



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:16 PM GMT

PDB ID : 4C0H  
Title : Extended interface between Pcf11p and Clp1p and structural basis for ATP loss in Gly135Arg point mutant  
Authors : Dupin, A.F.; Fribourg, S.  
Deposited on : 2013-08-02  
Resolution : 2.70 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  
<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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

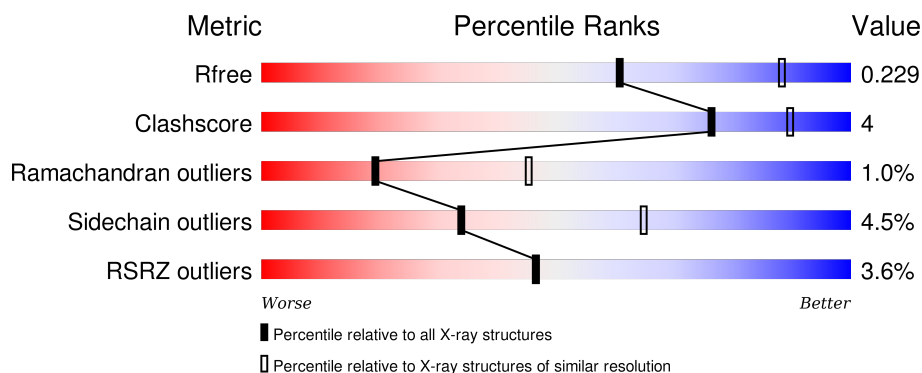
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	445	<div> <div>4%</div> <div>85%</div> <div>9%</div> <div>5%</div> </div>
1	B	445	<div> <div>3%</div> <div>85%</div> <div>9%</div> <div>• •</div> </div>
2	C	110	<div> <div>3%</div> <div>20%</div> <div>• •</div> <div>75%</div> </div>
2	D	110	<div> <div>2%</div> <div>23%</div> <div>6%</div> <div>•</div> <div>70%</div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 7336 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 MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	422	Total	C	N	O	S	0	1	0
			3391	2180	569	630	12			
1	B	427	Total	C	N	O	S	0	1	0
			3421	2198	574	637	12			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	135	ARG	GLY	ENGINEERED MUTATION	UNP Q08685
B	135	ARG	GLY	ENGINEERED MUTATION	UNP Q08685

- Molecule 2 is a protein called PCF11P.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	28	Total	C	N	O	0	0	0
			231	141	40	50			
2	D	33	Total	C	N	O	0	0	0
			268	165	47	56			

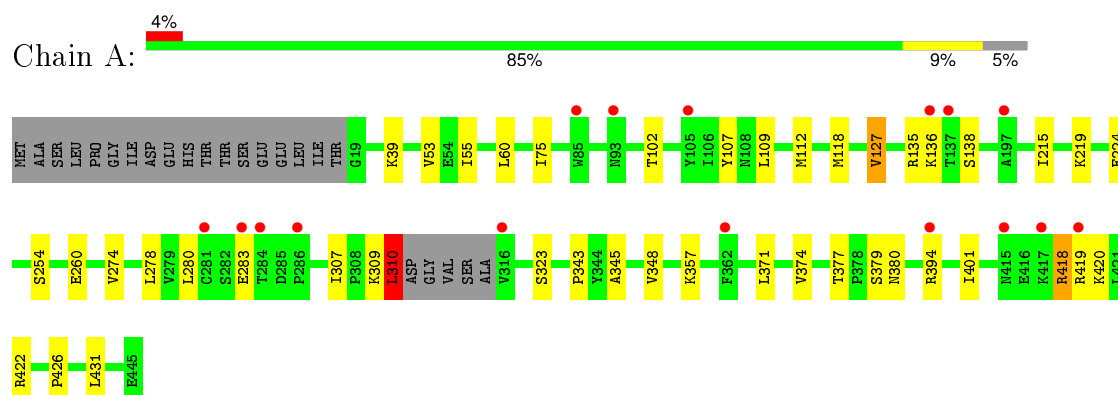
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	8	Total	O	0	0
			8	8		
3	B	14	Total	O	0	0
			14	14		
3	C	1	Total	O	0	0
			1	1		
3	D	2	Total	O	0	0
			2	2		

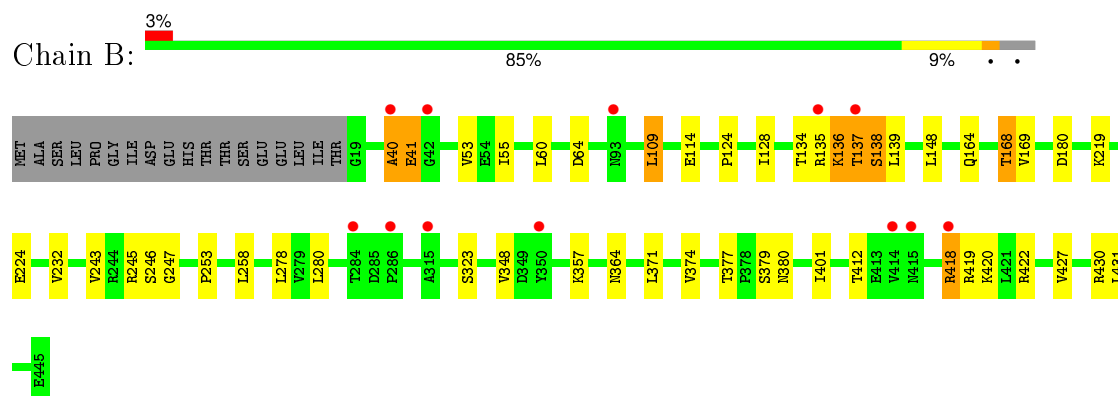
### 3 Residue-property plots [i](#)

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

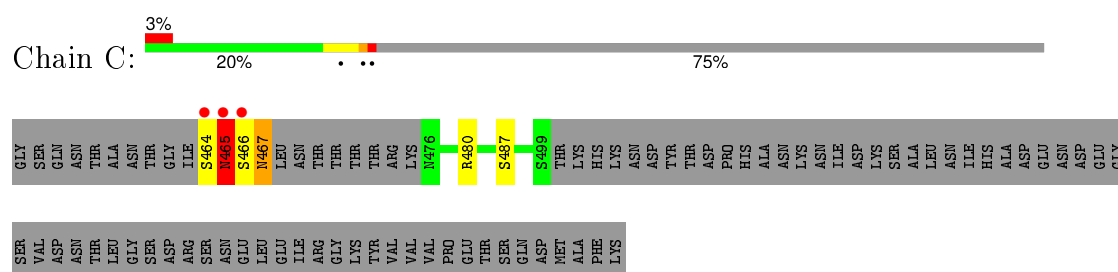
#### • Molecule 1: MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1



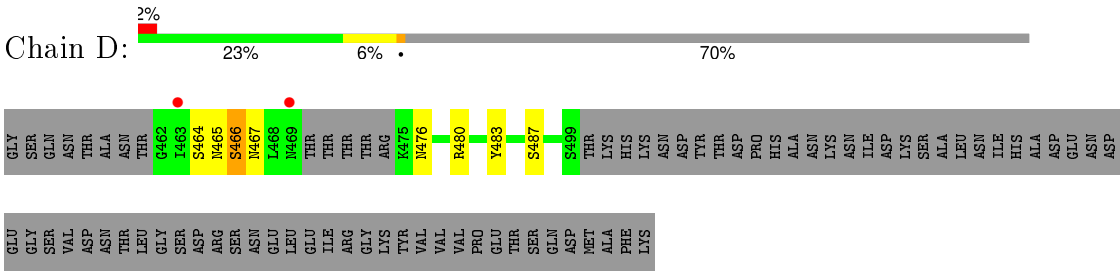
#### • Molecule 1: MRNA CLEAVAGE AND POLYADENYLATION FACTOR CLP1



#### • Molecule 2: PCF11P



#### • Molecule 2: PCF11P



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	89.39 Å 95.61 Å 182.06 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.80 – 2.70 46.24 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.80-2.70) 99.7 (46.24-2.70)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.191 , 0.223 0.201 , 0.229	Depositor DCC
$R_{free}$ test set	2188 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	55.4	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 52.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>1</sup>	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	1 of 43472 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7336	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.41% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.48	0/3467	0.74	1/4713 (0.0%)
1	B	0.51	0/3498	0.76	0/4758
2	C	0.50	0/235	0.70	0/317
2	D	0.62	0/272	0.75	0/366
All	All	0.50	0/7472	0.75	1/10154 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	310	LEU	N-CA-CB	5.00	120.41	110.40

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	3391	0	3430	21	0
1	B	3421	0	3459	26	0
2	C	231	0	200	10	0
2	D	268	0	244	5	0
3	A	8	0	0	0	0
3	B	14	0	0	0	0
3	C	1	0	0	0	0
3	D	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7336	0	7333	55	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 (55) 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:343:PRO:O	2:C:467:ASN:HA	1.70	0.91
1:B:168:THR:HG23	2:D:483:TYR:OH	1.75	0.85
1:A:345:ALA:HB3	2:C:466:SER:HA	1.62	0.80
1:A:127:VAL:HG22	1:A:274:VAL:HG11	1.66	0.78
2:C:467:ASN:ND2	2:C:467:ASN:H	1.87	0.72
1:B:348:VAL:HG11	1:B:431:LEU:HD22	1.70	0.72
2:D:466:SER:OG	2:D:467:ASN:N	2.22	0.71
1:B:412:THR:HG21	1:B:422:ARG:HE	1.56	0.70
1:B:135:ARG:HH22	1:B:253:PRO:C	1.99	0.65
2:C:467:ASN:HD22	2:C:467:ASN:H	1.47	0.62
1:A:53:VAL:HG13	1:A:75:ILE:HG23	1.82	0.62
1:B:124:PRO:O	1:B:247:GLY:HA3	2.00	0.60
1:A:348:VAL:HG11	1:A:431:LEU:HD22	1.84	0.59
1:B:137:THR:HG23	1:B:138:SER:H	1.70	0.57
2:C:465:ASN:ND2	2:C:466:SER:OG	2.37	0.56
1:A:371:LEU:HG	1:A:401:ILE:HD12	1.91	0.53
1:B:412:THR:HG21	1:B:422:ARG:NE	2.22	0.53
1:B:134:THR:O	1:B:138:SER:HB3	2.09	0.53
1:B:243:VAL:O	1:B:246:SER:O	2.27	0.52
1:A:357:LYS:HB2	1:A:374:VAL:HG22	1.92	0.52
1:A:55:ILE:HD12	1:A:60:LEU:HD21	1.92	0.51
1:B:55:ILE:HD12	1:B:60:LEU:HD21	1.91	0.51
1:B:128:ILE:CG2	1:B:135:ARG:HB2	2.40	0.51
1:B:371:LEU:HG	1:B:401:ILE:HD12	1.92	0.50
1:A:135:ARG:HB3	1:A:280:LEU:HD22	1.94	0.50
2:D:465:ASN:O	2:D:466:SER:HB3	2.12	0.49
1:B:137:THR:O	1:B:138:SER:C	2.50	0.49
1:B:109:LEU:HD21	1:B:278:LEU:HD21	1.95	0.49
1:A:53:VAL:CG1	1:A:75:ILE:HG23	2.42	0.49
1:B:148:LEU:HD21	1:B:180:ASP:C	2.33	0.49
1:B:168:THR:HG22	1:B:169:VAL:H	1.78	0.48
2:C:467:ASN:N	2:C:467:ASN:ND2	2.60	0.48
1:B:357:LYS:HB2	1:B:374:VAL:HG22	1.95	0.48

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:426:PRO:HB2	2:C:480:ARG:HB3	1.96	0.47
1:B:253:PRO:HG2	1:B:258:LEU:HD21	1.95	0.47
1:B:168:THR:HG23	2:D:483:TYR:HH	1.77	0.46
1:B:139:LEU:HD22	1:B:280:LEU:HD11	1.98	0.45
1:B:137:THR:CG2	1:B:138:SER:H	2.29	0.45
1:A:418:ARG:O	1:A:420:LYS:N	2.50	0.45
2:C:467:ASN:N	2:C:467:ASN:HD22	2.15	0.45
2:C:465:ASN:HA	2:C:466:SER:HA	1.73	0.43
1:B:40:ALA:HB3	1:B:41:GLU:OE1	2.19	0.43
1:B:418:ARG:O	1:B:420:LYS:N	2.50	0.43
1:A:135:ARG:HH21	1:A:254:SER:HA	1.83	0.43
1:B:136:LYS:HD3	1:B:164:GLN:HG2	2.00	0.43
1:A:377:THR:HG22	1:A:379:SER:H	1.83	0.42
1:A:135:ARG:HG3	1:A:136:LYS:N	2.35	0.41
1:B:427:VAL:HG22	2:D:480:ARG:O	2.20	0.41
1:B:377:THR:HG22	1:B:379:SER:H	1.83	0.41
1:A:102:THR:OG1	1:A:310:LEU:HB3	2.20	0.41
1:A:107:TYR:CD1	1:A:107:TYR:N	2.88	0.41
1:A:102:THR:HG21	1:A:310:LEU:HA	2.03	0.41
1:A:215:ILE:HG22	1:A:219:LYS:HD3	2.03	0.41
1:A:278:LEU:HB3	1:A:307:ILE:CD1	2.51	0.41
1:A:345:ALA:H	2:C:466:SER:HB3	1.86	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	419/445 (94%)	395 (94%)	21 (5%)	3 (1%)	26	55
1	B	426/445 (96%)	408 (96%)	15 (4%)	3 (1%)	26	55
2	C	24/110 (22%)	20 (83%)	3 (12%)	1 (4%)	3	7

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	29/110 (26%)	25 (86%)	2 (7%)	2 (7%)	1	2
All	All	898/1110 (81%)	848 (94%)	41 (5%)	9 (1%)	19	45

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	309	LYS
1	A	419	ARG
1	B	138	SER
1	B	419	ARG
2	D	464	SER
2	C	465	ASN
2	D	466	SER
1	B	40	ALA
1	A	138	SER

### 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	383/401 (96%)	369 (96%)	14 (4%)	41	72
1	B	386/401 (96%)	368 (95%)	18 (5%)	32	63
2	C	26/98 (26%)	22 (85%)	4 (15%)	3	8
2	D	30/98 (31%)	28 (93%)	2 (7%)	20	44
All	All	825/998 (83%)	787 (95%)	38 (5%)	34	64

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	LYS
1	A	109	LEU
1	A	112	MET
1	A	118	MET
1	A	127	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	224	GLU
1	A	260	GLU
1	A	283	GLU
1	A	310	LEU
1	A	323	SER
1	A	380	ASN
1	A	394	ARG
1	A	418	ARG
1	A	422	ARG
1	B	41	GLU
1	B	53	VAL
1	B	64	ASP
1	B	109	LEU
1	B	114[A]	GLU
1	B	114[B]	GLU
1	B	136	LYS
1	B	137	THR
1	B	168	THR
1	B	219	LYS
1	B	224	GLU
1	B	232	VAL
1	B	245	ARG
1	B	323	SER
1	B	364	ASN
1	B	380	ASN
1	B	418	ARG
1	B	430	ARG
2	C	464	SER
2	C	465	ASN
2	C	467	ASN
2	C	487	SER
2	D	476	ASN
2	D	487	SER

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

Mol	Chain	Res	Type
2	C	465	ASN
2	C	467	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 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 ⓘ

### 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	422/445 (94%)	0.09	16 (3%) 44 44	32, 62, 102, 122	0
1	B	427/445 (95%)	-0.03	12 (2%) 56 57	27, 49, 90, 113	0
2	C	28/110 (25%)	0.38	3 (10%) 8 6	32, 49, 102, 111	0
2	D	33/110 (30%)	0.06	2 (6%) 25 23	27, 40, 99, 105	0
All	All	910/1110 (81%)	0.04	33 (3%) 46 46	27, 55, 98, 122	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	466	SER	6.5
1	A	284	THR	4.2
1	B	137	THR	3.8
1	B	284	THR	3.6
2	D	469	ASN	3.0
1	A	417	LYS	2.9
1	A	283	GLU	2.8
1	B	418	ARG	2.8
1	B	135	ARG	2.8
1	B	93	ASN	2.7
1	B	315	ALA	2.5
2	C	464	SER	2.5
2	C	465	ASN	2.5
1	B	414	VAL	2.5
2	D	463	ILE	2.5
1	A	136	LYS	2.5
1	A	85	TRP	2.5
1	A	362	PHE	2.5
1	A	93	ASN	2.4
1	A	281	CYS	2.4
1	A	105	TYR	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	350	TYR	2.3
1	B	42	GLY	2.3
1	A	415	ASN	2.3
1	A	394	ARG	2.3
1	A	197	ALA	2.3
1	B	286	PRO	2.2
1	A	137	THR	2.2
1	A	316	VAL	2.2
1	B	40	ALA	2.2
1	A	286	PRO	2.1
1	B	415	ASN	2.1
1	A	419	ARG	2.0

## 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.