



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:14 PM GMT

PDB ID : 4L0D  
Title : Crystal structure of delta516-525 human cystathionine beta-synthase containing C-terminal 6xHis-tag  
Authors : Ereno, J.; Majtan, T.; Oyenarte, I.; Kraus, J.P.; Martinez-Cruz, L.A.  
Deposited on : 2013-05-31  
Resolution : 2.97 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

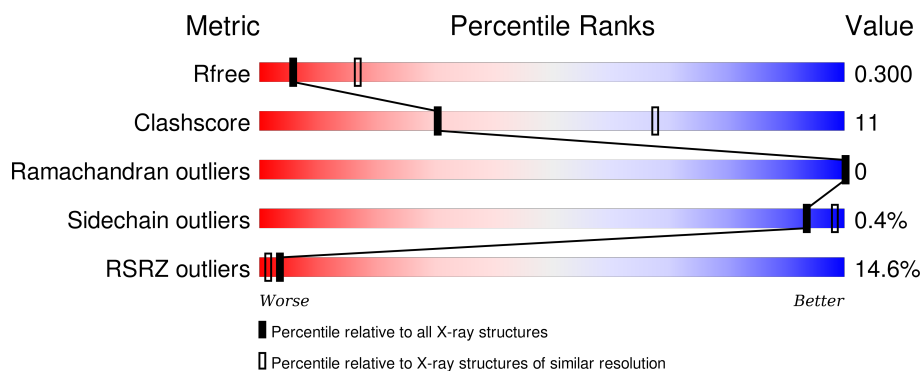
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.97 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	1992 (3.00-2.96)
Clashscore	102246	2349 (3.00-2.96)
Ramachandran outliers	100387	2274 (3.00-2.96)
Sidechain outliers	100360	2277 (3.00-2.96)
RSRZ outliers	91569	2007 (3.00-2.96)

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	549	<div> <div>11%</div> <div> <div></div> <div>67%</div> <div>23%</div> <div>10%</div> </div> </div>
1	B	549	<div> <div>15%</div> <div> <div></div> <div>68%</div> <div>22%</div> <div>10%</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7728 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 Cystathionine beta-synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	496	Total	C	N	O	S	0	1	0
			3807	2413	667	705	22			
1	B	496	Total	C	N	O	S	0	1	0
			3805	2412	666	705	22			

There are 38 discrepancies between the modelled and reference sequences:

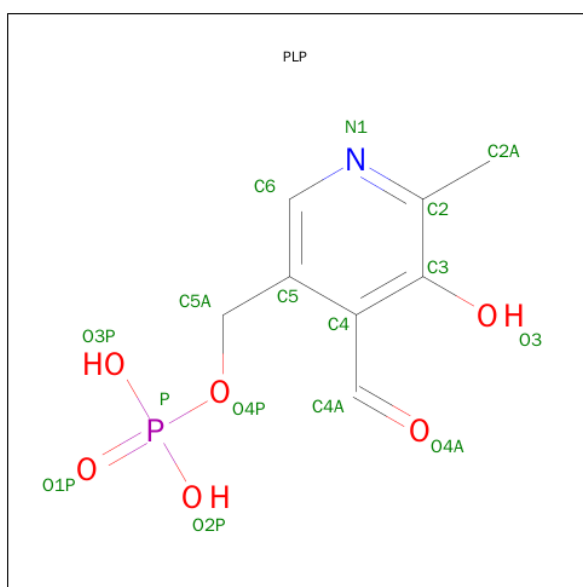
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	PRO	ENGINEERED MUTATION	UNP P35520
A	?	-	ILE	DELETION	UNP P35520
A	?	-	GLN	DELETION	UNP P35520
A	?	-	TYR	DELETION	UNP P35520
A	?	-	HIS	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	?	-	THR	DELETION	UNP P35520
A	?	-	GLY	DELETION	UNP P35520
A	?	-	LYS	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	?	-	SER	DELETION	UNP P35520
A	552	GLU	-	EXPRESSION TAG	UNP P35520
A	553	LEU	-	EXPRESSION TAG	UNP P35520
A	554	HIS	-	EXPRESSION TAG	UNP P35520
A	555	HIS	-	EXPRESSION TAG	UNP P35520
A	556	HIS	-	EXPRESSION TAG	UNP P35520
A	557	HIS	-	EXPRESSION TAG	UNP P35520
A	558	HIS	-	EXPRESSION TAG	UNP P35520
A	559	HIS	-	EXPRESSION TAG	UNP P35520
B	2	GLY	PRO	ENGINEERED MUTATION	UNP P35520
B	?	-	ILE	DELETION	UNP P35520
B	?	-	GLN	DELETION	UNP P35520
B	?	-	TYR	DELETION	UNP P35520
B	?	-	HIS	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520

*Continued on next page...*

Continued from previous page...

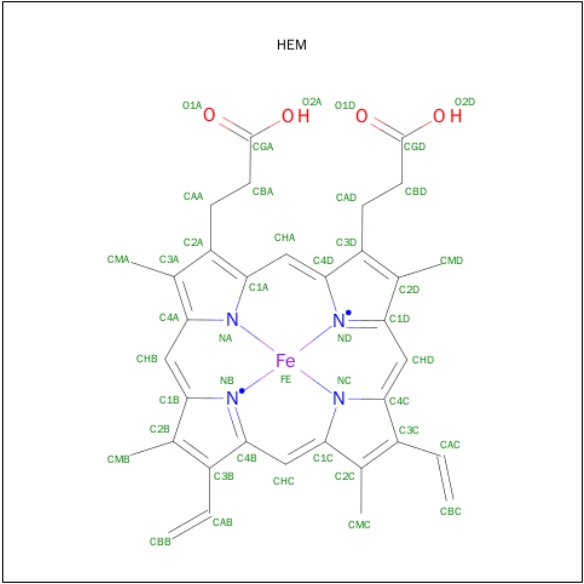
Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	THR	DELETION	UNP P35520
B	?	-	GLY	DELETION	UNP P35520
B	?	-	LYS	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520
B	?	-	SER	DELETION	UNP P35520
B	552	GLU	-	EXPRESSION TAG	UNP P35520
B	553	LEU	-	EXPRESSION TAG	UNP P35520
B	554	HIS	-	EXPRESSION TAG	UNP P35520
B	555	HIS	-	EXPRESSION TAG	UNP P35520
B	556	HIS	-	EXPRESSION TAG	UNP P35520
B	557	HIS	-	EXPRESSION TAG	UNP P35520
B	558	HIS	-	EXPRESSION TAG	UNP P35520
B	559	HIS	-	EXPRESSION TAG	UNP P35520

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula:  $C_8H_{10}NO_6P$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			15	8	1	5	1		
2	B	1	Total	C	N	O	P	0	0
			15	8	1	5	1		

- Molecule 3 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula:  $C_{34}H_{32}FeN_4O_4$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		
3	B	1	Total	C	Fe	N	O	0	0
			43	34	1	4	4		



P282	E283	G284	S285	I286	P290	Q295	T296	E297	Q298	T299	T300	V303	D309	F310	T313	D316	R317	T318	V319	V320	W323	D328	F334	L338	G346	G347	G348	S349	A350	L364	L374	P375	R379	K384	F385	L386	S387	D388	R389	Q393	K394	G395	F396																																																																																																																																																																																																																																																																																																																																																																																																																																																													
L397	K398	E399	E400	T403	W408	W409	W410	H411	L412	R413	V414	Q415	E416	L417	Q418	L419	L423	P427	T430	H433	F443	V449	D450	E451	A452	Q453	V454	V459	G462	L468	K472	G480	K481	V482	L483	F487	K488	Q489	I490	R491	L492	T493	D494	T495	L496																																																																																																																																																																																																																																																																																																																																																																																																																																																											
G497	R498	L499	S500	H501	I502	L503	E504	M505	D506	H507	F508	A509	L510	V511	V512	H513	E514	G515	ARG	Q528	M529	V530	F531	G532	V533	V534	T535	A536	L539	L540	N541	F542	V543	A544	E547	R548	D549	G550	L551	G552	L553	H554	H555	H556	H557	H558	H559	H560	H561	H562	H563	H564	H565	H566	H567	H568	H569	H570	H571	H572	H573	H574	H575	H576	H577	H578	H579	H580	H581	H582	H583	H584	H585	H586	H587	H588	H589	H590	H591	H592	H593	H594	H595	H596	H597	H598	H599	H600	H601	H602	H603	H604	H605	H606	H607	H608	H609	H610	H611	H612	H613	H614	H615	H616	H617	H618	H619	H620	H621	H622	H623	H624	H625	H626	H627	H628	H629	H630	H631	H632	H633	H634	H635	H636	H637	H638	H639	H640	H641	H642	H643	H644	H645	H646	H647	H648	H649	H650	H651	H652	H653	H654	H655	H656	H657	H658	H659	H660	H661	H662	H663	H664	H665	H666	H667	H668	H669	H670	H671	H672	H673	H674	H675	H676	H677	H678	H679	H680	H681	H682	H683	H684	H685	H686	H687	H688	H689	H690	H691	H692	H693	H694	H695	H696	H697	H698	H699	H700	H701	H702	H703	H704	H705	H706	H707	H708	H709	H710	H711	H712	H713	H714	H715	H716	H717	H718	H719	H720	H721	H722	H723	H724	H725	H726	H727	H728	H729	H730	H731	H732	H733	H734	H735	H736	H737	H738	H739	H740	H741	H742	H743	H744	H745	H746	H747	H748	H749	H750	H751	H752	H753	H754	H755	H756	H757	H758	H759	H760	H761	H762	H763	H764	H765	H766	H767	H768	H769	H770	H771	H772	H773	H774	H775	H776	H777	H778	H779	H780	H781	H782	H783	H784	H785	H786	H787	H788	H789	H790	H791	H792	H793	H794	H795	H796	H797	H798	H799	H800	H801	H802	H803	H804	H805	H806	H807	H808	H809	H810	H811	H812	H813	H814	H815	H816	H817	H818	H819	H820	H821	H822	H823	H824	H825	H826	H827	H828	H829	H830	H831	H832	H833	H834	H835	H836	H837	H838	H839	H840	H841	H842	H843	H844	H845	H846	H847	H848	H849	H850	H851	H852	H853	H854	H855	H856	H857	H858	H859	H860	H861	H862	H863	H864	H865	H866	H867	H868	H869	H870	H871	H872	H873	H874	H875	H876	H877	H878	H879	H880	H881	H882	H883	H884	H885	H886	H887	H888	H889	H890	H891	H892	H893	H894	H895	H896	H897	H898	H899	H900	H901	H902	H903	H904	H905	H906	H907	H908	H909	H910	H911	H912	H913	H914	H915	H916	H917	H918	H919	H920	H921	H922	H923	H924	H925	H926	H927	H928	H929	H930	H931	H932	H933	H934	H935	H936	H937	H938	H939	H940	H941	H942	H943	H944	H945	H946	H947	H948	H949	H950	H951	H952	H953	H954	H955	H956	H957	H958	H959	H960	H961	H962	H963	H964	H965	H966	H967	H968	H969	H970	H971	H972	H973	H974	H975	H976	H977	H978	H979	H980	H981	H982	H983	H984	H985	H986	H987	H988	H989	H990	H991	H992	H993	H994	H995	H996	H997	H998	H999	H1000

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.36Å 136.20Å 169.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.06 – 2.97 53.65 – 2.97	Depositor EDS
% Data completeness (in resolution range)	95.6 (53.06-2.97) 97.8 (53.65-2.97)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.07 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.240 , 0.284 0.262 , 0.300	Depositor DCC
$R_{free}$ test set	1498 reflections (5.38%)	DCC
Wilson B-factor (Å <sup>2</sup> )	70.6	Xtriage
Anisotropy	0.557	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 72.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	3 of 29365 reflections (0.010%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	7728	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.97% 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, 0.375 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: HEM, PLP

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	0.37	0/3878	0.56	0/5253
1	B	0.39	0/3881	0.57	0/5261
All	All	0.38	0/7759	0.57	0/10514

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	3807	0	3832	90	0
1	B	3805	0	3834	84	0
2	A	15	0	7	3	0
2	B	15	0	7	3	0
3	A	43	0	30	5	0
3	B	43	0	30	6	0
All	All	7728	0	7740	170	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 11.

The worst 5 of 170 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:413:ARG:NH1	1:A:493:THR:O	1.89	1.05
1:B:513:HIS:HB2	1:B:531:PHE:HE2	1.43	0.84
1:B:266:ARG:HD2	3:B:602:HEM:HBC2	1.63	0.81
1:B:110:GLU:OE2	1:B:121:ARG:NE	2.14	0.81
1:A:58:ARG:NE	1:A:62:GLU:OE1	2.16	0.78

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	489/549 (89%)	474 (97%)	15 (3%)	0	100	100
1	B	493/549 (90%)	479 (97%)	14 (3%)	0	100	100
All	All	982/1098 (89%)	953 (97%)	29 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	407/463 (88%)	406 (100%)	1 (0%)	95	99
1	B	408/463 (88%)	406 (100%)	2 (0%)	92	98

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	815/926 (88%)	812 (100%)	3 (0%)	93	98

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	460	THR
1	B	299	THR
1	B	454	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	380	ASN

### 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 carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 ligands 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	PLP	A	601	1	15,15,16	0.99	1 (6%)	21,22,23	1.26	2 (9%)
3	HEM	A	602	1	30,50,50	4.42	10 (33%)	24,82,82	3.10	10 (41%)
2	PLP	B	601	1	15,15,16	1.03	1 (6%)	21,22,23	1.29	3 (14%)
3	HEM	B	602	1	30,50,50	4.51	11 (36%)	24,82,82	3.01	9 (37%)

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	PLP	A	601	1	-	0/6/6/8	0/1/1/1
3	HEM	A	602	1	-	0/10/54/54	0/0/8/8
2	PLP	B	601	1	-	0/6/6/8	0/1/1/1
3	HEM	B	602	1	-	0/10/54/54	0/0/8/8

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	602	HEM	C3B-C4B	-14.34	1.39	1.51
3	A	602	HEM	C3B-C4B	-13.55	1.39	1.51
3	A	602	HEM	C2D-C3D	-12.38	1.17	1.54
3	B	602	HEM	C2D-C3D	-12.23	1.17	1.54
3	B	602	HEM	C3D-C4D	-9.97	1.38	1.51

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	602	HEM	C3B-C4B-NB	-4.11	103.77	111.63
3	B	602	HEM	C3B-C4B-NB	-4.05	103.89	111.63
3	A	602	HEM	CBA-CAA-C2A	-2.84	107.44	112.53
3	B	602	HEM	CBA-CAA-C2A	-2.81	107.50	112.53
3	A	602	HEM	C1D-CHD-C4C	-2.79	121.16	125.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	PLP	3	0
3	A	602	HEM	5	0
2	B	601	PLP	3	0
3	B	602	HEM	6	0

## 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	496/549 (90%)	1.00	60 (12%) <b>6</b> <b>3</b>	53, 93, 150, 202	0
1	B	496/549 (90%)	1.10	85 (17%) <b>2</b> <b>1</b>	46, 88, 142, 184	0
All	All	992/1098 (90%)	1.05	145 (14%) <b>3</b> <b>1</b>	46, 90, 147, 202	0

The worst 5 of 145 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	397	LEU	6.0
1	A	403	THR	5.7
1	B	403	THR	5.2
1	B	411	HIS	5.2
1	B	416	GLU	4.8

### 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 carbohydrates in this entry.

### 6.4 Ligands [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( $\text{\AA}^2$ )	Q<0.9
3	HEM	B	602	43/43	0.92	0.34	0.77	101,104,104,104	0
3	HEM	A	602	43/43	0.93	0.29	0.61	107,108,109,109	0
2	PLP	B	601	15/16	0.90	0.22	-0.70	53,66,75,76	0
2	PLP	A	601	15/16	0.92	0.20	-1.23	58,75,85,86	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.