



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 8, 2021 – 10:08 AM JST

PDB ID : 7CVQ  
Title : crystal structure of Arabidopsis CO CCT domain in complex with NF-YB2/YC3 and FT CORE1 DNA  
Authors : Lv, X.; Du, J.  
Deposited on : 2020-08-26  
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

<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
Xtriage (Phenix)	:	1.13
EDS	:	2.20
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)	:	2.20

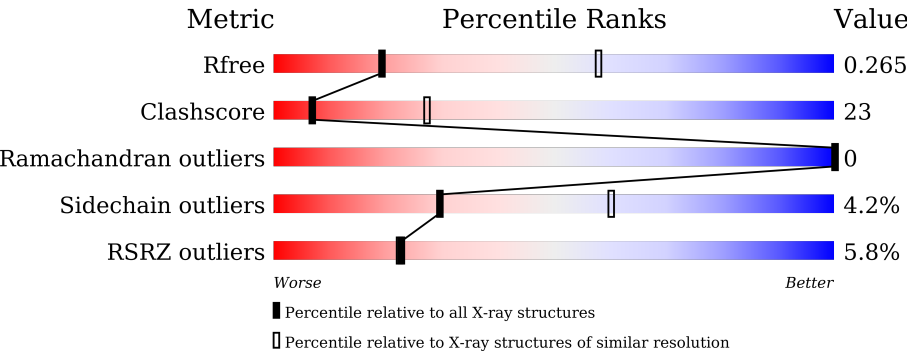
# 1 Overall quality at a glance

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))
$R_{free}$	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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	174	<div> <div>3%</div> <div> <div>45%</div> <div>27%</div> <div>28%</div> </div> </div>
1	F	174	<div> <div>%</div> <div> <div>44%</div> <div>27%</div> <div>28%</div> </div> </div>
1	K	174	<div> <div>8%</div> <div> <div>41%</div> <div>24%</div> <div>33%</div> </div> </div>
1	P	174	<div> <div>10%</div> <div> <div>37%</div> <div>29%</div> <div>30%</div> </div> </div>
2	B	97	<div> <div>2%</div> <div> <div>52%</div> <div>34%</div> <div>12%</div> </div> </div>
2	G	97	<div> <div>%</div> <div> <div>53%</div> <div>35%</div> <div>11%</div> </div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	L	97	
2	Q	97	
3	D	25	
3	I	25	
3	N	25	
3	T	25	
4	E	25	
4	J	25	
4	O	25	
4	U	25	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10910 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 Chimera of Nuclear transcription factor Y subunit C-3 and Zinc finger protein CONSTANS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	S	0	0	0
			1050	662	204	179	5			
1	F	125	Total	C	N	O	S	0	0	0
			1050	662	204	179	5			
1	K	117	Total	C	N	O	S	0	0	0
			978	615	188	170	5			
1	P	121	Total	C	N	O	S	0	0	0
			1011	637	195	174	5			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	278	GLY	-	linker	UNP Q9ZVL3
A	279	SER	-	linker	UNP Q9ZVL3
A	280	GLY	-	linker	UNP Q9ZVL3
A	281	SER	-	linker	UNP Q9ZVL3
A	282	GLY	-	linker	UNP Q9ZVL3
A	283	SER	-	linker	UNP Q9ZVL3
A	284	GLY	-	linker	UNP Q9ZVL3
A	285	SER	-	linker	UNP Q9ZVL3
A	286	GLY	-	linker	UNP Q9ZVL3
A	287	SER	-	linker	UNP Q9ZVL3
A	288	GLY	-	linker	UNP Q9ZVL3
A	289	SER	-	linker	UNP Q9ZVL3
F	278	GLY	-	linker	UNP Q9ZVL3
F	279	SER	-	linker	UNP Q9ZVL3
F	280	GLY	-	linker	UNP Q9ZVL3
F	281	SER	-	linker	UNP Q9ZVL3
F	282	GLY	-	linker	UNP Q9ZVL3
F	283	SER	-	linker	UNP Q9ZVL3
F	284	GLY	-	linker	UNP Q9ZVL3
F	285	SER	-	linker	UNP Q9ZVL3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	286	GLY	-	linker	UNP Q9ZVL3
F	287	SER	-	linker	UNP Q9ZVL3
F	288	GLY	-	linker	UNP Q9ZVL3
F	289	SER	-	linker	UNP Q9ZVL3
K	278	GLY	-	linker	UNP Q9ZVL3
K	279	SER	-	linker	UNP Q9ZVL3
K	280	GLY	-	linker	UNP Q9ZVL3
K	281	SER	-	linker	UNP Q9ZVL3
K	282	GLY	-	linker	UNP Q9ZVL3
K	283	SER	-	linker	UNP Q9ZVL3
K	284	GLY	-	linker	UNP Q9ZVL3
K	285	SER	-	linker	UNP Q9ZVL3
K	286	GLY	-	linker	UNP Q9ZVL3
K	287	SER	-	linker	UNP Q9ZVL3
K	288	GLY	-	linker	UNP Q9ZVL3
K	289	SER	-	linker	UNP Q9ZVL3
P	278	GLY	-	linker	UNP Q9ZVL3
P	279	SER	-	linker	UNP Q9ZVL3
P	280	GLY	-	linker	UNP Q9ZVL3
P	281	SER	-	linker	UNP Q9ZVL3
P	282	GLY	-	linker	UNP Q9ZVL3
P	283	SER	-	linker	UNP Q9ZVL3
P	284	GLY	-	linker	UNP Q9ZVL3
P	285	SER	-	linker	UNP Q9ZVL3
P	286	GLY	-	linker	UNP Q9ZVL3
P	287	SER	-	linker	UNP Q9ZVL3
P	288	GLY	-	linker	UNP Q9ZVL3
P	289	SER	-	linker	UNP Q9ZVL3

- Molecule 2 is a protein called Nuclear transcription factor Y subunit B-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	85	Total	C	N	O	S	0	0	0
			678	430	114	129	5			
2	G	86	Total	C	N	O	S	0	0	0
			689	439	115	130	5			
2	L	86	Total	C	N	O	S	0	0	0
			689	439	115	130	5			
2	Q	86	Total	C	N	O	S	0	0	0
			689	439	115	130	5			

- Molecule 3 is a DNA chain called FT CORE1 DNA forward strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	25	Total 516	C 247	N 98	O 147	P 24	0	0	0
3	I	25	Total 516	C 247	N 98	O 147	P 24	0	0	0
3	N	25	Total 516	C 247	N 98	O 147	P 24	0	0	0
3	T	25	Total 516	C 247	N 98	O 147	P 24	0	0	0

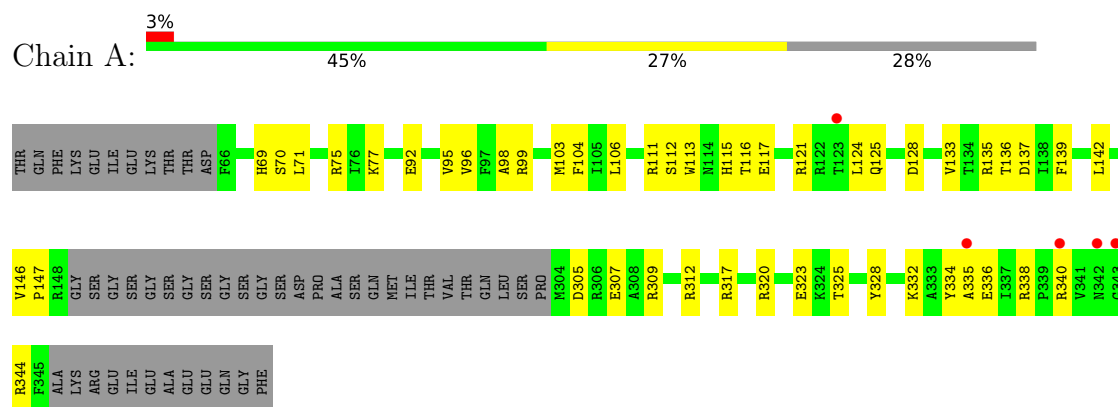
- Molecule 4 is a DNA chain called FT CORE1 DNA reverse strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	25	Total 503	C 243	N 87	O 149	P 24	0	0	0
4	J	25	Total 503	C 243	N 87	O 149	P 24	0	0	0
4	O	25	Total 503	C 243	N 87	O 149	P 24	0	0	0
4	U	25	Total 503	C 243	N 87	O 149	P 24	0	0	0

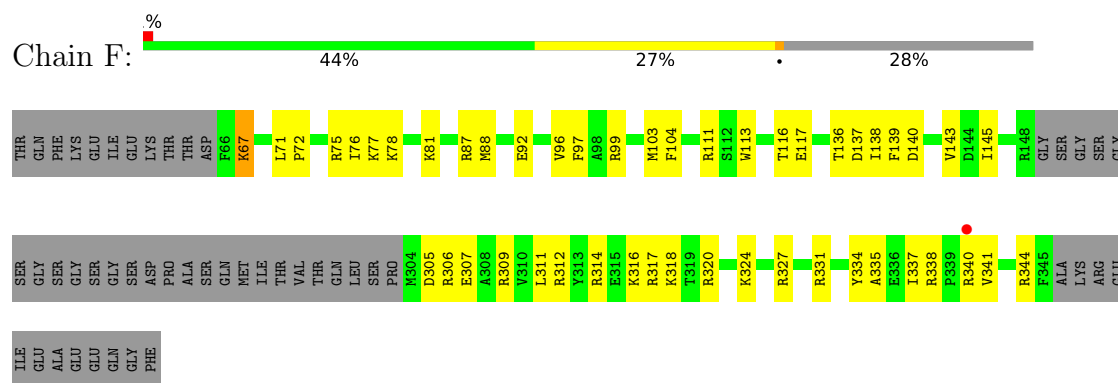
### 3 Residue-property plots [i](#)

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.

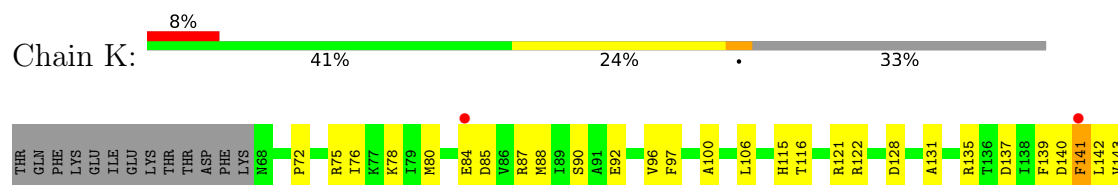
- Molecule 1: Chimera of Nuclear transcription factor Y subunit C-3 and Zinc finger protein CONSTANS

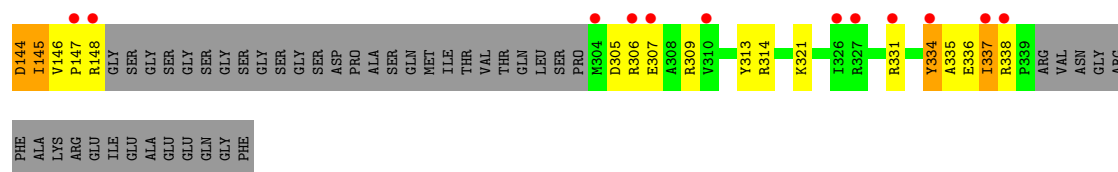


- Molecule 1: Chimera of Nuclear transcription factor Y subunit C-3 and Zinc finger protein CONSTANS

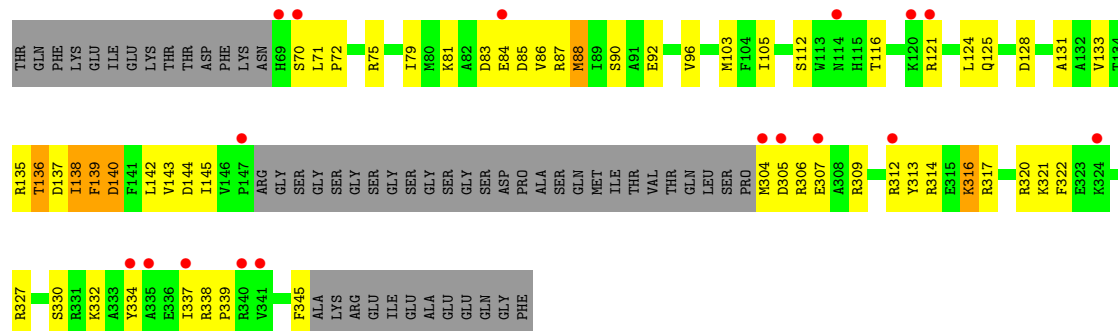


- Molecule 1: Chimera of Nuclear transcription factor Y subunit C-3 and Zinc finger protein CONSTANS

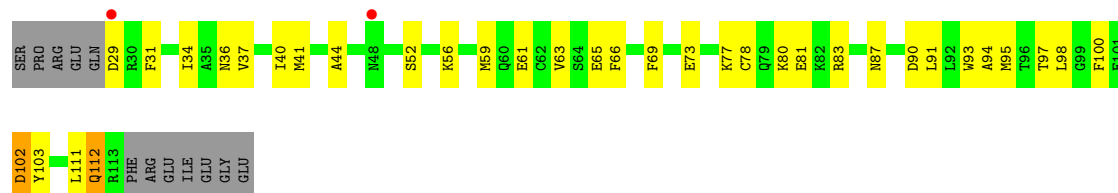




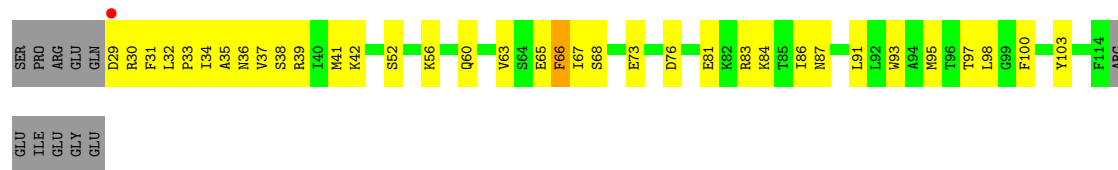
- Molecule 1: Chimera of Nuclear transcription factor Y subunit C-3 and Zinc finger protein CONSTANS



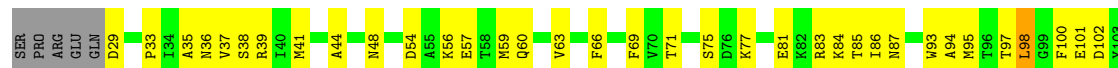
- Molecule 2: Nuclear transcription factor Y subunit B-2



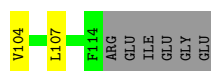
- Molecule 2: Nuclear transcription factor Y subunit B-2



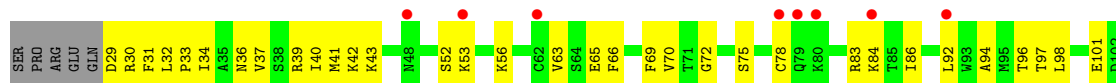
- Molecule 2: Nuclear transcription factor Y subunit B-2







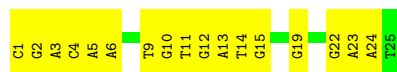
- Molecule 2: Nuclear transcription factor Y subunit B-2



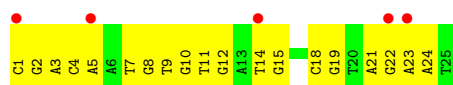
- Molecule 3: FT CORE1 DNA forward strand



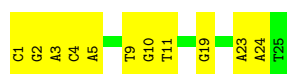
- Molecule 3: FT CORE1 DNA forward strand



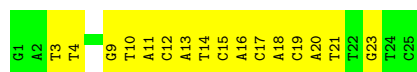
- Molecule 3: FT CORE1 DNA forward strand



- Molecule 3: FT CORE1 DNA forward strand



- Molecule 4: FT CORE1 DNA reverse strand

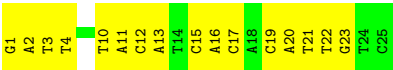


- Molecule 4: FT CORE1 DNA reverse strand

Chain J: 

36%

64%



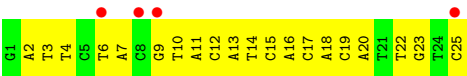
● Molecule 4: FT CORE1 DNA reverse strand

Chain O: 

16%

20%

80%



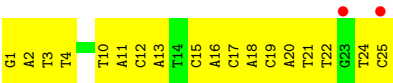
● Molecule 4: FT CORE1 DNA reverse strand

Chain U: 

8%

28%

72%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.24Å 65.27Å 138.28Å 90.00° 91.17° 90.00°	Depositor
Resolution (Å)	35.18 – 3.30 37.65 – 3.29	Depositor EDS
% Data completeness (in resolution range)	98.8 (35.18-3.30) 98.6 (37.65-3.29)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.48 (at 3.32Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.241 , 0.264 0.244 , 0.265	Depositor DCC
$R_{free}$ test set	1423 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	95.4	Xtriage
Anisotropy	0.827	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 73.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Estimated twinning fraction	0.085 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	10910	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	127.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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	0.28	0/1066	0.48	0/1425
1	F	0.35	0/1066	0.52	0/1425
1	K	0.36	0/992	0.58	0/1328
1	P	0.34	0/1026	0.56	0/1373
2	B	0.33	0/688	0.46	0/923
2	G	0.33	0/700	0.54	0/939
2	L	0.29	0/700	0.49	0/939
2	Q	0.27	0/700	0.47	0/939
3	D	0.58	0/580	0.95	0/895
3	I	0.61	0/580	0.97	0/895
3	N	0.56	0/580	0.93	0/895
3	T	0.55	0/580	0.95	0/895
4	E	0.61	0/562	1.01	0/864
4	J	0.58	0/562	0.97	0/864
4	O	0.61	0/562	1.00	0/864
4	U	0.62	0/562	1.01	0/864
All	All	0.45	0/11506	0.75	0/16327

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	1050	0	1093	50	0
1	F	1050	0	1093	55	0
1	K	978	0	1018	75	0
1	P	1011	0	1052	78	0
2	B	678	0	692	36	0
2	G	689	0	701	39	0
2	L	689	0	701	43	0
2	Q	689	0	701	45	0
3	D	516	0	284	15	0
3	I	516	0	284	18	0
3	N	516	0	284	18	0
3	T	516	0	284	9	0
4	E	503	0	285	20	0
4	J	503	0	285	24	0
4	O	503	0	285	27	0
4	U	503	0	285	22	0
All	All	10910	0	9327	449	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 23.

The worst 5 of 449 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:K:335:ALA:HA	1:K:338:ARG:CD	1.72	1.20
1:K:140:ASP:HB2	1:K:306:ARG:HH12	1.07	1.17
1:P:139:PHE:HB3	1:P:142:LEU:HD12	1.27	1.13
1:P:84:GLU:OE2	1:P:321:LYS:HD2	1.50	1.11
2:L:98:LEU:HD22	2:L:100:PHE:CZ	1.89	1.07

There are no symmetry-related clashes.

## 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	121/174 (70%)	116 (96%)	5 (4%)	0	100	100
1	F	121/174 (70%)	117 (97%)	4 (3%)	0	100	100
1	K	113/174 (65%)	107 (95%)	6 (5%)	0	100	100
1	P	117/174 (67%)	113 (97%)	4 (3%)	0	100	100
2	B	83/97 (86%)	81 (98%)	2 (2%)	0	100	100
2	G	84/97 (87%)	82 (98%)	2 (2%)	0	100	100
2	L	84/97 (87%)	84 (100%)	0	0	100	100
2	Q	84/97 (87%)	83 (99%)	1 (1%)	0	100	100
All	All	807/1084 (74%)	783 (97%)	24 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	111/150 (74%)	110 (99%)	1 (1%)	78	87
1	F	111/150 (74%)	107 (96%)	4 (4%)	35	63
1	K	104/150 (69%)	95 (91%)	9 (9%)	10	34
1	P	107/150 (71%)	100 (94%)	7 (6%)	17	46
2	B	75/86 (87%)	71 (95%)	4 (5%)	22	53
2	G	76/86 (88%)	75 (99%)	1 (1%)	69	82
2	L	76/86 (88%)	72 (95%)	4 (5%)	22	53
2	Q	76/86 (88%)	75 (99%)	1 (1%)	69	82
All	All	736/944 (78%)	705 (96%)	31 (4%)	30	60

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	148	ARG
2	L	66	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	K	337	ILE
2	L	102	ASP
1	P	316	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	112	GLN
1	K	115	HIS

### 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 monosaccharides 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 [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	125/174 (71%)	0.39	5 (4%) 38 36	70, 103, 134, 162	0
1	F	125/174 (71%)	0.19	1 (0%) 86 86	57, 87, 120, 143	0
1	K	117/174 (67%)	0.72	14 (11%) 4 3	103, 141, 212, 234	0
1	P	121/174 (69%)	0.75	17 (14%) 2 2	102, 155, 205, 226	0
2	B	85/97 (87%)	0.19	2 (2%) 59 56	66, 86, 111, 127	0
2	G	86/97 (88%)	0.03	1 (1%) 79 78	55, 77, 98, 110	0
2	L	86/97 (88%)	0.19	0 100 100	105, 134, 157, 178	0
2	Q	86/97 (88%)	0.62	9 (10%) 6 6	105, 165, 188, 194	0
3	D	25/25 (100%)	0.07	0 100 100	89, 116, 136, 148	0
3	I	25/25 (100%)	0.11	0 100 100	86, 108, 137, 146	0
3	N	25/25 (100%)	0.78	5 (20%) 1 1	154, 176, 210, 230	0
3	T	25/25 (100%)	0.20	0 100 100	107, 136, 171, 193	0
4	E	25/25 (100%)	-0.12	0 100 100	91, 108, 152, 156	0
4	J	25/25 (100%)	0.07	0 100 100	87, 107, 130, 135	0
4	O	25/25 (100%)	0.93	4 (16%) 1 2	161, 183, 209, 219	0
4	U	25/25 (100%)	0.38	2 (8%) 12 11	115, 137, 165, 196	0
All	All	1031/1284 (80%)	0.38	60 (5%) 23 22	55, 121, 191, 234	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	69	HIS	4.7
1	K	334	TYR	4.2
1	P	312	ARG	4.0
2	B	29	ASP	3.9
1	P	340	ARG	3.8



## 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.