



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2024 – 12:38 PM EST

PDB ID : 4DOP  
Title : Crystal structure of the CusBA heavy-metal efflux complex from Escherichia coli, R mutant  
Authors : Su, C.-C.; Long, F.; Yu, E.  
Deposited on : 2012-02-10  
Resolution : 4.20 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 : 2.36  
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) : 2.36

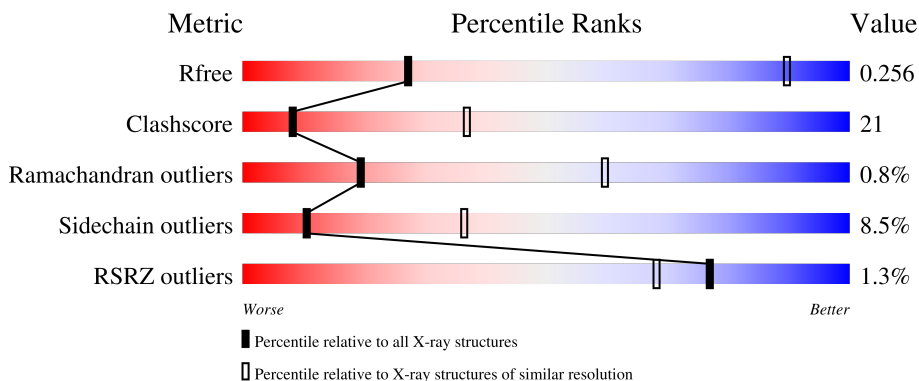
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.20 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))
$R_{free}$	130704	1005 (4.62-3.78)
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)
RSRZ outliers	127900	1063 (4.70-3.70)

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 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	B	413	 52% 22% 22%
1	C	413	 52% 24% 22%
2	A	1054	 2% 51% 42% 5%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12891 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 Cation efflux system protein CusB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	322	Total	C	N	O	S	0	0	0
			2458	1555	428	469	6			
1	C	324	Total	C	N	O	S	0	0	0
			2473	1563	430	474	6			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	408	HIS	-	expression tag	UNP P77239
B	409	HIS	-	expression tag	UNP P77239
B	410	HIS	-	expression tag	UNP P77239
B	411	HIS	-	expression tag	UNP P77239
B	412	HIS	-	expression tag	UNP P77239
B	413	HIS	-	expression tag	UNP P77239
C	408	HIS	-	expression tag	UNP P77239
C	409	HIS	-	expression tag	UNP P77239
C	410	HIS	-	expression tag	UNP P77239
C	411	HIS	-	expression tag	UNP P77239
C	412	HIS	-	expression tag	UNP P77239
C	413	HIS	-	expression tag	UNP P77239

- Molecule 2 is a protein called Cation efflux system protein CusA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	1031	Total	C	N	O	S	0	0	0
			7942	5137	1330	1438	37			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	expression tag	UNP P38054
A	-5	HIS	-	expression tag	UNP P38054

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	HIS	-	expression tag	UNP P38054
A	-3	HIS	-	expression tag	UNP P38054
A	-2	HIS	-	expression tag	UNP P38054
A	-1	HIS	-	expression tag	UNP P38054
A	0	HIS	-	expression tag	UNP P38054
A	669	ALA	ARG	engineered mutation	UNP P38054

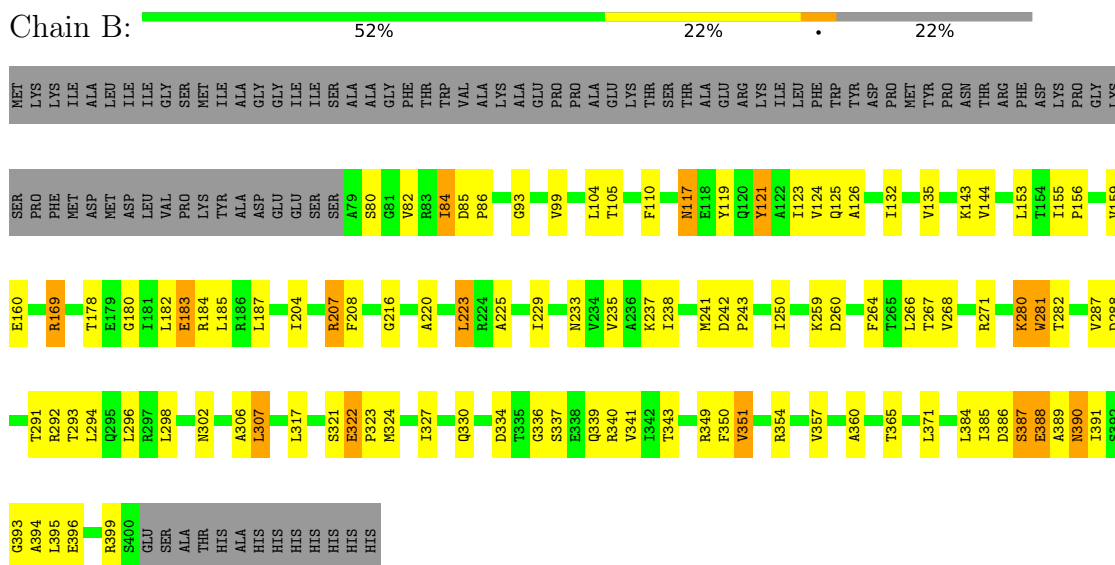
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	4	Total O 4 4	0	0
3	C	11	Total O 11 11	0	0
3	A	3	Total O 3 3	0	0

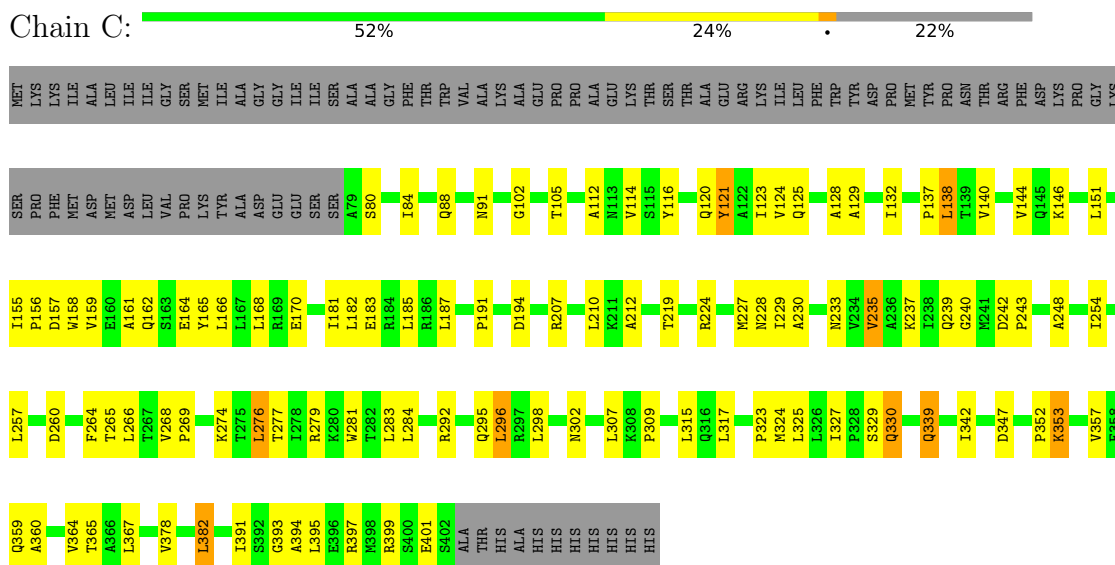
### 3 Residue-property plots [i](#)

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: Cation efflux system protein CusB



- Molecule 1: Cation efflux system protein CusB



- Molecule 2: Cation efflux system protein CusA



MET	HIS	HIS	HIS	HIS	HIS	HIS	M1	L2	E3	W4	I5	I6	R7	R8	S9	V10	A11	M12	R13	F14	L15	V16	L17	M18	G19	A20	L21	F22	L23	R24	L25	W26	G27	S126	T28	W29	L31	L32	V36	D37	A38	L39	P40	D41	R157	L144	V145	D146	R260	D261	L262	A263	L159	F163	P269	E270	L169	M271	P172	D173	V177	K185	L69
M73	K80	T81	P85	G89	Y92	V93	Y94	F97	D102	S113	N114	Q215	I222	E223	L224	Y229	G235	Y236	L237	Q238	L118	T239	L240	L121	V125	A127	E128	L129	E141	V255	L144	V145	D146	L259	R260	D261	L262	A263	L159	F163	P269	E270	L169	M271	P172	D173	V177	K185	L69														
Q188	I191	D192	P193	Q194	R195	L196	L201	S202	L203	L210	A212	N214	Q215	I222	E223	L224	Y229	G235	Y236	L237	Q238	L118	T239	L240	L121	V125	A127	E128	L129	E141	V255	L144	V145	D146	L259	R260	D261	L262	A263	L159	F163	P269	E270	L169	M271	P172	D173	V177	K185	L69													
V284	V288	I290	S293	G294	K295	A297	R298	E299	V300	I301	L311	L315	V322	Y325	S328	I335	L338	S339	Q340	G341	K341	L342	L343	E344	E345	V348	V349	A350	V351	V352	C353	A354	L355	F356	W358	H359	V360	R361	S362	A363	L364	R367	I368	S369	L370	P371	L69																
L372	G373	L374	C375	I376	A377	F378	I379	V380	M381	M389	I390	R391	S392	G395	I396	I397	A398	A399	V400	M403	V404	D405	I408	V409	M410	Q411	I412	M413	A414	H415	K416	R417	L418	E419	A420	W421	Q422	H423	Q424	H425	P426	D427	A428	T429	L430	D431	M432	K433	T434	R435	W436	Q437	V438	I439	T440	L69							
D441	A442	S443	V444	E445	G447	P448	A449	V450	L450	S453	L454	L455	L456	L457	T458	S460	F461	P463	T466	G469	R473	L474	F475	G476	P477	T481	K482	M486	A487	G488	G589	E590	F591	L592	L593	L594	D595	L570	L571	Y572	I499	L500	M501	W504	I505	ARG	GLY	L506	PRO	I439	T440	L69											
SER	SER	ASN	PRO	L517	N518	R519	F520	L521	L522	R523	V524	Y525	H526	P527	L528	L529	L530	K531	V532	L533	H534	E535	T536	P537	T538	T539	S546	V547	W548	L549	V550	V551	L552	P553	L554	G559	E560	F561	L562	P563	D569	L570	L571	Y572	I499	L500	M501	W504	I505	ARG	GLY	L506	PRO	I439	T440	L69							
K594	L595	I596	V599	P600	E601	V602	A603	R604	G607	K608	T609	P620	L621	E622	H623	E624	E625	T626	T627	L628	E634	H635	W636	R637	P638	M642	E647	L649	R650	M651	T652	V653	R654	L655	P656	G657	L658	A659	L571	Y572	I499	L500	M501	W504	I505	ARG	GLY	L506	PRO	I439	T440	L69											
I677	K678	S679	L683	T688	V689	L690	L693	M696	A697	E701	L702	E702	L703	R705	T706	V707	P708	L718	R722	Y723	L724	N725	L728	N729	K732	Y736	F746	S749	A750	L661	V662	V663	P664	P665	R667	L673	S674	T675	L676	N769	L770	L69																					
R771	Y772	P773	Q774	S775	S779	P780	Q781	R784	Q785	T788	L789	I796	T797	A802	D803	I804	R805	S807	K814	T815	A818	R819	G820	H821	H822	H823	H824	H825	H826	A827	A828	R831	D832	M833	V834	V837	H838	D839	L840	Q841	K842	A843	I844	A845	E846	K847	L850	R851	P852	G853	T854	S855	V856	L69									
A857	F858	S859	G860	E863	L864	L865	K871	L872	K873	L874	M875	P877	T878	T879	L880	M881	I882	I883	F884	L887	A890	F891	R892	R893	V894	G895	E896	I900	V904	P905	L908	V909	G910	G911	I912	H913	L914	S923	V924	A925	F930	I931	A932	L933	V936	A937	E938	I939	F940	L69													
G941	V942	Y943	N944	L945	M946	Y947	A951	L952	E953	A954	V955	P956	S957	L958	N959	P960	P961	Q962	T963	F964	S965	K968	L969	A972	L973	Y974	H975	G976	A977	Y978	R980	V981	P982	P983	K984	A985	M986	V990	I991	L992	A993	G994	L995	L996	P997	G1001	T1002	G1005	S1006	E1007	V1008	L69											
M1009	S1010	R1011	I1012	A1013	A1014	P1015	M1016	I1017	M1020	I1021	T1022	A1023	P1024	L1025	L1026	I1031	P1032	A1033	A1034	Y1035	K1036	L1037	M1038	W1039	L1040	H1043	ARG	VAL	ARG	LNS	R980	V981	P982	P983	K984	A985	M986	V990	I991	L992	A993	G994	L995	L996	P997	G1001	T1002	G1005	S1006	E1007	V1008	L69											

## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.52Å 160.52Å 681.33Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	97.31 – 4.20 136.21 – 4.20	Depositor EDS
% Data completeness (in resolution range)	92.2 (97.31-4.20) 99.9 (136.21-4.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.33	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 4.15Å)	Xtrriage
Refinement program	PHENIX 1.6.4_486	Depositor
R, $R_{free}$	0.222 , 0.268 0.209 , 0.256	Depositor DCC
$R_{free}$ test set	1284 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.2	Xtrriage
Anisotropy	0.443	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 128.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	12891	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	111.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% 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	B	0.22	0/2498	0.43	0/3401
1	C	0.22	0/2513	0.42	0/3421
2	A	0.22	0/8108	0.41	0/11041
All	All	0.22	0/13119	0.41	0/17863

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	B	2458	0	2522	92	0
1	C	2473	0	2533	80	0
2	A	7942	0	8187	399	0
3	A	3	0	0	0	0
3	B	4	0	0	0	0
3	C	11	0	0	0	0
All	All	12891	0	13242	550	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 21.

The worst 5 of 550 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)
2:A:573:MET:HG2	2:A:678:LYS:NZ	1.80	0.96
1:C:242:ASP:HB3	1:C:243:PRO:HD3	1.49	0.94
2:A:696:MET:HG3	2:A:854:THR:HG21	1.50	0.93
2:A:573:MET:SD	2:A:678:LYS:HD2	2.11	0.90
2:A:574:PRO:HB2	2:A:658:LEU:HD11	1.52	0.89

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	B	320/413 (78%)	281 (88%)	36 (11%)	3 (1%)	17	56
1	C	322/413 (78%)	300 (93%)	21 (6%)	1 (0%)	41	76
2	A	1027/1054 (97%)	922 (90%)	95 (9%)	10 (1%)	15	54
All	All	1669/1880 (89%)	1503 (90%)	152 (9%)	14 (1%)	19	60

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	388	GLU
2	A	638	PRO
1	B	337	SER
1	B	390	ASN
2	A	894	VAL

### 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	B	263/338 (78%)	243 (92%)	20 (8%)	13	40
1	C	265/338 (78%)	245 (92%)	20 (8%)	13	40
2	A	849/871 (98%)	772 (91%)	77 (9%)	9	32
All	All	1377/1547 (89%)	1260 (92%)	117 (8%)	10	36

5 of 117 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	260	ARG
2	A	945	LEU
2	A	417	ARG
2	A	940	PHE
2	A	765	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	660	ASN
2	A	729	ASN
2	A	795	GLN
1	C	302	ASN
1	C	263	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](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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	B	322/413 (77%)	0.04	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	36, 86, 147, 228	0
1	C	324/413 (78%)	0.11	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	35, 88, 138, 216	0
2	A	1031/1054 (97%)	0.27	21 (2%) <span style="border: 1px solid blue; padding: 2px;">65</span> <span style="border: 1px solid blue; padding: 2px;">56</span>	30, 110, 226, 401	0
All	All	1677/1880 (89%)	0.20	21 (1%) <span style="border: 1px solid blue; padding: 2px;">77</span> <span style="border: 1px solid blue; padding: 2px;">68</span>	30, 99, 210, 401	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	446	VAL	6.9
2	A	445	GLU	5.1
2	A	954	ALA	4.4
2	A	429	THR	3.9
2	A	450	LEU	3.6

### 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

There are no such residues in this entry.