



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 25, 2018 – 01:28 PM EDT

PDB ID : 1Q0C  
Title : Anerobic Substrate Complex of Homoproteocatechuate 2,3-Dioxygenase from *Brevibacterium fuscum*. (Complex with 3,4-Dihydroxyphenylacetate)  
Authors : Vetting, M.W.; Wackett, L.P.; Que, L.; Lipscomb, J.D.; Ohlendorf, D.H.  
Deposited on : 2003-07-15  
Resolution : 2.10 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 : rb-20031021  
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) : rb-20031021

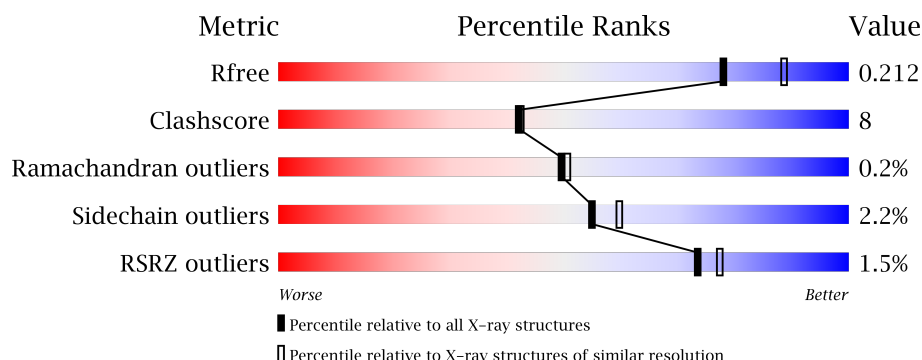
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	4608 (2.10-2.10)
Clashscore	122126	5109 (2.10-2.10)
Ramachandran outliers	120053	5059 (2.10-2.10)
Sidechain outliers	120020	5060 (2.10-2.10)
RSRZ outliers	108989	4497 (2.10-2.10)

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	365	<div> <div>0.2%</div> <div> <div></div> <div>71%</div> <div>15%</div> <div>•</div> <div>13%</div> </div> </div>
1	B	365	<div> <div>3%</div> <div> <div></div> <div>70%</div> <div>16%</div> <div>•</div> <div>13%</div> </div> </div>
1	C	365	<div> <div></div> <div> <div></div> <div>71%</div> <div>15%</div> <div>••</div> <div>13%</div> </div> </div>
1	D	365	<div> <div>0.2%</div> <div> <div></div> <div>75%</div> <div>12%</div> <div>•</div> <div>13%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10829 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.

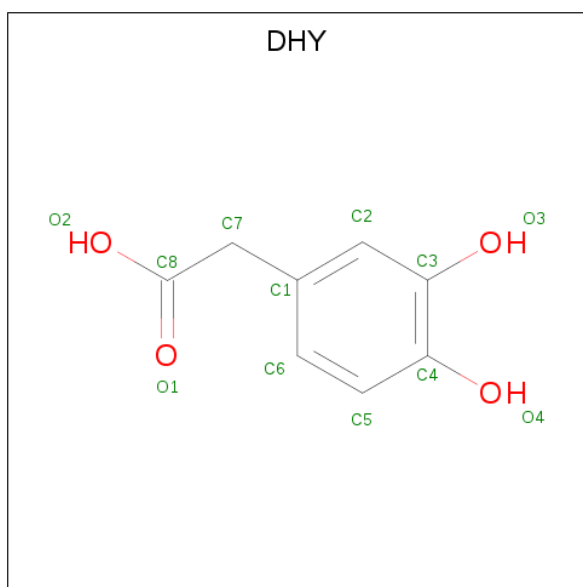
- Molecule 1 is a protein called homoprotocatechuate 2,3-dioxygenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2601	1648	459	487	7			
1	B	319	Total	C	N	O	S	0	0	0
			2601	1648	459	487	7			
1	C	319	Total	C	N	O	S	0	0	0
			2601	1648	459	487	7			
1	D	319	Total	C	N	O	S	0	0	0
			2601	1648	459	487	7			

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Fe	0	0
			1	1		
2	A	1	Total	Fe	0	0
			1	1		
2	D	1	Total	Fe	0	0
			1	1		
2	C	1	Total	Fe	0	0
			1	1		

- Molecule 3 is 2-(3,4-DIHYDROXYPHENYL)ACETIC ACID (three-letter code: DHY) (formula: C<sub>8</sub>H<sub>8</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			12	8	4		
3	B	1	Total	C	O	0	0
			12	8	4		
3	C	1	Total	C	O	0	0
			12	8	4		
3	D	1	Total	C	O	0	0
			12	8	4		

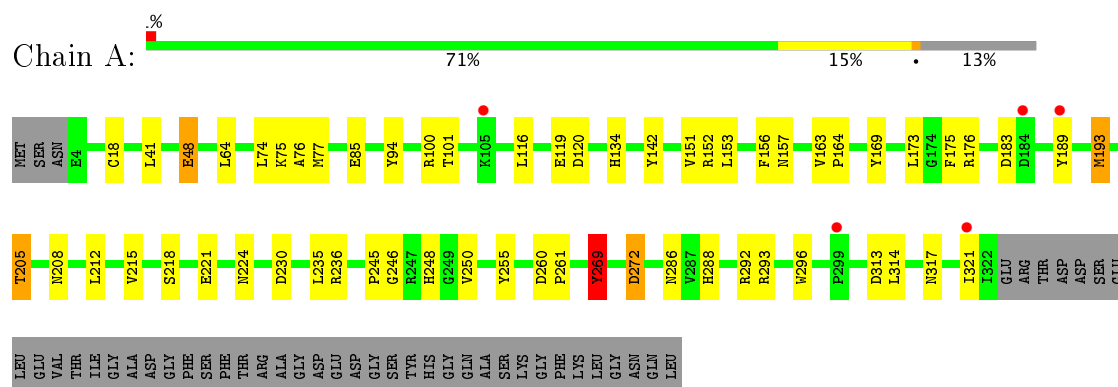
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	84	Total	O	0	0
			84	84		
4	B	67	Total	O	0	0
			67	67		
4	C	107	Total	O	0	0
			107	107		
4	D	115	Total	O	0	0
			115	115		

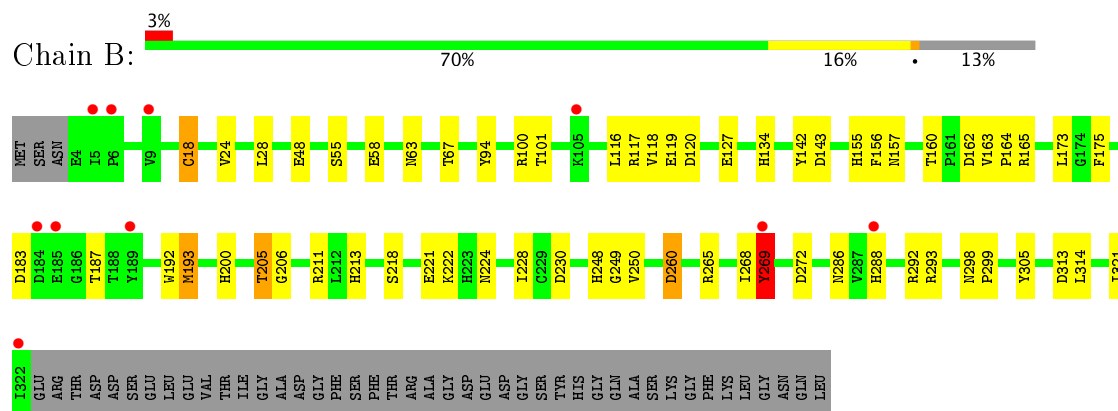
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

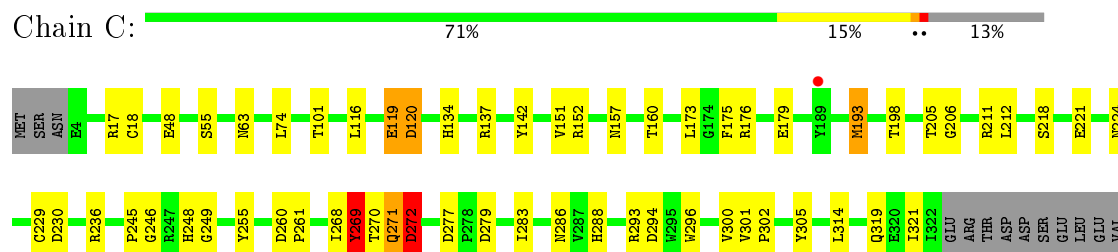
- Molecule 1: homoprotocatechuate 2,3-dioxygenase



- Molecule 1: homoprotocatechuate 2,3-dioxygenase



- Molecule 1: homoprotocatechuate 2,3-dioxygenase



THR	ILE	GLY	ALA	ASP	GLY	PHE	SER	PHE	THR	ARG	ALA	GLY	ASP	GLU	ASP	GLY	SER	TYR	HIS	GLY	GLN	ALA	SER	LYS	GLY	PHE	LYS	LEU	GLY	ASN	GLN	LEU
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

• Molecule 1: homoprotocatechuate 2,3-dioxygenase



MET	SER	ASN	E4	V9	R32	Y45	E46	D47	E48	I51	S55	N63	Y94	K110	D120	H134	R137	Y142	V151	R152	N157	D162	Y163	P164	Y169	L173	G174	F175	M193	H194	R195	T205	L212	V215	S218	E221
-----	-----	-----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

N224	I225	D230	G246	E247	E248	D262	N263	E267	I268	Y269	T270	Q271	D272	N286	V287	E288	D289	N290	W296	G297	N298	P299	V300	I322	GLU	ARG	THR	ASP	ASP	SER	GLU	LEU	LEU	VAL	THR	THR	ILE	GLY	ALA	ASP	GLY	PHE	SER	PHE	THR	ARG	ALA	GLY	ASP	GLU	ASP	GLY
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

SER	THR	HIS	GLY	GLN	ALA	SER	LYS	PHE	LYS	LEU	GLY	ASN	GLN	LEU
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	157.00Å 157.00Å 122.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10 41.02 – 2.11	Depositor EDS
% Data completeness (in resolution range)	96.0 (20.00-2.10) 97.4 (41.02-2.11)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	0.05	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.89 (at 2.10Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.170 , 0.234 0.170 , 0.212	Depositor DCC
$R_{free}$ test set	4152 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	36.9	Xtriage
Anisotropy	0.066	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 44.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.004 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	10829	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 2.85% 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: FE, DHY

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	0.73	1/2672 (0.0%)	0.90	4/3634 (0.1%)
1	B	0.73	2/2672 (0.1%)	0.93	7/3634 (0.2%)
1	C	0.84	2/2672 (0.1%)	0.97	10/3634 (0.3%)
1	D	0.81	0/2672	0.93	5/3634 (0.1%)
All	All	0.78	5/10688 (0.0%)	0.93	26/14536 (0.2%)

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	1
1	C	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	18	CYS	CB-SG	-6.18	1.71	1.82
1	B	269	TYR	CE1-CZ	5.78	1.46	1.38
1	B	18	CYS	CB-SG	-5.52	1.72	1.81
1	C	229	CYS	CB-SG	-5.18	1.73	1.81
1	A	269	TYR	CE1-CZ	5.04	1.45	1.38

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	269	TYR	CB-CG-CD2	-7.00	116.80	121.00
1	D	269	TYR	CA-CB-CG	6.50	125.75	113.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	269	TYR	N-CA-CB	-6.37	99.13	110.60
1	D	272	ASP	CB-CG-OD1	-6.31	112.62	118.30
1	A	272	ASP	CB-CG-OD1	-6.18	112.74	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	269	TYR	Sidechain
1	C	269	TYR	Sidechain

## 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	2601	0	2485	38	0
1	B	2601	0	2485	45	0
1	C	2601	0	2485	44	0
1	D	2601	0	2485	34	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	12	0	5	0	0
3	B	12	0	5	1	0
3	C	12	0	5	2	0
3	D	12	0	5	1	0
4	A	84	0	0	2	0
4	B	67	0	0	1	0
4	C	107	0	0	1	0
4	D	115	0	0	2	0
All	All	10829	0	9960	154	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 8.

The worst 5 of 154 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:C:271:GLN:HG2	1:C:271:GLN:O	1.73	0.88
1:D:271:GLN:HG2	1:D:271:GLN:O	1.74	0.86
1:C:286:ASN:ND2	1:C:288:HIS:H	1.73	0.85
1:D:286:ASN:ND2	1:D:288:HIS:H	1.78	0.81
1:D:48:GLU:H	1:D:48:GLU:CD	1.84	0.80

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	317/365 (87%)	305 (96%)	12 (4%)	0	100	100
1	B	317/365 (87%)	303 (96%)	14 (4%)	0	100	100
1	C	317/365 (87%)	305 (96%)	11 (4%)	1 (0%)	43	43
1	D	317/365 (87%)	307 (97%)	9 (3%)	1 (0%)	43	43
All	All	1268/1460 (87%)	1220 (96%)	46 (4%)	2 (0%)	49	51

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	271	GLN
1	D	271	GLN

### 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	277/313 (88%)	270 (98%)	7 (2%)	50	55
1	B	277/313 (88%)	270 (98%)	7 (2%)	50	55
1	C	277/313 (88%)	272 (98%)	5 (2%)	62	67
1	D	277/313 (88%)	272 (98%)	5 (2%)	62	67
All	All	1108/1252 (88%)	1084 (98%)	24 (2%)	55	60

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	205	THR
1	B	272	ASP
1	D	272	ASP
1	B	260	ASP
1	B	269	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	286	ASN
1	C	95	GLN
1	D	224	ASN
1	C	63	ASN
1	C	134	HIS

### 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

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 for Mogul analysis.

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
3	DHY	A	999	2	9,12,12	2.19	6 (66%)	13,16,16	2.09	5 (38%)
3	DHY	B	999	2	9,12,12	2.26	5 (55%)	13,16,16	2.39	6 (46%)
3	DHY	C	999	2	9,12,12	2.57	6 (66%)	13,16,16	2.02	5 (38%)
3	DHY	D	999	2	9,12,12	1.87	5 (55%)	13,16,16	2.06	5 (38%)

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
3	DHY	A	999	2	-	0/2/4/4	0/1/1/1
3	DHY	B	999	2	-	0/2/4/4	0/1/1/1
3	DHY	C	999	2	-	0/2/4/4	0/1/1/1
3	DHY	D	999	2	-	0/2/4/4	0/1/1/1

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	999	DHY	C2-C1	2.09	1.42	1.39
3	D	999	DHY	C6-C1	2.10	1.43	1.38
3	A	999	DHY	C5-C4	2.15	1.43	1.39
3	D	999	DHY	O4-C4	2.23	1.41	1.36
3	B	999	DHY	O4-C4	2.27	1.41	1.36

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	999	DHY	C7-C1-C2	-3.83	113.00	120.69
3	B	999	DHY	O4-C4-C3	-3.35	109.78	118.50

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	999	DHY	O3-C3-C4	-3.18	110.23	118.50
3	D	999	DHY	O3-C3-C4	-3.05	110.56	118.50
3	C	999	DHY	C7-C1-C2	-3.04	114.59	120.69

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	999	DHY	1	0
3	C	999	DHY	2	0
3	D	999	DHY	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	319/365 (87%)	-0.22	5 (1%) 72 76	28, 41, 59, 76	0
1	B	319/365 (87%)	-0.17	10 (3%) 49 56	27, 40, 60, 75	0
1	C	319/365 (87%)	-0.40	1 (0%) 93 94	23, 32, 49, 62	0
1	D	319/365 (87%)	-0.30	3 (0%) 84 86	25, 33, 53, 69	0
All	All	1276/1460 (87%)	-0.27	19 (1%) 73 77	23, 37, 57, 76	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	5	ILE	3.4
1	B	184	ASP	3.3
1	A	321	ILE	3.3
1	B	6	PRO	2.9
1	C	189	TYR	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	DHY	C	999	12/12	0.88	0.11	43,45,46,47	0
3	DHY	B	999	12/12	0.89	0.12	40,43,51,52	0
3	DHY	A	999	12/12	0.94	0.11	45,46,50,50	0
3	DHY	D	999	12/12	0.95	0.11	42,44,49,50	0
2	FE	A	500	1/1	0.98	0.05	43,43,43,43	0
2	FE	B	500	1/1	0.99	0.06	42,42,42,42	0
2	FE	C	500	1/1	0.99	0.05	34,34,34,34	0
2	FE	D	500	1/1	1.00	0.06	37,37,37,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.