



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:09 AM GMT

PDB ID : 2PNJ  
Title : Crystal structure of human ferrochelatase mutant with Phe 337 replaced by Ala  
Authors : Dailey, H.A.; Wu, C.-K.; Horanyi, P.; Medlock, A.E.; Najahi-Missaoui, W.; Burden, A.E.; Dailey, T.A.; Rose, J.P.  
Deposited on : 2007-04-24  
Resolution : 2.35 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

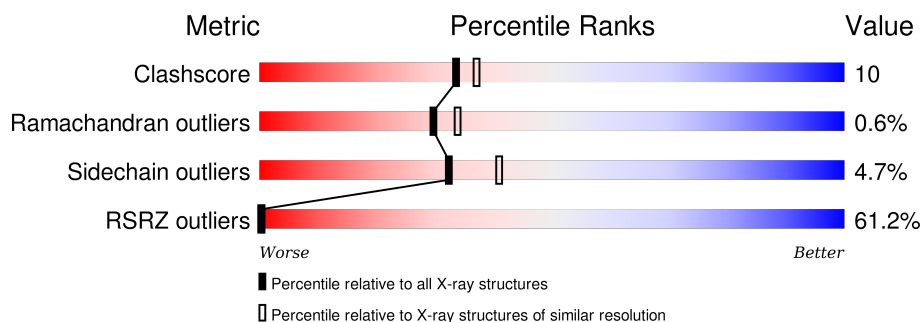
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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	359	<div> <div>36%</div> <div>79%</div> <div>16%</div> <div>..</div> </div>
1	B	359	<div> <div>85%</div> <div>78%</div> <div>18%</div> <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	CHD	B	501	-	-	-	X
3	CHD	B	502	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CHD	B	503	-	-	-	X

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6166 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.

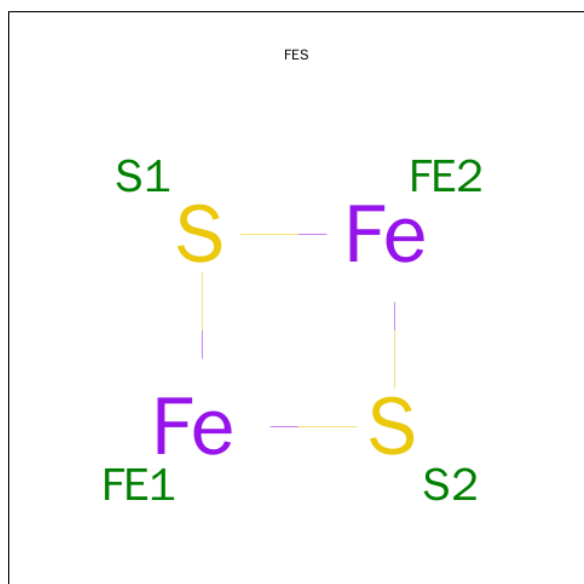
- Molecule 1 is a protein called Ferrochelatase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	355	Total	C	N	O	S	0	0	0
			2815	1794	481	522	18			
1	B	356	Total	C	N	O	S	0	0	0
			2827	1803	484	522	18			

There are 4 discrepancies between the modelled and reference sequences:

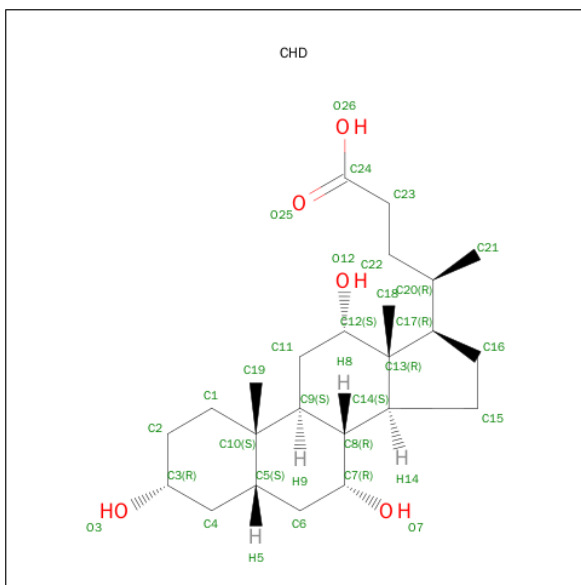
Chain	Residue	Modelled	Actual	Comment	Reference
A	115	LEU	ARG	ENGINEERED	UNP P22830
A	337	ALA	PHE	ENGINEERED	UNP P22830
B	115	LEU	ARG	ENGINEERED	UNP P22830
B	337	ALA	PHE	ENGINEERED	UNP P22830

- Molecule 2 is FE2/S2 (INORGANIC) CLUSTER (three-letter code: FES) (formula: Fe<sub>2</sub>S<sub>2</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe S 4 2 2	0	0
2	B	1	Total Fe S 4 2 2	0	0

- Molecule 3 is CHOLIC ACID (three-letter code: CHD) (formula:  $C_{24}H_{40}O_5$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 29 24 5	0	0
3	A	1	Total C O 29 24 5	0	0
3	A	1	Total C O 29 24 5	0	0
3	B	1	Total C O 29 24 5	0	0
3	B	1	Total C O 29 24 5	0	0
3	B	1	Total C O 29 24 5	0	0

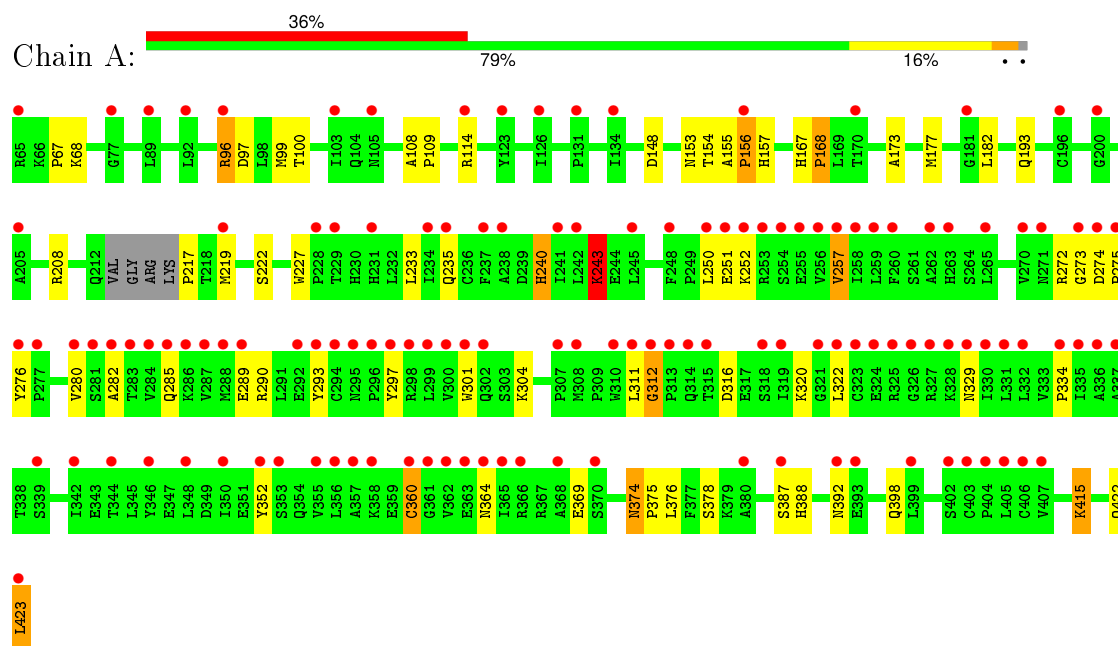
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	172	Total O 172 172	0	0
4	B	170	Total O 170 170	0	0

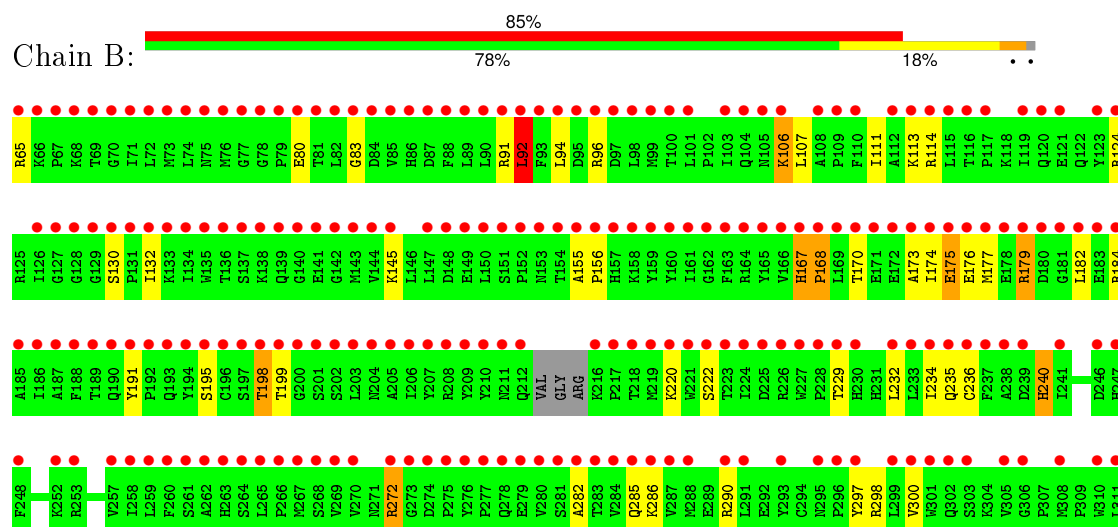
### 3 Residue-property plots

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: Ferrochelatase, mitochondrial



#### • Molecule 1: Ferrochelatase, mitochondrial



F377	F378	F379	A380	A381	A382	D383	D384	F385	F386	F387	F388	F389	F390	F391	F392	F393	F394	F395	F396	F397	Q398	F399	T400	L401	S402	C403	P404	L405	C406	F407	N408	P409	V410	C411	R412	E413	T414	K415	S416	F417	F418	T419	S420	Q421	Q422	L423	F377	F378	F379	A380	A381	A382	D383	D384	F385	F386	F387	F388	F389	F390	F391	F392	F393	F394	F395	F396	F397	Q398	F399	T400	L401	S402	C403	P404	L405	C406	F407	N408	P409	V410	C411	R412	E413	T414	K415	S416	F417	F418	T419	S420	Q421	Q422	L423	F377	F378	F379	A380	A381	A382	D383	D384	F385	F386	F387	F388	F389	F390	F391	F392	F393	F394	F395	F396	F397	Q398	F399	T400	L401	S402	C403	P404	L405	C406	F407	N408	P409	V410	C411	R412	E413	T414	K415	S416	F417	F418	T419	S420	Q421	Q422	L423
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.69Å 94.05Å 113.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.44 – 2.35 48.52 – 1.95	Depositor EDS
% Data completeness (in resolution range)	96.4 (43.44-2.35) 80.5 (48.52-1.95)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.24 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.2	Depositor
R, $R_{free}$	0.192 , 0.253 0.401 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	20.6	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 30.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	5 of 58876 reflections (0.008%)	Xtriage
$F_o, F_c$ correlation	0.75	EDS
Total number of atoms	6166	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	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 71.66 % 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 2.5280e-06. 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: CHD, FES

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	0.76	1/2883 (0.0%)	0.78	2/3916 (0.1%)
1	B	0.77	0/2895	0.81	4/3931 (0.1%)
All	All	0.76	1/5778 (0.0%)	0.80	6/7847 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	B	0	5
All	All	0	9

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	312	GLY	CA-C	-5.00	1.43	1.51

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	167	HIS	C-N-CD	-9.73	99.19	120.60
1	B	92	LEU	CA-CB-CG	7.40	132.33	115.30
1	A	312	GLY	N-CA-C	-6.72	96.29	113.10
1	B	91	ARG	NE-CZ-NH2	-5.85	117.37	120.30
1	A	243	LYS	CD-CE-NZ	5.49	124.32	111.70

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	155	ALA	Peptide
1	A	167	HIS	Mainchain,Peptide
1	A	311	LEU	Peptide
1	B	155	ALA	Mainchain,Peptide
1	B	167	HIS	Peptide

## 5.2 Too-close contacts ⓘ

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	2815	0	2769	56	8
1	B	2827	0	2790	64	5
2	A	4	0	0	0	0
2	B	4	0	0	0	0
3	A	87	0	117	2	0
3	B	87	0	117	4	0
4	A	172	0	0	13	2
4	B	170	0	0	19	7
All	All	6166	0	5793	118	11

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 10.

The worst 5 of 118 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:243:LYS:HD2	4:A:616:HOH:O	1.28	1.31
1:B:410:VAL:HB	4:B:633:HOH:O	1.43	1.18
1:B:323:CYS:HB3	4:B:602:HOH:O	1.47	1.13
1:A:398:GLN:HE22	1:B:297:TYR:H	1.15	0.94
1:B:272:ARG:NH1	4:B:673:HOH:O	2.01	0.93

The worst 5 of 11 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:645:HOH:O	4:B:628:HOH:O[3_655]	0.89	1.31
1:B:80:GLU:OE1	4:A:620:HOH:O[3_645]	1.08	1.12
1:A:252:LYS:CE	4:B:599:HOH:O[3_655]	1.09	1.11
1:A:252:LYS:CD	4:B:599:HOH:O[3_655]	1.66	0.54
1:A:251:GLU:CB	1:B:83:GLY:O[3_655]	1.75	0.45

## 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	351/359 (98%)	338 (96%)	11 (3%)	2 (1%)	30	34
1	B	352/359 (98%)	340 (97%)	10 (3%)	2 (1%)	30	34
All	All	703/718 (98%)	678 (96%)	21 (3%)	4 (1%)	30	34

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	156	PRO
1	B	168	PRO
1	A	168	PRO
1	B	156	PRO

### 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	310/323 (96%)	296 (96%)	14 (4%)	34	43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	311/323 (96%)	296 (95%)	15 (5%)	31	40
All	All	621/646 (96%)	592 (95%)	29 (5%)	32	41

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	423	LEU
1	B	114	ARG
1	B	362	VAL
1	B	92	LEU
1	B	175	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	390	GLN
1	A	421	GLN
1	B	398	GLN
1	A	392	ASN
1	A	398	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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	FES	A	499	1	0,4,4	0.00	-	0,4,4	0.00	-
3	CHD	A	501	-	29,32,32	0.57	0	48,51,51	1.80	14 (29%)
3	CHD	A	502	-	29,32,32	0.52	0	48,51,51	1.83	15 (31%)
3	CHD	A	503	-	29,32,32	0.54	0	48,51,51	1.42	8 (16%)
2	FES	B	499	1	0,4,4	0.00	-	0,4,4	0.00	-
3	CHD	B	501	-	29,32,32	0.68	0	48,51,51	2.07	11 (22%)
3	CHD	B	502	-	29,32,32	0.44	0	48,51,51	1.81	10 (20%)
3	CHD	B	503	-	29,32,32	0.51	0	48,51,51	1.50	10 (20%)

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	FES	A	499	1	-	0/0/4/4	0/1/1/1
3	CHD	A	501	-	-	0/7/74/74	0/4/4/4
3	CHD	A	502	-	-	0/7/74/74	0/4/4/4
3	CHD	A	503	-	-	0/7/74/74	0/4/4/4
2	FES	B	499	1	-	0/0/4/4	0/1/1/1
3	CHD	B	501	-	-	0/7/74/74	0/4/4/4
3	CHD	B	502	-	-	0/7/74/74	0/4/4/4
3	CHD	B	503	-	-	0/7/74/74	0/4/4/4

There are no bond length outliers.

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	501	CHD	C23-C22-C20	-7.18	106.29	114.75
3	B	501	CHD	C13-C17-C20	-4.61	113.89	119.50
3	A	501	CHD	C6-C5-C4	-4.60	105.91	111.05
3	A	502	CHD	C6-C5-C4	-4.47	106.05	111.05
3	B	501	CHD	C6-C5-C4	-4.44	106.08	111.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	CHD	1	0
3	A	502	CHD	1	0
3	B	501	CHD	3	0
3	B	503	CHD	1	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	355/359 (98%)	1.86	131 (36%) 0 0	14, 23, 40, 50	0
1	B	356/359 (99%)	4.05	304 (85%) 0 0	13, 23, 41, 48	0
All	All	711/718 (99%)	2.95	435 (61%) 0 0	13, 23, 41, 50	0

The worst 5 of 435 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	161	ILE	11.4
1	B	169	LEU	10.9
1	B	206	ILE	10.6
1	B	410	VAL	9.9
1	B	83	GLY	9.7

### 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	CHD	A	501	29/29	0.74	0.28	1.08	26,31,47,48	0
3	CHD	B	503	29/29	0.59	0.45	1.02	45,48,63,64	0
3	CHD	B	501	29/29	0.64	0.42	0.89	28,29,41,45	0
3	CHD	A	502	29/29	0.66	0.26	0.88	37,41,49,51	0
3	CHD	B	502	29/29	0.51	0.44	0.65	34,38,51,53	0
3	CHD	A	503	29/29	0.54	0.31	-0.09	43,47,50,51	0
2	FES	B	499	4/4	0.51	0.25	-2.14	20,20,22,23	0
2	FES	A	499	4/4	0.88	0.12	-3.50	16,17,18,18	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.