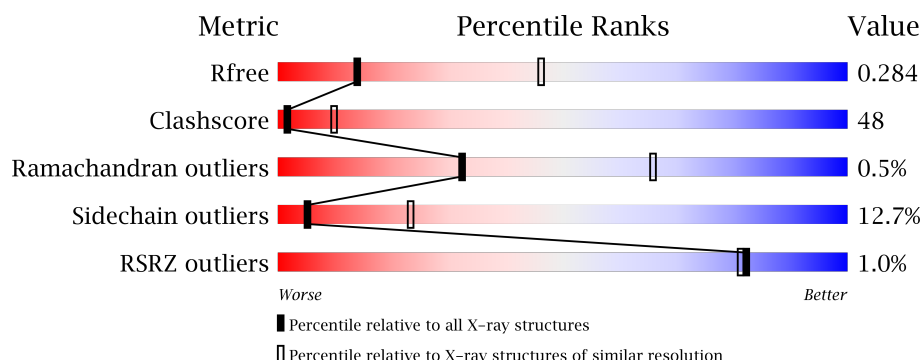
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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))
$R_{free}$	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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. 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  $\leq 5\%$ . 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	446	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>27%</span> <span>31%</span> <span>•</span> <span>39%</span> </div> </div>
1	B	446	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, orange, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>25%</span> <span>31%</span> <span>7%</span> <span>38%</span> </div> </div>
1	C	446	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, orange, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>24%</span> <span>33%</span> <span>5%</span> <span>37%</span> </div> </div>
1	D	446	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>21%</span> <span>35%</span> <span>5%</span> <span>39%</span> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8821 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 CBL-INTERACTING SERINE/THREONINE-PROTEIN KINASE 24.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2192	1404	382	399	7			
1	B	277	Total	C	N	O	S	0	0	0
			2217	1419	388	403	7			
1	C	279	Total	C	N	O	S	0	0	0
			2234	1430	390	407	7			
1	D	271	Total	C	N	O	S	0	0	0
			2178	1397	379	395	7			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	81	LYS	PRO	engineered mutation	UNP Q9LDI3
A	107	LYS	GLU	engineered mutation	UNP Q9LDI3
A	109	ASP	SER	engineered mutation	UNP Q9LDI3
A	127	SER	CYS	engineered mutation	UNP Q9LDI3
A	167	ASN	ARG	conflict	UNP Q9LDI3
A	168	ASP	THR	engineered mutation	UNP Q9LDI3
A	228	ASP	SER	engineered mutation	UNP Q9LDI3
A	266	LYS	LEU	engineered mutation	UNP Q9LDI3
B	81	LYS	PRO	engineered mutation	UNP Q9LDI3
B	107	LYS	GLU	engineered mutation	UNP Q9LDI3
B	109	ASP	SER	engineered mutation	UNP Q9LDI3
B	127	SER	CYS	engineered mutation	UNP Q9LDI3
B	167	ASN	ARG	conflict	UNP Q9LDI3
B	168	ASP	THR	engineered mutation	UNP Q9LDI3
B	228	ASP	SER	engineered mutation	UNP Q9LDI3
B	266	LYS	LEU	engineered mutation	UNP Q9LDI3
C	81	LYS	PRO	engineered mutation	UNP Q9LDI3
C	107	LYS	GLU	engineered mutation	UNP Q9LDI3
C	109	ASP	SER	engineered mutation	UNP Q9LDI3
C	127	SER	CYS	engineered mutation	UNP Q9LDI3

*Continued on next page...*

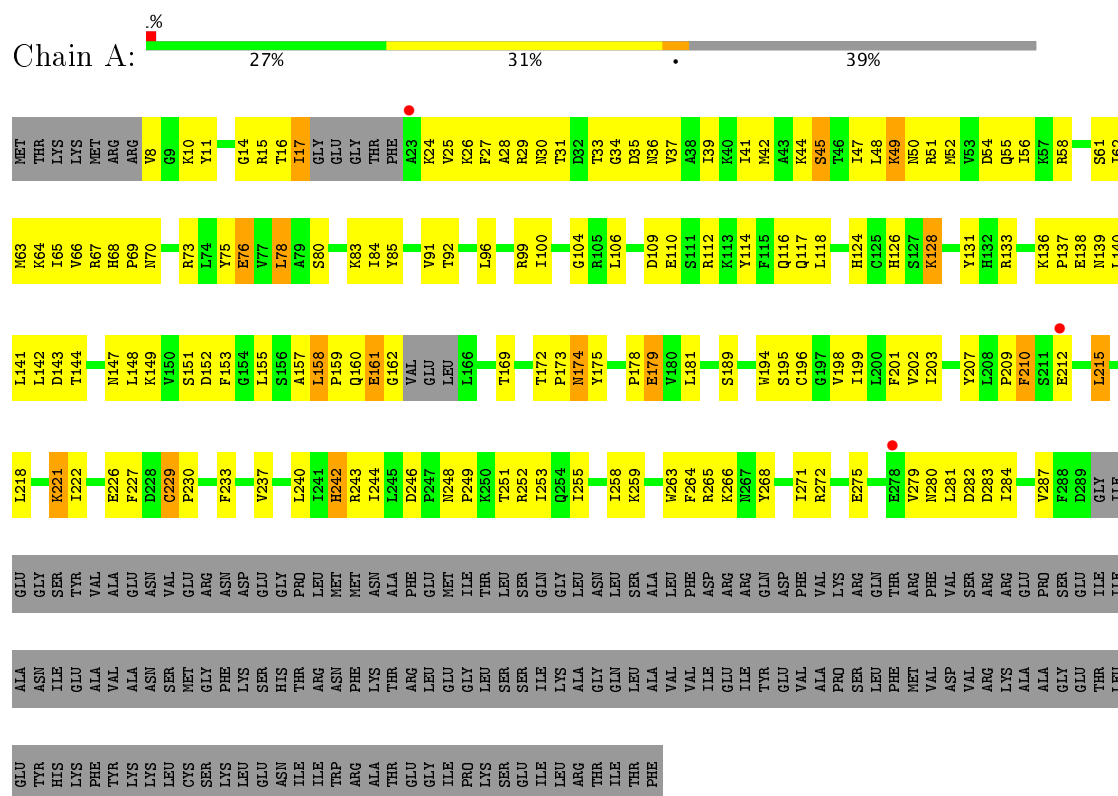
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	167	ASN	ARG	conflict	UNP Q9LDI3
C	168	ASP	THR	engineered mutation	UNP Q9LDI3
C	228	ASP	SER	engineered mutation	UNP Q9LDI3
C	266	LYS	LEU	engineered mutation	UNP Q9LDI3
D	81	LYS	PRO	engineered mutation	UNP Q9LDI3
D	107	LYS	GLU	engineered mutation	UNP Q9LDI3
D	109	ASP	SER	engineered mutation	UNP Q9LDI3
D	127	SER	CYS	engineered mutation	UNP Q9LDI3
D	167	ASN	ARG	conflict	UNP Q9LDI3
D	168	ASP	THR	engineered mutation	UNP Q9LDI3
D	228	ASP	SER	engineered mutation	UNP Q9LDI3
D	266	LYS	LEU	engineered mutation	UNP Q9LDI3

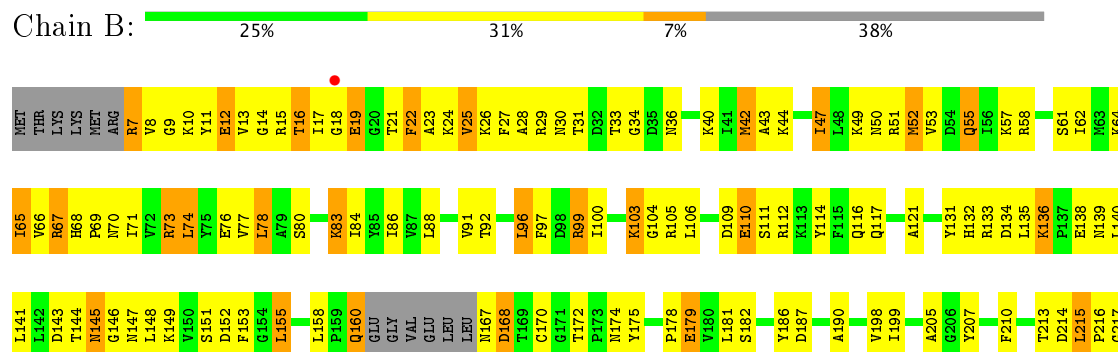
### 3 Residue-property plots

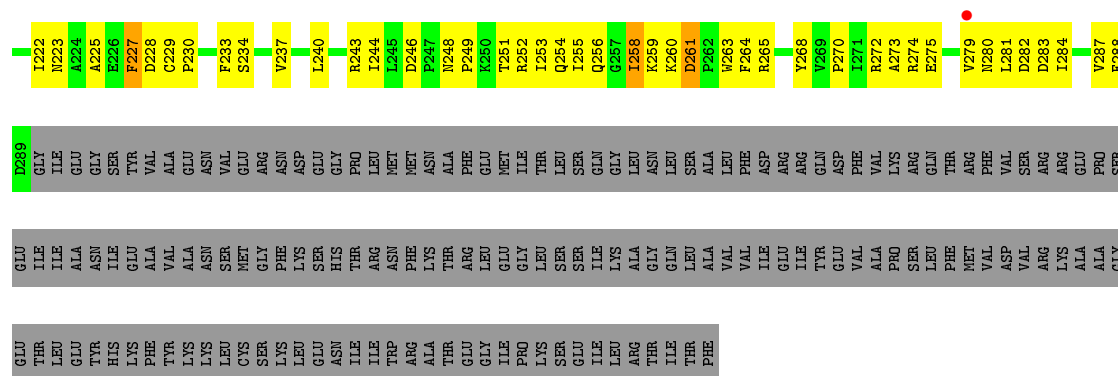
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: CBL-INTERACTING SERINE/THREONINE-PROTEIN KINASE 24

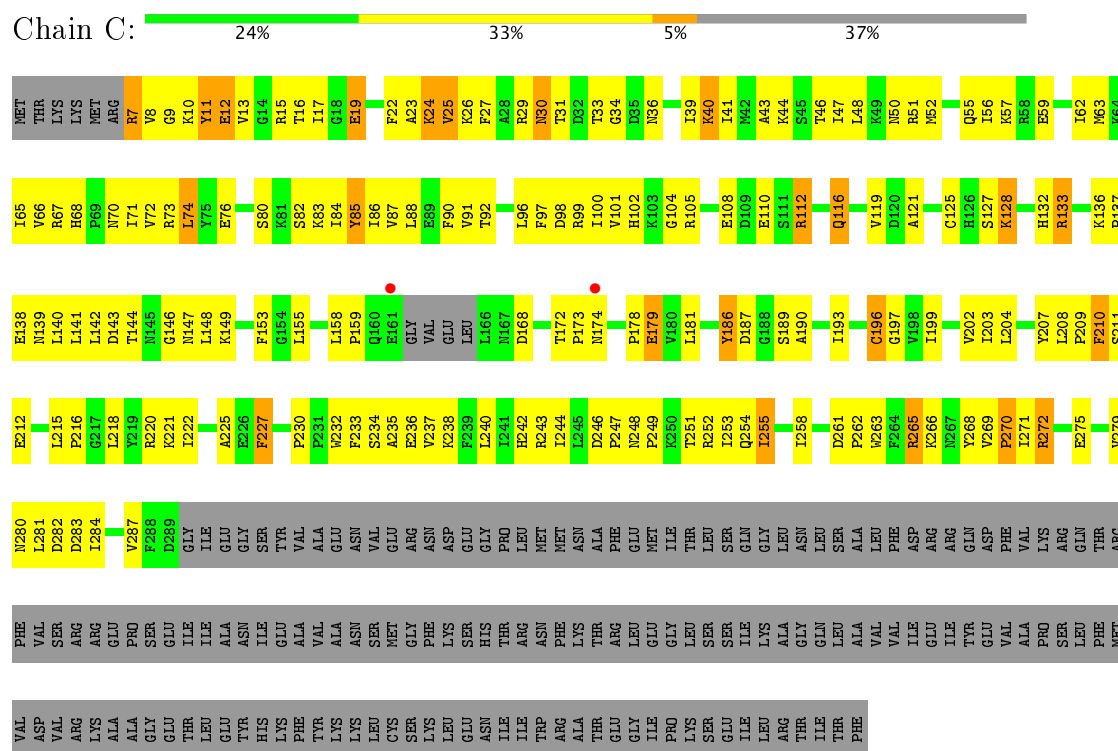


#### • Molecule 1: CBL-INTERACTING SERINE/THREONINE-PROTEIN KINASE 24

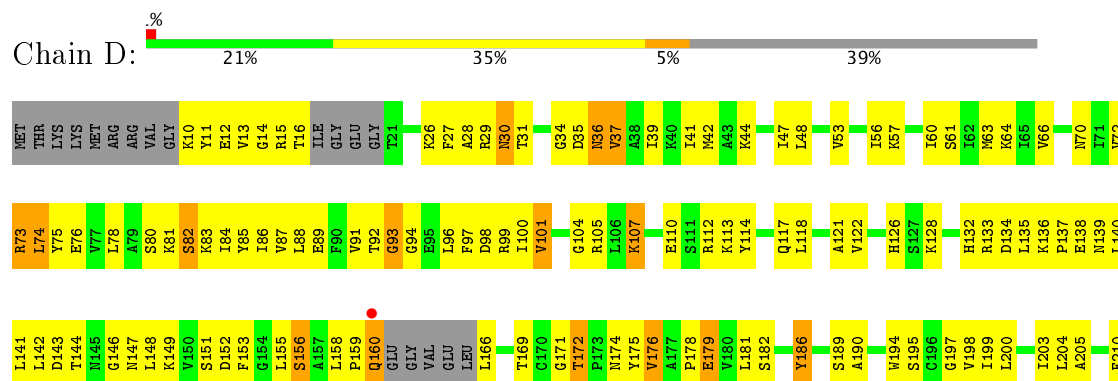




• Molecule 1: CBL-INTERACTING SERINE/THREONINE-PROTEIN KINASE 24



• Molecule 1: CBL-INTERACTING SERINE/THREONINE-PROTEIN KINASE 24



LEU	GLN	E277	S211
PHE	THR	E278	E212
MET	ARG	V279	T213
VAL	PHE	N280	D214
ASP	VAL	L281	L215
VAL	SER	D282	P216
ARG	ARG	D283	C217
LYS	ARG	I284	L218
ALA	GLU	V287	Y219
ALA	PRO	F288	R220
GLY	SER	D289	K221
GLU	GLU	GLY	I222
THR	ILE	ILE	E226
LEU	ILE	GLU	F227
GLU	ALA	GLY	P230
TYR	ASN	SER	P231
HIS	ILE	TYR	W232
LYS	GLU	VAL	F233
PHE	ALA	ALA	
TYR	VAL	GLU	E236
LYS	ALA	ASN	V237
LYS	ASN	VAL	K238
LEU	SER	GLU	F239
LEU	SER	MET	L240
CYS	MET	GLY	I241
SER	GLY	ARG	H242
LYS	PHE	ASN	R243
LEU	LYS	ASP	I244
GLU	SER	GLY	PRO
ASN	HIS	GLY	L245
ILE	THR	PRO	LEU
ILE	THR	ARG	D246
TRP	ARG	ASN	P247
ARG	ASN	MET	N248
ALA	PHE	MET	P249
ALA	LYS	ALA	K250
THR	THR	ARG	T251
GLU	ARG	GLU	R252
GLY	LEU	LEU	I253
ILE	GLU	MET	Q254
PRO	GLY	ILE	I255
LYS	LEU	THR	Q256
SER	SER	SER	G257
GLU	SER	SER	I258
ILE	ILE	GLN	K259
LEU	LYS	GLY	D261
ANG	ALA	LEU	P262
ARG	GLY	ASN	W263
ILE	GLN	LEU	F264
THR	LEU	SER	R265
THR	ALA	LEU	K266
PHE	VAL	PHE	M267
	VAL	ASP	Y268
	ILE	ARG	V269
	GLU	ARG	P270
	TYR	GLN	R274
	GLU	ASP	E275
	VAL	PHE	
	ALA	VAL	
	PRO	LYS	
	SER	ARG	

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.11Å 71.35Å 77.83Å 104.85° 100.32° 118.96°	Depositor
Resolution (Å)	70.23 – 3.30 70.23 – 3.30	Depositor EDS
% Data completeness (in resolution range)	90.8 (70.23-3.30) 76.9 (70.23-3.30)	Depositor EDS
$R_{merge}$	0.27	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.73 (at 3.33Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.271 , 0.283 0.275 , 0.284	Depositor DCC
$R_{free}$ test set	810 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	72.5	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 21.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.045 for k,h,-h-k-l 0.043 for -k,-h,l 0.069 for -h,-k,h+k+l	Xtriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	8821	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.13% 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	0.30	0/2236	0.56	0/3020
1	B	0.31	0/2263	0.56	0/3057
1	C	0.29	0/2280	0.54	0/3080
1	D	0.30	0/2223	0.54	0/3003
All	All	0.30	0/9002	0.55	0/12160

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	2192	0	2222	171	0
1	B	2217	0	2244	202	0
1	C	2234	0	2261	241	0
1	D	2178	0	2206	254	0
All	All	8821	0	8933	860	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 48.

The worst 5 of 860 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:174:ASN:OD1	1:C:212:GLU:OE2	1.60	1.15
1:D:86:ILE:HG22	1:D:88:LEU:HD11	1.17	1.14
1:B:92:THR:OG1	1:B:144:THR:HG23	1.49	1.12
1:B:7:ARG:HD2	1:B:12:GLU:OE2	1.48	1.11
1:D:175:TYR:O	1:D:194:TRP:NE1	1.83	1.10

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	268/446 (60%)	254 (95%)	14 (5%)	0	100	100
1	B	273/446 (61%)	249 (91%)	23 (8%)	1 (0%)	38	71
1	C	275/446 (62%)	263 (96%)	11 (4%)	1 (0%)	38	71
1	D	265/446 (59%)	247 (93%)	15 (6%)	3 (1%)	17	52
All	All	1081/1784 (61%)	1013 (94%)	63 (6%)	5 (0%)	32	66

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	93	GLY
1	B	23	ALA
1	D	176	VAL
1	D	267	ASN
1	C	270	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	239/388 (62%)	217 (91%)	22 (9%)	11	37
1	B	241/388 (62%)	198 (82%)	43 (18%)	2	10
1	C	243/388 (63%)	214 (88%)	29 (12%)	6	26
1	D	238/388 (61%)	210 (88%)	28 (12%)	6	26
All	All	961/1552 (62%)	839 (87%)	122 (13%)	5	23

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	182	SER
1	C	19	GLU
1	D	186	TYR
1	B	186	TYR
1	B	229	CYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	139	ASN
1	B	242	HIS
1	D	242	HIS
1	B	160	GLN
1	B	223	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 [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, 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	274/446 (61%)	-0.26	3 (1%) 80 79	40, 61, 100, 124	0
1	B	277/446 (62%)	-0.21	2 (0%) 87 87	39, 64, 125, 146	0
1	C	279/446 (62%)	-0.18	2 (0%) 87 87	44, 65, 121, 136	0
1	D	271/446 (60%)	-0.24	4 (1%) 74 70	47, 67, 109, 132	0
All	All	1101/1784 (61%)	-0.22	11 (0%) 82 81	39, 64, 116, 146	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	18	GLY	3.7
1	B	279	VAL	3.2
1	D	160	GLN	2.4
1	A	212	GLU	2.4
1	D	277	GLU	2.3

### 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 carbohydrates 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.