



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:24 AM GMT

PDB ID : 2ABR  
Title : Structure of D280A arginine deiminase with L-arginine forming a S-alkylthiouronium reaction intermediate  
Authors : Galkin, A.; Lu, X.; Dunaway-Mariano, D.; Herzberg, O.  
Deposited on : 2005-07-15  
Resolution : 2.90 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  
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

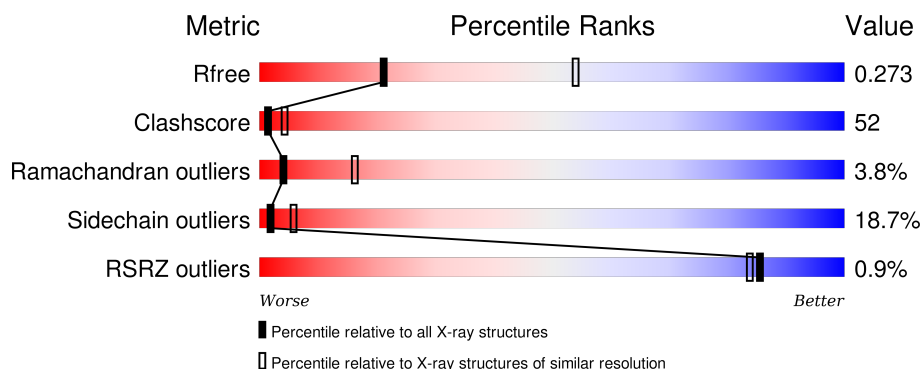
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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	418	<div> <div>%</div> <div> <div></div> <div>30%</div> <div>51%</div> <div>14%</div> <div>• •</div> </div> </div>
1	B	418	<div> <div></div> <div> <div>32%</div> <div>50%</div> <div>15%</div> <div>•</div> </div> </div>
1	C	418	<div> <div>%</div> <div> <div></div> <div>31%</div> <div>50%</div> <div>14%</div> <div>• •</div> </div> </div>
1	D	418	<div> <div>%</div> <div> <div></div> <div>33%</div> <div>49%</div> <div>14%</div> <div>•</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12637 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.

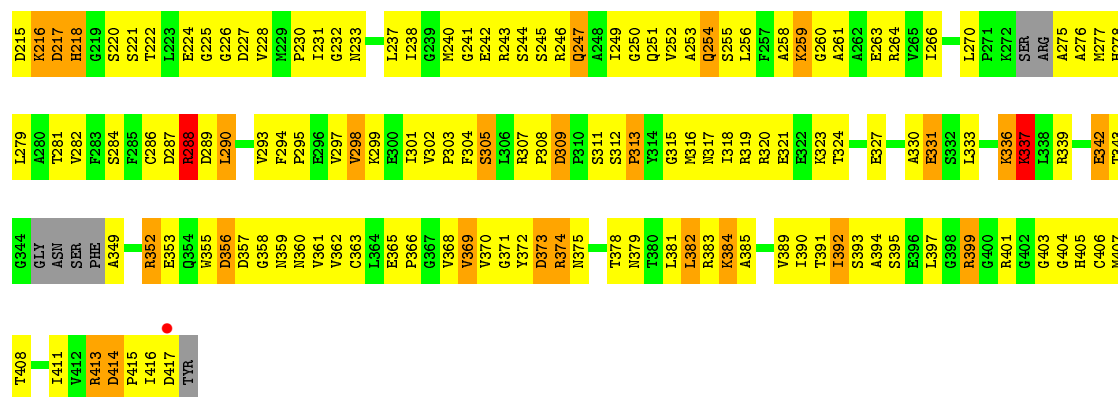
- Molecule 1 is a protein called Arginine deiminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	0	0
			3151	1994	549	591	17			
1	B	406	Total	C	N	O	S	0	0	0
			3174	2008	553	596	17			
1	C	403	Total	C	N	O	S	0	0	0
			3158	1998	550	593	17			
1	D	402	Total	C	N	O	S	0	0	0
			3154	1996	552	589	17			

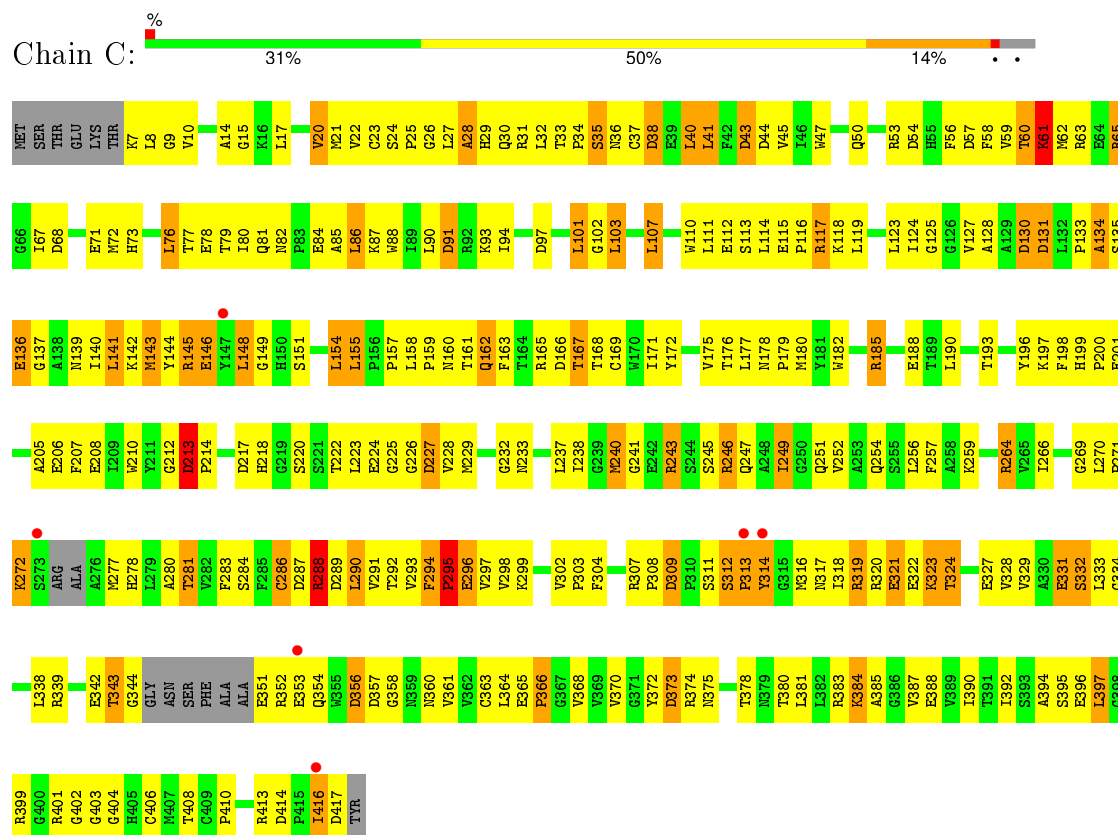
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	280	ALA	ASP	ENGINEERED	UNP P13981
A	406	CYR	CYS	MODIFIED RESIDUE	UNP P13981
B	280	ALA	ASP	ENGINEERED	UNP P13981
B	406	CYR	CYS	MODIFIED RESIDUE	UNP P13981
C	280	ALA	ASP	ENGINEERED	UNP P13981
C	406	CYR	CYS	MODIFIED RESIDUE	UNP P13981
D	280	ALA	ASP	ENGINEERED	UNP P13981
D	406	CYR	CYS	MODIFIED RESIDUE	UNP P13981

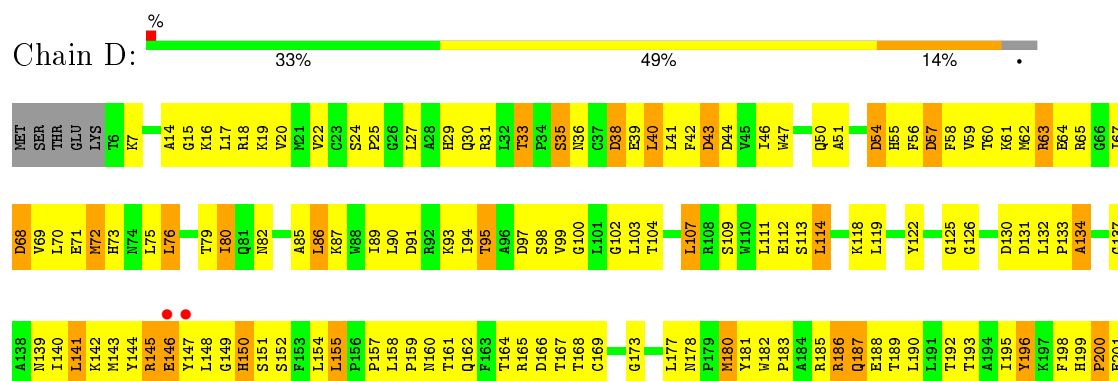




• Molecule 1: Arginine deiminase



• Molecule 1: Arginine deiminase



V412	R413	D414	P415	T416	ASP	TYR	ASN	SER	PHE	ALA	ALA	GLU	R352	E353	Q354	W355	D356	D357	G358	R359	N360	L364	E365	P366	G367	V368	V369	V370	D373	R374	Y377	T378	N379	T380	L381	L382	R383	K384	V387	I390	T391	I392	S393	A394	S395	E396	L397	G398	R399	G400	R401	G402	G403	G404	H405	C406	M407	T408	C409	P410	I411
F202	A203	E208	I209	W210	D213	P214	D215	K216	D217	H218	G219	S220	S221	T222	L223	F224	G225	G226	D227	P230	I231	G232	N233	G234	V235	V236	L237	T238	G239	W240	G241	E242	R243	S244	S245	R246	Q247	A248	T249	G250	Q251	V252	L256	F257	R264	V265	I266	K272	S273	R274	ALA	ALA	N277								
H278	L279	V282	F283	S284	F285	C286	D287	R288	D289	L290	V291	T292	V293	F294	P295	E296	V297	V298	K299	F304	S305	L306	R307	P308	D309	P310	V311	S312	P313	Y314	I318	R319	R320	E321	T324	E327	V328	V329	A330	E331	S332	L333	L335	K336	K337	L338	R339	V340	V341	E342	T343	G344	GLY								

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.60Å 120.60Å 147.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.90 19.92 – 2.90	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.90) 96.0 (19.92-2.90)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.76 (at 2.88Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.214 , 0.274 0.212 , 0.273	Depositor DCC
$R_{free}$ test set	1714 reflections (4.91%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.7	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 56.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	1 of 34924 reflections (0.003%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	12637	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.67 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.7701e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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: CYR

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	0.47	0/3199	0.76	13/4336 (0.3%)
1	B	0.47	0/3222	0.75	11/4368 (0.3%)
1	C	0.46	0/3206	0.78	17/4345 (0.4%)
1	D	0.47	0/3202	0.77	14/4339 (0.3%)
All	All	0.47	0/12829	0.77	55/17388 (0.3%)

There are no bond length outliers.

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	217	ASP	CB-CG-OD2	6.44	124.09	118.30
1	B	373	ASP	CB-CG-OD2	6.29	123.96	118.30
1	A	227	ASP	CB-CG-OD2	6.15	123.83	118.30
1	B	44	ASP	CB-CG-OD2	6.12	123.81	118.30
1	D	217	ASP	CB-CG-OD2	5.93	123.64	118.30

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	3151	0	3139	345	0
1	B	3174	0	3160	298	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3158	0	3143	347	0
1	D	3154	0	3148	388	0
All	All	12637	0	12590	1319	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 52.

The worst 5 of 1319 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:D:43:ASP:CB	1:D:401:ARG:HH12	1.37	1.36
1:D:43:ASP:HB2	1:D:401:ARG:NH1	1.55	1.20
1:D:14:ALA:O	1:D:366:PRO:HD3	1.39	1.19
1:D:352:ARG:HB3	1:D:352:ARG:NH1	1.59	1.17
1:A:306:LEU:HD13	1:A:318:ILE:HG23	1.21	1.17

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	395/418 (94%)	313 (79%)	65 (16%)	17 (4%)	3	13
1	B	399/418 (96%)	332 (83%)	50 (12%)	17 (4%)	3	13
1	C	396/418 (95%)	323 (82%)	61 (15%)	12 (3%)	5	22
1	D	395/418 (94%)	327 (83%)	54 (14%)	14 (4%)	4	18
All	All	1585/1672 (95%)	1295 (82%)	230 (14%)	60 (4%)	4	16

5 of 60 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	63	ARG
1	A	117	ARG
1	A	203	ALA
1	A	221	SER
1	A	249	ILE

### 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	342/353 (97%)	278 (81%)	64 (19%)	2	6
1	B	342/353 (97%)	279 (82%)	63 (18%)	2	6
1	C	342/353 (97%)	276 (81%)	66 (19%)	2	5
1	D	342/353 (97%)	279 (82%)	63 (18%)	2	6
All	All	1368/1412 (97%)	1112 (81%)	256 (19%)	2	6

5 of 256 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	342	GLU
1	C	119	LEU
1	D	288	ARG
1	B	374	ARG
1	C	35	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	178	ASN
1	D	30	GLN
1	C	218	HIS
1	C	162	GLN
1	C	199	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 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	$\# Z  > 2$
1	CYR	A	406	1	11,16,17	2.31	3 (27%)	6,19,21	1.67	1 (16%)
1	CYR	B	406	1	11,16,17	2.35	3 (27%)	6,19,21	2.17	2 (33%)
1	CYR	C	406	1	11,16,17	2.26	3 (27%)	6,19,21	2.07	3 (50%)
1	CYR	D	406	1	11,16,17	2.43	3 (27%)	6,19,21	1.87	2 (33%)

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
1	CYR	A	406	1	-	0/10/18/20	0/0/0/0
1	CYR	B	406	1	-	0/10/18/20	0/0/0/0
1	CYR	C	406	1	-	1/10/18/20	0/0/0/0
1	CYR	D	406	1	-	0/10/18/20	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	406	CYR	CB-SG	-4.56	1.75	1.81
1	B	406	CYR	CB-SG	-4.45	1.75	1.81
1	C	406	CYR	CB-SG	-3.60	1.77	1.81
1	A	406	CYR	CB-SG	-3.50	1.77	1.81
1	B	406	CYR	C7-N6	4.25	1.45	1.36

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	406	CYR	N6-C7-N7	-4.37	109.51	120.88
1	C	406	CYR	N6-C7-N7	-3.89	110.76	120.88
1	D	406	CYR	N6-C7-N7	-3.43	111.96	120.88
1	A	406	CYR	N6-C7-N7	-3.38	112.09	120.88
1	C	406	CYR	O-C-CA	-2.37	119.32	125.49

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	C	406	CYR	SG-C7-N6-C5

There are no ring outliers.

3 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	406	CYR	3	0
1	C	406	CYR	3	0
1	D	406	CYR	5	0

## 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	401/418 (95%)	-0.49	3 (0%) 89 88	12, 40, 72, 99	0
1	B	405/418 (96%)	-0.54	1 (0%) 95 95	13, 37, 73, 99	0
1	C	402/418 (96%)	-0.43	6 (1%) 76 74	14, 45, 79, 98	0
1	D	401/418 (95%)	-0.42	5 (1%) 81 78	17, 43, 79, 98	0
All	All	1609/1672 (96%)	-0.47	15 (0%) 85 84	12, 41, 77, 99	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	273	SER	4.4
1	D	147	TYR	3.5
1	A	134	ALA	3.4
1	D	353	GLU	3.2
1	C	313	PRO	2.9

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	CYR	D	406	17/18	0.97	0.10	-	15,40,60,62	0
1	CYR	B	406	17/18	0.93	0.15	-	22,42,60,63	0
1	CYR	C	406	17/18	0.95	0.12	-	21,34,46,50	0
1	CYR	A	406	17/18	0.96	0.13	-	14,45,58,60	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.