



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 13, 2018 – 02:31 pm GMT

PDB ID : 3VQR  
Title : Structure of a dye-linked L-proline dehydrogenase mutant from the aerobic hyperthermophilic archaeon, *Aeropyrum pernix*  
Authors : Sakuraba, H.; Ohshima, T.; Satomura, T.; Yoneda, K.  
Deposited on : 2012-03-29  
Resolution : 2.01 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 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

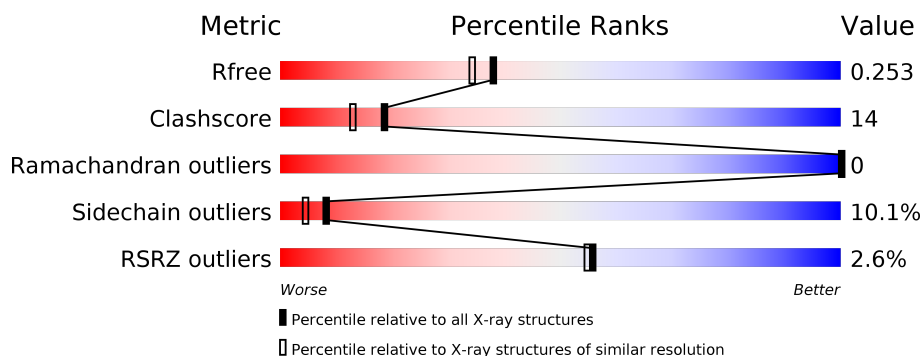
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.01 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	7193 (2.00-2.00)
Clashscore	122126	8267 (2.00-2.00)
Ramachandran outliers	120053	8166 (2.00-2.00)
Sidechain outliers	120020	8165 (2.00-2.00)
RSRZ outliers	108989	7011 (2.00-2.00)

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	447	<div> <div>2%</div> <div> <div></div> <div>71%</div> <div>20%</div> <div>• 5%</div> </div> </div>
1	B	447	<div> <div>3%</div> <div> <div></div> <div>69%</div> <div>22%</div> <div>• 5%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACT	A	1002	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6893 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 Putative oxidoreductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	425	Total	C	N	O	S	0	0	0
			3275	2081	573	612	9			
1	B	425	Total	C	N	O	S	0	0	0
			3275	2081	573	612	9			

There are 42 discrepancies between the modelled and reference sequences:

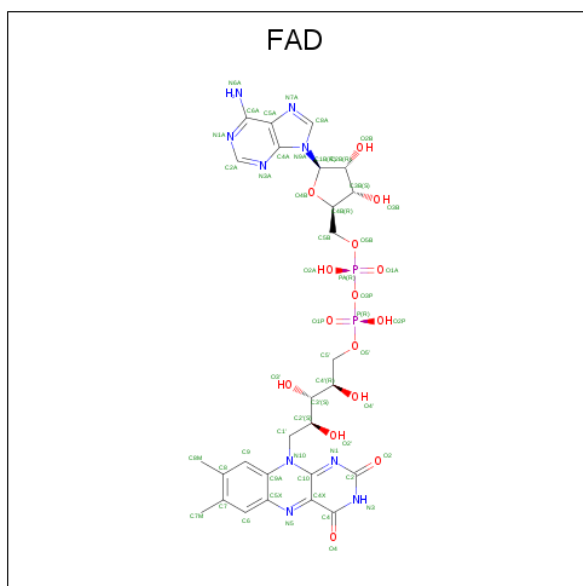
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q9YCJ0
A	-18	GLY	-	EXPRESSION TAG	UNP Q9YCJ0
A	-17	SER	-	EXPRESSION TAG	UNP Q9YCJ0
A	-16	SER	-	EXPRESSION TAG	UNP Q9YCJ0
A	-15	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-14	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-13	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-12	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-11	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-10	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	-9	SER	-	EXPRESSION TAG	UNP Q9YCJ0
A	-8	SER	-	EXPRESSION TAG	UNP Q9YCJ0
A	-7	GLY	-	EXPRESSION TAG	UNP Q9YCJ0
A	-6	LEU	-	EXPRESSION TAG	UNP Q9YCJ0
A	-5	VAL	-	EXPRESSION TAG	UNP Q9YCJ0
A	-4	PRO	-	EXPRESSION TAG	UNP Q9YCJ0
A	-3	ARG	-	EXPRESSION TAG	UNP Q9YCJ0
A	-2	GLY	-	EXPRESSION TAG	UNP Q9YCJ0
A	-1	SER	-	EXPRESSION TAG	UNP Q9YCJ0
A	0	HIS	-	EXPRESSION TAG	UNP Q9YCJ0
A	?	-	LEU	DELETION	UNP Q9YCJ0
B	-19	MET	-	EXPRESSION TAG	UNP Q9YCJ0
B	-18	GLY	-	EXPRESSION TAG	UNP Q9YCJ0
B	-17	SER	-	EXPRESSION TAG	UNP Q9YCJ0
B	-16	SER	-	EXPRESSION TAG	UNP Q9YCJ0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-14	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-13	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-12	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-11	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-10	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-9	SER	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-8	SER	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-7	GLY	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-6	LEU	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-5	VAL	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-4	PRO	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-3	ARG	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-2	GLY	-	EXPRESSION TAG	UNP Q9YJCJ0
B	-1	SER	-	EXPRESSION TAG	UNP Q9YJCJ0
B	0	HIS	-	EXPRESSION TAG	UNP Q9YJCJ0
B	?	-	LEU	DELETION	UNP Q9YJCJ0

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula:  $\text{C}_{27}\text{H}_{33}\text{N}_9\text{O}_{15}\text{P}_2$ ).



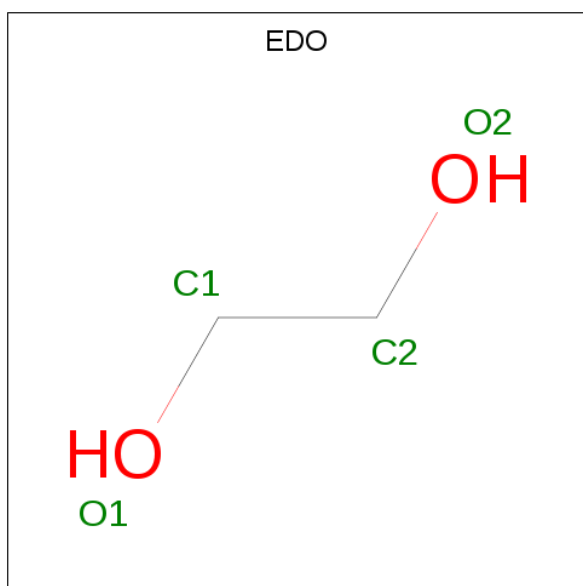
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 53	C 27	N 9	O 15	P 2	0	0
2	B	1	Total 53	C 27	N 9	O 15	P 2	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula:  $\text{C}_2\text{H}_3\text{O}_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	104	Total	O	0	0
			104	104		
5	B	109	Total	O	0	0
			109	109		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	47.99Å 70.57Å 73.09Å 71.65° 70.85° 70.15°	Depositor
Resolution (Å)	24.20 – 2.01 24.20 – 2.01	Depositor EDS
% Data completeness (in resolution range)	(Not available) (24.20-2.01) 95.1 (24.20-2.01)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	8.16 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.210 , 0.255 0.212 , 0.253	Depositor DCC
$R_{free}$ test set	2654 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.8	Xtriage
Anisotropy	0.655	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 34.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.489 for h,h-k,h-l 0.015 for -h,-l,-k 0.014 for -h,-h+l,-h+k	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6893	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.89% 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: EDO, FAD, ACT

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	1.07	2/3349 (0.1%)	1.02	5/4538 (0.1%)
1	B	1.04	2/3349 (0.1%)	1.01	6/4538 (0.1%)
All	All	1.06	4/6698 (0.1%)	1.01	11/9076 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	97	TRP	CD2-CE2	5.81	1.48	1.41
1	B	348	TRP	CD2-CE2	5.81	1.48	1.41
1	B	300	TRP	CD2-CE2	5.49	1.48	1.41
1	A	413	TRP	CD2-CE2	5.37	1.47	1.41

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	202	LEU	CB-CG-CD1	5.93	121.08	111.00
1	A	52	ARG	NE-CZ-NH1	5.88	123.24	120.30
1	A	126	LEU	CA-CB-CG	5.86	128.78	115.30
1	B	122	LEU	CA-CB-CG	5.61	128.20	115.30
1	B	202	LEU	CB-CG-CD2	5.56	120.45	111.00

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	3275	0	3234	86	0
1	B	3275	0	3234	102	0
2	A	53	0	31	0	0
2	B	53	0	31	3	0
3	A	4	0	3	6	0
3	B	4	0	3	1	0
4	A	8	0	12	0	0
4	B	8	0	12	2	0
5	A	104	0	0	8	0
5	B	109	0	0	4	0
All	All	6893	0	6560	188	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 14.

The worst 5 of 188 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:B:336:LEU:HD13	1:B:339:PHE:CD2	1.83	1.13
1:A:261:THR:HG22	1:A:264:LEU:H	1.09	1.09
1:B:332:LEU:HD12	1:B:336:LEU:HD11	1.25	1.07
1:B:191:PRO:HG2	1:B:194:GLU:HG3	1.39	1.05
1:A:76:ARG:HH11	1:A:76:ARG:HG2	1.30	0.96

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	423/447 (95%)	411 (97%)	12 (3%)	0	<b>100</b> <b>100</b>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	423/447 (95%)	417 (99%)	6 (1%)	0	100	100
All	All	846/894 (95%)	828 (98%)	18 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	335/354 (95%)	297 (89%)	38 (11%)	6	3
1	B	335/354 (95%)	305 (91%)	30 (9%)	10	6
All	All	670/708 (95%)	602 (90%)	68 (10%)	8	4

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	315	GLU
1	B	48	MET
1	B	336	LEU
1	A	331	ILE
1	A	408	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	75	GLN
1	B	338	GLN
1	B	239	ASN
1	A	239	ASN
1	B	306	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 ⓘ

8 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	FAD	A	1001	-	51,58,58	2.10	11 (21%)	57,89,89	2.16	13 (22%)
3	ACT	A	1002	-	1,3,3	0.61	0	0,3,3	0.00	-
4	EDO	A	1003	-	3,3,3	0.95	0	2,2,2	0.44	0
4	EDO	A	1004	-	3,3,3	0.46	0	2,2,2	1.29	0
2	FAD	B	1001	-	51,58,58	2.05	11 (21%)	57,89,89	2.15	9 (15%)
3	ACT	B	1002	-	1,3,3	3.87	1 (100%)	0,3,3	0.00	-
4	EDO	B	1003	-	3,3,3	1.03	0	2,2,2	0.48	0
4	EDO	B	1004	-	3,3,3	0.57	0	2,2,2	2.42	1 (50%)

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	FAD	A	1001	-	-	0/28/50/50	0/6/6/6

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ACT	A	1002	-	-	0/0/0/0	0/0/0/0
4	EDO	A	1003	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1004	-	-	0/1/1/1	0/0/0/0
2	FAD	B	1001	-	-	0/28/50/50	0/6/6/6
3	ACT	B	1002	-	-	0/0/0/0	0/0/0/0
4	EDO	B	1003	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1004	-	-	0/1/1/1	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1002	ACT	CH3-C	-3.87	1.43	1.48
2	B	1001	FAD	P-O2P	-2.40	1.43	1.55
2	A	1001	FAD	P-O2P	-2.35	1.43	1.55
2	A	1001	FAD	O4'-C4'	-2.35	1.38	1.43
2	A	1001	FAD	C5A-N7A	-2.16	1.32	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1001	FAD	N3A-C2A-N1A	-9.01	121.15	128.86
2	B	1001	FAD	N3A-C2A-N1A	-8.67	121.44	128.86
2	A	1001	FAD	C4B-O4B-C1B	-4.48	105.16	109.83
2	B	1001	FAD	C4X-C10-N10	-4.01	116.67	120.40
2	A	1001	FAD	C4X-C10-N10	-3.66	117.00	120.40

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1002	ACT	6	0
2	B	1001	FAD	3	0
3	B	1002	ACT	1	0
4	B	1004	EDO	2	0

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

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	425/447 (95%)	0.07	9 (2%) 63 62	16, 35, 67, 93	0
1	B	425/447 (95%)	0.08	13 (3%) 49 48	16, 35, 68, 99	0
All	All	850/894 (95%)	0.07	22 (2%) 56 55	16, 35, 68, 99	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	269	ARG	3.5
1	A	131	ARG	3.4
1	A	98	ARG	3.4
1	A	97	TRP	3.0
1	A	425	ARG	2.9

### 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. 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
3	ACT	B	1002	4/4	0.71	0.18	26,30,34,49	0
3	ACT	A	1002	4/4	0.76	0.28	20,25,36,52	0
4	EDO	A	1003	4/4	0.93	0.12	18,21,22,22	0
4	EDO	A	1004	4/4	0.93	0.13	21,27,27,32	0
4	EDO	B	1003	4/4	0.93	0.09	18,20,21,22	0
2	FAD	B	1001	53/53	0.95	0.11	16,20,31,41	0
2	FAD	A	1001	53/53	0.97	0.11	15,20,32,42	0
4	EDO	B	1004	4/4	0.98	0.10	20,29,37,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.