



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2024 – 07:55 am GMT

PDB ID : 2J04  
Title : The tau60-tau91 subcomplex of yeast transcription factor IIIC  
Authors : Mylona, A.; Fernandez-Tornero, C.; Legrand, P.; Muller, C.W.  
Deposited on : 2006-07-31  
Resolution : 3.20 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

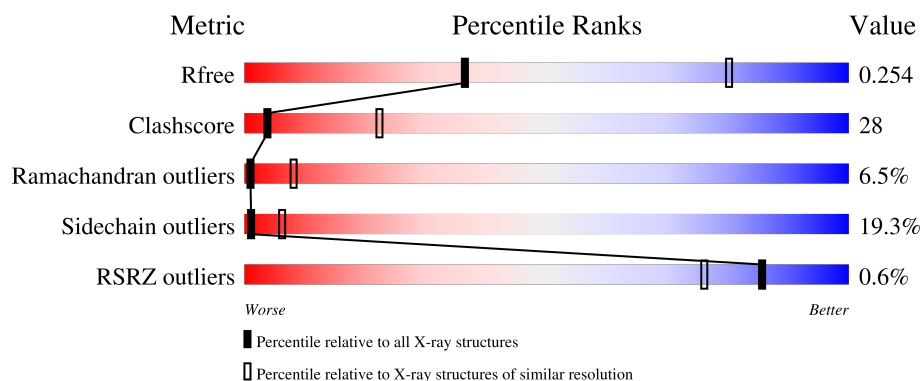
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
$R_{free}$	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (3.20-3.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	588	<div> <div>%</div> <div> <div></div> <div>52%</div> <div>34%</div> <div>11%</div> <div>.</div> </div> </div>
1	C	588	<div> <div>%</div> <div> <div></div> <div>48%</div> <div>36%</div> <div>13%</div> <div>.</div> </div> </div>
2	B	524	<div> <div></div> <div> <div>43%</div> <div>32%</div> <div>13%</div> <div>.</div> <div>10%</div> </div> </div>
2	D	524	<div> <div>%</div> <div> <div></div> <div>43%</div> <div>33%</div> <div>12%</div> <div>.</div> <div>10%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16935 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 HYPOTHETICAL PROTEIN YPL007C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	587	Total	C	N	O	S	0	0	0
			4760	3058	777	902	23			
1	C	586	Total	C	N	O	S	0	0	0
			4751	3052	775	901	23			

- Molecule 2 is a protein called YDR362CP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	472	Total	C	N	O	S	0	0	1
			3715	2384	613	703	15			
2	D	469	Total	C	N	O	S	0	0	1
			3704	2379	609	701	15			

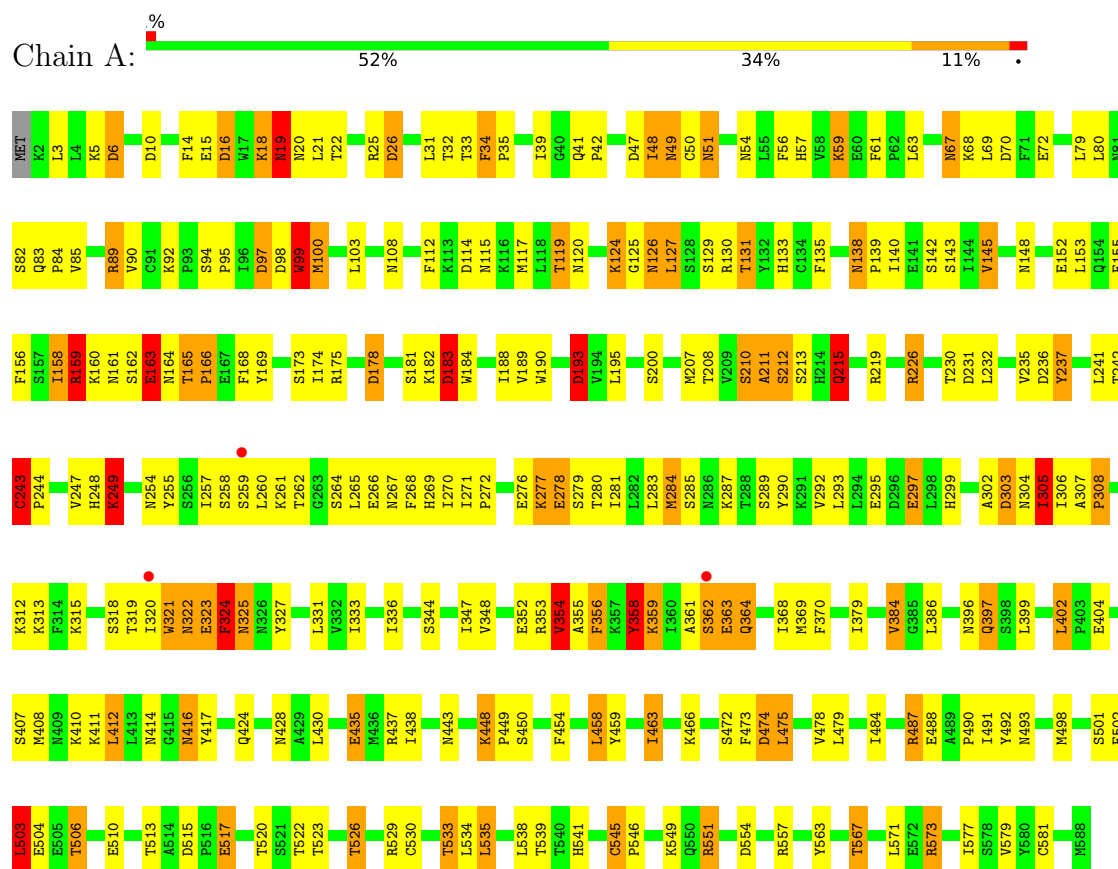
- Molecule 3 is water.

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

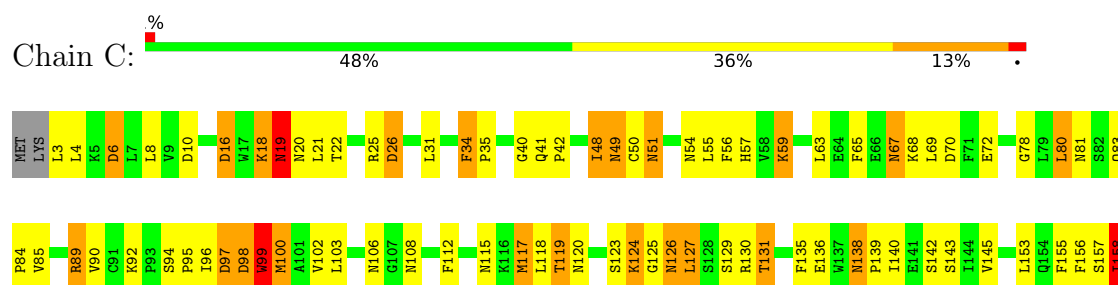
### 3 Residue-property plots

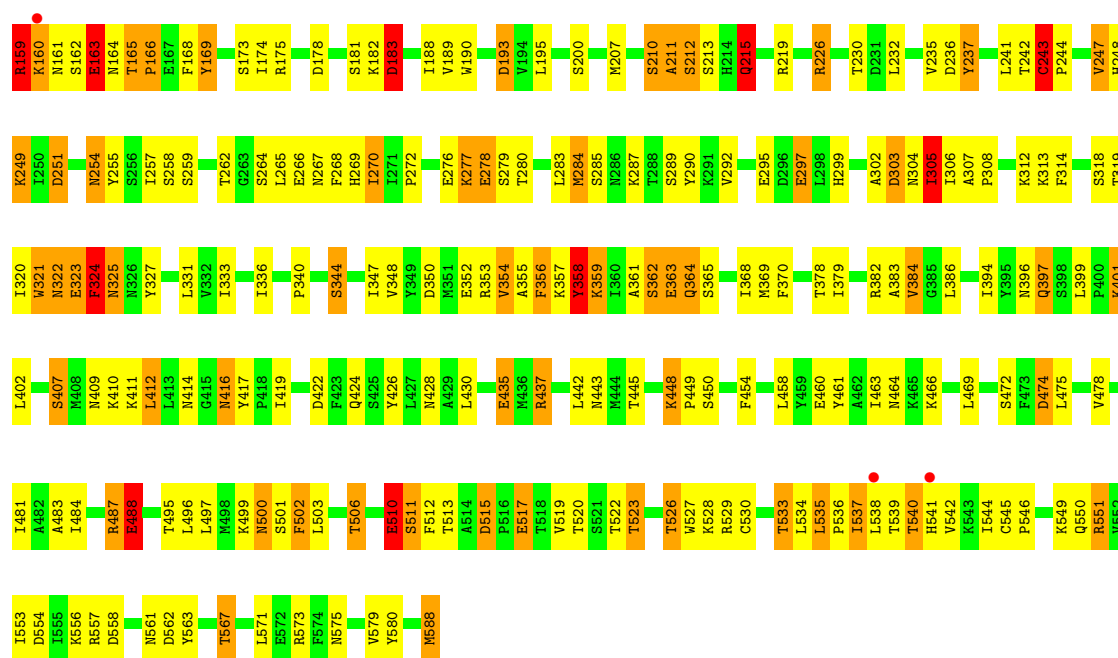
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: HYPOTHETICAL PROTEIN YPL007C



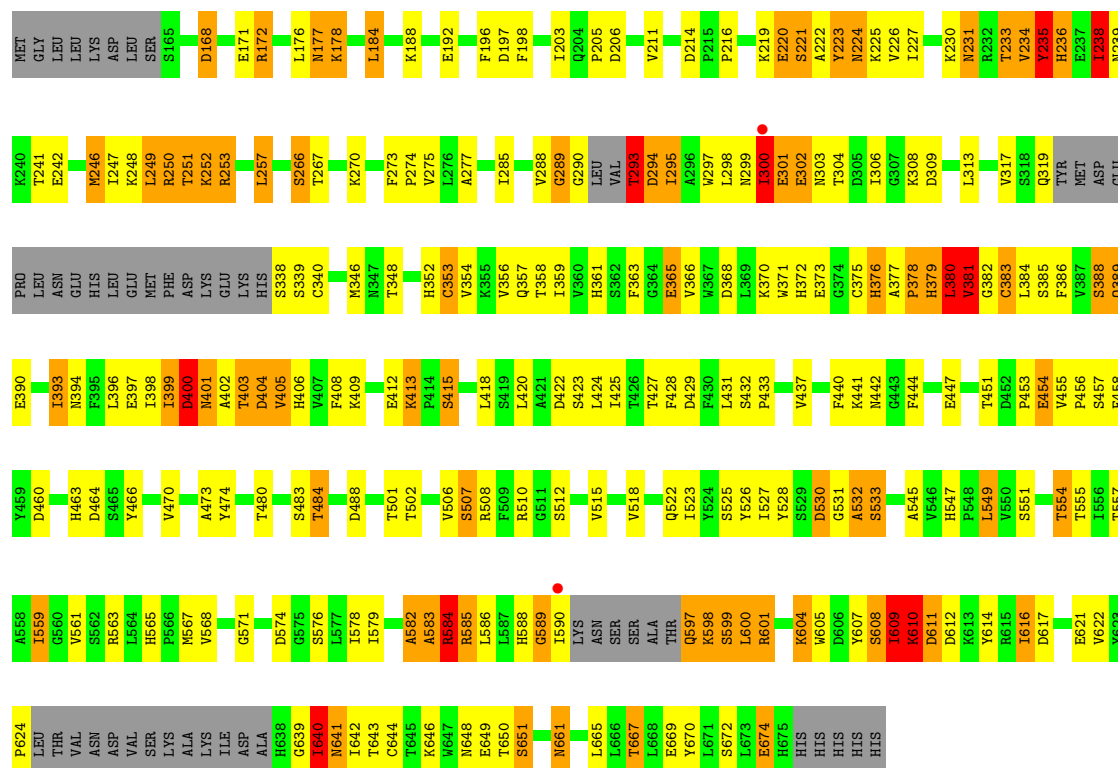
#### • Molecule 1: HYPOTHETICAL PROTEIN YPL007C





### • Molecule 2: YDR362CP

Chain B: 43% 32% 13% 10%



### • Molecule 2: YDR362CP

Chain D: 43% 33% 12% 10%

LYS	ALA	LYS	ILE	ASP	A637	H638	G639	D640	N641	I642	T643	C644	T645	K646	W647	N648	E649	T650	S651	M661	S662	L665	L666	T667	L668	S672	L673	E674	H675	HIS	HIS	HIS	HIS																																																														
R563	L564	H565	P566	M567	V568	L569	D574	I578	I579	T580	N581	A582	A583	R584	R585	L586	L587	H588	G589	ILE	LYS	ASN	SER	SER	ALA	THR	GLN	K598	S599	L600	R601	K604	W605	D606	Y607	S608	I609	K610	D611	D612	K613	Y614	R615	I616	E621	V622	Y623	P624	LEU	THR	VAL	ASN	ASP	VAL	SER																																								
F458	H463	D464	S469	V470	A473	Y474	D476	D479	T480	S483	T484	D488	T501	V506	S507	R508	F509	R510	G511	S512	V515	C520	P521	Q522	I523	Y524	S525	Y526	I527	D530	G531	A532	S533	A545	V546	H547	P548	L549	T554	T555	I556	T557	A558	I559	S385	F386	V387	S388	Q389	E390	I393	N394	F395	L396	E397	I398	I399	D400	N401	T403	D404	V405	H406	V407	K409	E412	K413	P414	G415	L418	S419	L420	S423	L424	I425	T426	T427	F428	D429	F430	L431	S432	P433	V437	F440	K441	E447	T451	D452	P453	E454	V455	G382	C383	L384
Y320	Met	ASP	GLU	PRO	LEU	ASN	GLU	HIS	LEU	Met	PHE	ASP	LYS	GLU	LYS	HIS	S338	S339	C340	K345	F408	N347	T348	H352	C353	V354	K355	V356	Q357	T358	I359	H361	V360	S362	F363	G364	E365	V366	V367	D368	L369	K370	N371	H372	E373	G374	C375	H376	A377	P378	H379	L380	V381	G382	C383	L384																																							
E242	N246	I247	K248	L249	R250	T251	K252	R253	L257	I258	D262	S266	T267	K270	I271	E272	F273	P274	V275	L276	A277	I285	V288	G289	GLY	LEU	VAL	T293	D294	I295	A296	W297	L298	N299	I300	E301	E302	N303	T304	D305	I306	G307	K308	D309	I310	Q311	V312	L313	V317	S318	Q319																																												
Met	GLY	LEU	LEU	LYS	ASP	LEU	SER	SER	ALA	R167	D168	E171	R172	L176	N177	K178	L184	K188	E192	F196	D197	F198	I203	Q204	P205	V211	P216	K219	E220	S221	A222	Y223	N224	K225	V226	I227	K230	N231	R232	T233	V234	Y235	H236	F237	I238	N239	K240	T241																																															

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.42Å 125.80Å 210.47Å 90.00° 94.49° 90.00°	Depositor
Resolution (Å)	208.51 – 3.20 208.51 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.7 (208.51-3.20) 98.7 (208.51-3.20)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.212 , 0.257 0.209 , 0.254	Depositor DCC
$R_{free}$ test set	2619 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	77.5	Xtriage
Anisotropy	0.356	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 128.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	16935	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 3.63% 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

### 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	0.83	5/4872 (0.1%)	0.94	15/6610 (0.2%)
1	C	0.80	5/4863 (0.1%)	0.95	16/6599 (0.2%)
2	B	0.76	4/3798 (0.1%)	0.98	21/5153 (0.4%)
2	D	0.77	4/3788 (0.1%)	0.95	15/5140 (0.3%)
All	All	0.79	18/17321 (0.1%)	0.95	67/23502 (0.3%)

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	2
1	C	0	3
2	B	0	8
2	D	0	7
All	All	0	20

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	413	LYS	CD-CE	18.55	1.97	1.51
2	B	413	LYS	CD-CE	14.62	1.87	1.51
1	A	249	LYS	CD-CE	13.65	1.85	1.51
1	A	297	GLU	CD-OE1	11.57	1.38	1.25
1	C	249	LYS	CD-CE	11.35	1.79	1.51

The worst 5 of 67 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	413	LYS	CD-CE-NZ	-8.44	92.29	111.70
2	B	413	LYS	CD-CE-NZ	-8.38	92.42	111.70
2	D	413	LYS	CG-CD-CE	-7.55	89.24	111.90

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	6	ASP	CB-CG-OD2	7.20	124.78	118.30
1	A	231	ASP	CB-CG-OD2	7.00	124.60	118.30

There are no chirality outliers.

5 of 20 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	243	CYS	Peptide
1	A	98	ASP	Peptide
2	B	277	ALA	Peptide
2	B	293	THR	Peptide
2	B	379	HIS	Peptide

## 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	4760	0	4722	220	0
1	C	4751	0	4709	244	0
2	B	3715	0	3674	258	0
2	D	3704	0	3660	238	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
All	All	16935	0	16765	942	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 28.

The worst 5 of 942 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:249:LYS:CE	1:C:249:LYS:CD	1.79	1.60
2:B:413:LYS:CE	2:B:413:LYS:CD	1.87	1.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:249:LYS:CE	1:A:249:LYS:CD	1.85	1.50
2:B:413:LYS:CE	2:B:413:LYS:NZ	1.74	1.47
2:D:413:LYS:NZ	2:D:413:LYS:CE	1.76	1.45

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	585/588 (100%)	475 (81%)	74 (13%)	36 (6%)	1	9
1	C	584/588 (99%)	475 (81%)	69 (12%)	40 (7%)	1	7
2	B	462/524 (88%)	362 (78%)	68 (15%)	32 (7%)	1	7
2	D	459/524 (88%)	356 (78%)	76 (17%)	27 (6%)	1	10
All	All	2090/2224 (94%)	1668 (80%)	287 (14%)	135 (6%)	1	8

5 of 135 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	80	LEU
1	A	124	LYS
1	A	159	ARG
1	A	164	ASN
1	A	166	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	543/544 (100%)	434 (80%)	109 (20%)	1	5
1	C	542/544 (100%)	434 (80%)	108 (20%)	1	5
2	B	415/465 (89%)	340 (82%)	75 (18%)	1	7
2	D	414/465 (89%)	336 (81%)	78 (19%)	1	6
All	All	1914/2018 (95%)	1544 (81%)	370 (19%)	1	6

5 of 370 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	270	ILE
1	C	540	THR
1	C	312	LYS
1	C	435	GLU
2	D	227	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 60 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	648	ASN
2	D	565	HIS
1	C	138	ASN
2	D	547	HIS
2	D	661	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 oligosaccharides 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	587/588 (99%)	-0.06	3 (0%) 87 78	61, 71, 71, 81	0
1	C	586/588 (99%)	0.04	3 (0%) 87 78	60, 71, 71, 87	0
2	B	472/524 (90%)	0.01	2 (0%) 89 81	70, 71, 71, 72	0
2	D	469/524 (89%)	-0.09	5 (1%) 77 63	71, 71, 71, 75	0
All	All	2114/2224 (95%)	-0.02	13 (0%) 85 76	60, 71, 71, 87	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	300	ILE	3.1
2	B	590	ILE	2.7
2	D	300	ILE	2.5
2	D	306	ILE	2.2
2	D	402	ALA	2.2

### 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.