



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 13, 2018 – 01:14 pm GMT

PDB ID : 4UN0  
Title : Crystal structure of the human CDK12-cyclinK complex  
Authors : Dixon Clarke, S.E.; Elkins, J.M.; Pike, A.C.W.; Chaikuad, A.; Goubin, S.; Krojer, T.; Sorrell, F.J.; Nowak, R.; Williams, E.; Kopec, J.; Mahajan, R.P.; Burgess-Brown, N.; Carpenter, E.P.; Knapp, S.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Bullock, A.  
Deposited on : 2014-05-22  
Resolution : 3.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk31020

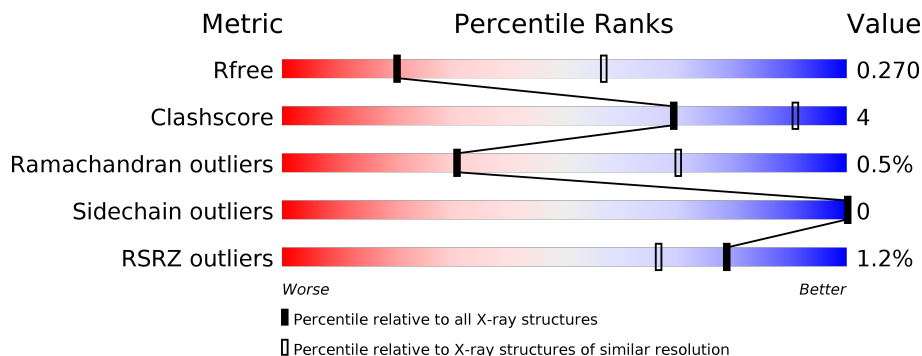
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.15 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))
$R_{free}$	111664	1389 (3.20-3.12)
Clashscore	122126	1522 (3.20-3.12)
Ramachandran outliers	120053	1493 (3.20-3.12)
Sidechain outliers	120020	1492 (3.20-3.12)
RSRZ outliers	108989	1344 (3.20-3.12)

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	259	 85% 7% 8%
1	B	259	 83% 8% 8%
2	C	326	 82% 12% 5%
2	D	326	 4% 76% 9% 15%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8190 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 CYCLIN-K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	239	Total 1929	C 1258	N 312	O 348	S 11	0	0	0
1	B	237	Total 1868	C 1227	N 300	O 330	S 11	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	SER	-	expression tag	UNP O75909
A	10	MET	-	expression tag	UNP O75909
B	9	SER	-	expression tag	UNP O75909
B	10	MET	-	expression tag	UNP O75909

- Molecule 2 is a protein called CYCLIN-DEPENDENT KINASE 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	P				S
2	C	309	Total 2325	C 1511	N 378	O 420	P 1	S 15	0	0	0
2	D	277	Total 2068	C 1342	N 338	O 372	P 1	S 15	0	0	0


There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	713	SER	-	expression tag	UNP Q9NYV4
C	714	MET	-	expression tag	UNP Q9NYV4
D	713	SER	-	expression tag	UNP Q9NYV4
D	714	MET	-	expression tag	UNP Q9NYV4

### 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 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.

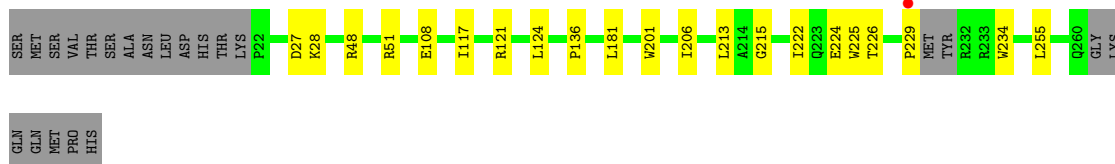
- Molecule 1: CYCLIN-K

Chain A: 




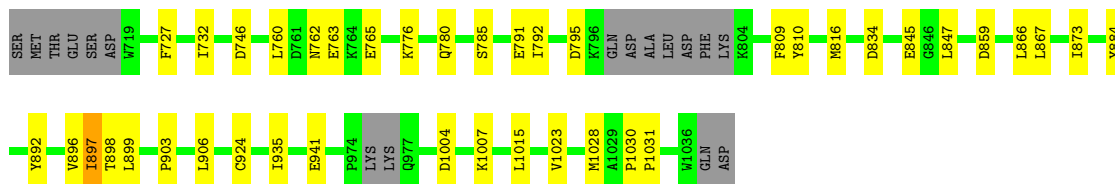
- Molecule 1: CYCLIN-K

Chain B: 




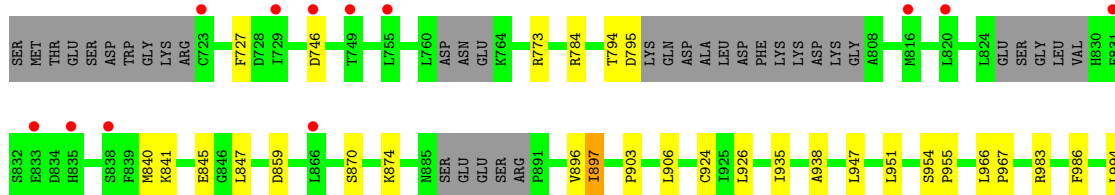
- Molecule 2: CYCLIN-DEPENDENT KINASE 12

Chain C: 



- Molecule 2: CYCLIN-DEPENDENT KINASE 12

Chain D: 



E1024

LEU  
SER  
LYS  
MET  
ALA  
PRO  
PRO  
ASP  
LEU  
PRO  
HIS  
TRP  
GLN  
ASP

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.58Å 148.78Å 92.15Å 90.00° 94.22° 90.00°	Depositor
Resolution (Å)	57.82 – 3.15 57.82 – 3.15	Depositor EDS
% Data completeness (in resolution range)	99.2 (57.82-3.15) 99.2 (57.82-3.15)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.75 (at 3.13Å)	Xtrriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.215 , 0.271 0.218 , 0.270	Depositor DCC
$R_{free}$ test set	1189 reflections (5.19%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.9	Xtrriage
Anisotropy	0.048	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 47.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	8190	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: TPO

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	0.21	0/1984	0.36	0/2703
1	B	0.21	0/1921	0.37	0/2620
2	C	0.22	0/2367	0.41	0/3222
2	D	0.22	0/2099	0.39	0/2855
All	All	0.22	0/8371	0.38	0/11400

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	1929	0	1821	13	0
1	B	1868	0	1754	13	0
2	C	2325	0	2180	23	0
2	D	2068	0	1918	17	0
All	All	8190	0	7673	63	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 4.

All (63) 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:A:156:GLN:OE1	2:C:780:GLN:NE2	2.11	0.84
2:C:896:VAL:HG12	2:C:897:ILE:HG12	1.60	0.80
2:C:763:GLU:HB3	2:C:765:GLU:HG2	1.63	0.78
1:A:27:ASP:OD1	1:A:28:LYS:N	2.16	0.78
1:B:27:ASP:OD1	1:B:28:LYS:N	2.22	0.72
2:D:896:VAL:HG12	2:D:897:ILE:HG12	1.70	0.72
2:D:983:ARG:HG2	2:D:994:LEU:HD21	1.81	0.63
1:A:221:GLU:OE1	1:A:223:GLN:NE2	2.32	0.62
2:C:732:ILE:HD12	2:C:732:ILE:H	1.67	0.59
2:C:834:ASP:HB3	2:C:1028:MET:HE1	1.84	0.59
2:D:840:MET:HG3	2:D:926:LEU:HD13	1.85	0.59
1:B:117:ILE:HD13	1:B:136:PRO:HB2	1.85	0.58
1:A:117:ILE:HD13	1:A:136:PRO:HB2	1.86	0.55
2:D:841:LYS:O	2:D:845:GLU:HG3	2.06	0.55
2:D:784:ARG:O	2:D:874:LYS:NZ	2.35	0.55
1:B:215:GLY:HA3	1:B:222:ILE:HD11	1.90	0.54
2:D:903:PRO:HD2	2:D:906:LEU:HD12	1.90	0.54
1:A:134:ASP:OD1	1:A:134:ASP:N	2.41	0.53
1:B:226:THR:HG21	1:B:234:TRP:HB2	1.90	0.53
2:C:903:PRO:HD2	2:C:906:LEU:HD12	1.90	0.53
1:A:48:ARG:HA	1:A:51:ARG:HE	1.74	0.53
2:D:951:LEU:HD11	2:D:986:PHE:HE1	1.74	0.53
2:D:727:PHE:HA	2:D:746:ASP:HA	1.90	0.52
2:D:794:THR:OG1	2:D:795:ASP:N	2.43	0.52
1:B:48:ARG:HA	1:B:51:ARG:HE	1.75	0.51
1:A:106:VAL:HG13	2:C:776:LYS:HD3	1.92	0.51
1:B:121:ARG:NH1	1:B:124:LEU:O	2.45	0.50
1:B:181:LEU:HD22	1:B:225:TRP:HZ3	1.77	0.50
2:C:785:SER:OG	2:C:785:SER:O	2.31	0.49
1:A:231:TYR:HE2	1:A:233:ARG:HB2	1.77	0.48
2:C:762:ASN:HB2	2:C:763:GLU:HG3	1.95	0.48
1:B:224:GLU:HA	1:B:229:PRO:HB3	1.95	0.47
1:A:231:TYR:CE2	1:A:233:ARG:HB2	2.50	0.47
2:D:938:ALA:HB2	2:D:947:LEU:HD12	1.97	0.47
2:C:727:PHE:CE2	2:C:746:ASP:HB3	2.50	0.46
2:C:795:ASP:OD1	2:C:810:TYR:HE1	1.98	0.46
2:C:845:GLU:HA	2:C:1015:LEU:HD11	1.96	0.46
2:C:760:LEU:HD11	2:C:809:PHE:HE1	1.81	0.46
2:C:867:LEU:HD13	2:C:873:ILE:HG13	1.97	0.46
1:A:222:ILE:H	1:A:222:ILE:HD12	1.82	0.45
2:C:1030:PRO:HA	2:C:1031:PRO:HD3	1.85	0.45

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:897:ILE:O	2:C:898:THR:OG1	2.30	0.45
1:B:222:ILE:H	1:B:222:ILE:HD12	1.82	0.44
2:C:924:CYS:HA	2:C:935:ILE:HD11	1.99	0.44
2:D:924:CYS:HA	2:D:935:ILE:HD11	2.00	0.44
2:C:1004:ASP:HB3	2:C:1007:LYS:HB2	2.00	0.44
2:C:899:LEU:HD11	2:C:941:GLU:HA	1.99	0.44
2:D:966:LEU:HA	2:D:967:PRO:HD3	1.87	0.43
1:A:76:ILE:HD12	1:A:159:HIS:CE1	2.55	0.42
1:B:201:TRP:HB2	1:B:206:ILE:HD11	2.02	0.42
2:C:816:MET:HB2	2:C:866:LEU:HD13	2.02	0.42
2:C:884:TYR:HA	2:C:892:TYR:OH	2.20	0.42
2:D:847:LEU:HA	2:D:847:LEU:HD23	1.87	0.41
2:C:847:LEU:HA	2:C:847:LEU:HD23	1.83	0.41
2:D:870:SER:O	2:D:870:SER:OG	2.31	0.41
1:A:224:GLU:HA	1:A:229:PRO:HB3	2.03	0.41
1:B:213:LEU:HD11	1:B:255:LEU:HD21	2.03	0.41
2:D:954:SER:HA	2:D:955:PRO:HD3	1.89	0.41
1:A:48:ARG:HB2	1:A:51:ARG:HH21	1.86	0.40
2:C:791:GLU:HG2	2:C:792:ILE:H	1.85	0.40
2:D:983:ARG:HG2	2:D:994:LEU:CD2	2.48	0.40
1:B:48:ARG:HB2	1:B:51:ARG:HH21	1.86	0.40
1:B:108:GLU:OE2	2:D:773:ARG:HD3	2.21	0.40

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	237/259 (92%)	235 (99%)	2 (1%)	0	100	100
1	B	233/259 (90%)	231 (99%)	2 (1%)	0	100	100
2	C	302/326 (93%)	272 (90%)	27 (9%)	3 (1%)	17	56

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	266/326 (82%)	240 (90%)	24 (9%)	2 (1%)	21	61
All	All	1038/1170 (89%)	978 (94%)	55 (5%)	5 (0%)	31	70

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	897	ILE
2	D	897	ILE
2	C	1023	VAL
2	C	859	ASP
2	D	859	ASP

### 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	198/233 (85%)	198 (100%)	0	100	100
1	B	186/233 (80%)	186 (100%)	0	100	100
2	C	225/293 (77%)	225 (100%)	0	100	100
2	D	196/293 (67%)	196 (100%)	0	100	100
All	All	805/1052 (76%)	805 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	156	GLN
1	B	33	HIS
2	C	780	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	TPO	C	893	2	9,10,11	1.28	1 (11%)	11,14,16	1.78	1 (9%)
2	TPO	D	893	2	9,10,11	1.27	1 (11%)	11,14,16	1.72	1 (9%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	TPO	C	893	2	-	0/8/11/13	0/0/0/0
2	TPO	D	893	2	-	0/8/11/13	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	893	TPO	CA-C	2.42	1.53	1.50
2	C	893	TPO	CA-C	2.44	1.53	1.50

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	893	TPO	P-OG1-CB	-5.33	107.11	123.21
2	D	893	TPO	P-OG1-CB	-5.10	107.79	123.21

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [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	239/259 (92%)	-0.24	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	24, 39, 61, 77	0
1	B	237/259 (91%)	-0.15	1 (0%) <span style="border: 1px solid blue; padding: 2px;">92</span> <span style="border: 1px solid blue; padding: 2px;">89</span>	34, 53, 77, 93	0
2	C	308/326 (94%)	-0.05	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	36, 55, 82, 95	0
2	D	276/326 (84%)	0.31	12 (4%) <span style="border: 1px solid red; padding: 2px;">35</span> <span style="border: 1px solid red; padding: 2px;">20</span>	46, 67, 94, 109	0
All	All	1060/1170 (90%)	-0.02	13 (1%) <span style="border: 1px solid blue; padding: 2px;">79</span> <span style="border: 1px solid blue; padding: 2px;">67</span>	24, 55, 85, 109	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	749	THR	3.8
2	D	723	CYS	2.9
2	D	755	LEU	2.6
2	D	833	GLU	2.5
2	D	838	SER	2.5
1	B	229	PRO	2.4
2	D	866	LEU	2.4
2	D	820	LEU	2.3
2	D	816	MET	2.1
2	D	831	PHE	2.1
2	D	746	ASP	2.1
2	D	729	ILE	2.0
2	D	835	HIS	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	TPO	D	893	11/12	0.94	0.15	48,49,66,68	0
2	TPO	C	893	11/12	0.95	0.14	47,54,62,67	0

### 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.