



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 05:40 PM GMT

PDB ID : 3C8D  
Title : Crystal structure of the enterobactin esterase FES from *Shigella flexneri* in the presence of 2,3-Di-hydroxy-N-benzoyl-glycine  
Authors : Kim, Y.; Maltseva, N.; Abergel, R.; Holzle, D.; Raymond, K.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2008-02-11  
Resolution : 1.80 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

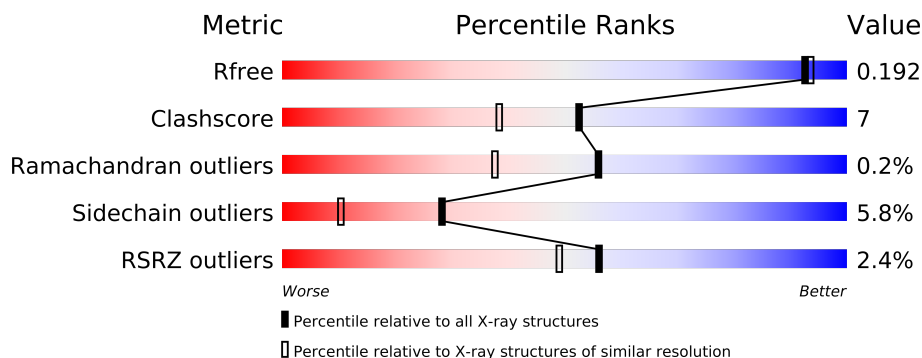
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 1.80 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))
$R_{free}$	66092	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	403	
1	B	403	
1	C	403	
1	D	403	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	CIT	A	501	-	X
2	CIT	B	501	-	X
2	CIT	C	501	-	X
2	CIT	C	502	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13727 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Enterochelin esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	378	Total	C	N	O	S	0	11	0
			3146	2019	553	562	12			
1	B	379	Total	C	N	O	S	0	9	0
			3134	2013	548	561	12			
1	C	379	Total	C	N	O	S	0	8	0
			3118	2003	544	559	12			
1	D	378	Total	C	N	O	S	0	11	0
			3157	2029	552	564	12			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q83SB9
A	-1	ASN	-	EXPRESSION TAG	UNP Q83SB9
A	0	ALA	-	EXPRESSION TAG	UNP Q83SB9
B	-2	SER	-	EXPRESSION TAG	UNP Q83SB9
B	-1	ASN	-	EXPRESSION TAG	UNP Q83SB9
B	0	ALA	-	EXPRESSION TAG	UNP Q83SB9
C	-2	SER	-	EXPRESSION TAG	UNP Q83SB9
C	-1	ASN	-	EXPRESSION TAG	UNP Q83SB9
C	0	ALA	-	EXPRESSION TAG	UNP Q83SB9
D	-2	SER	-	EXPRESSION TAG	UNP Q83SB9
D	-1	ASN	-	EXPRESSION TAG	UNP Q83SB9
D	0	ALA	-	EXPRESSION TAG	UNP Q83SB9

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).

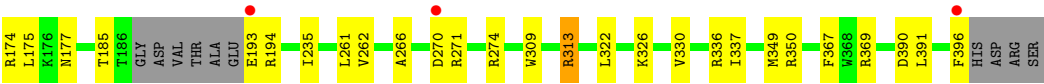


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	247	Total	O	0	0
			247	247		
3	B	314	Total	O	0	0
			314	314		
3	C	270	Total	O	0	0
			270	270		
3	D	289	Total	O	0	0
			289	289		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.30Å 155.97Å 48.48Å 90.00° 90.01° 90.00°	Depositor
Resolution (Å)	48.48 – 1.80 48.48 – 1.80	Depositor EDS
% Data completeness (in resolution range)	90.0 (48.48-1.80) 98.0 (48.48-1.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 1.79Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.150 , 0.194 0.150 , 0.192	Depositor DCC
$R_{free}$ test set	14896 reflections (9.97%)	DCC
Wilson B-factor (Å <sup>2</sup> )	26.6	Xtriage
Anisotropy	0.527	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 41.4	EDS
Estimated twinning fraction	0.368 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 149411 reflections	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	13727	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	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 41.79 % 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 2.2427e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CIT

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	0.34	0/3244	0.51	0/4425
1	B	0.35	0/3233	0.55	2/4409 (0.0%)
1	C	0.33	0/3217	0.53	1/4390 (0.0%)
1	D	0.34	0/3255	0.53	1/4438 (0.0%)
All	All	0.34	0/12949	0.53	4/17662 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	193	GLU	N-CA-C	-7.79	89.97	111.00
1	D	193	GLU	N-CA-C	-5.94	94.96	111.00
1	B	187	GLY	N-CA-C	5.82	127.64	113.10
1	C	193	GLU	N-CA-C	-5.43	96.33	111.00

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3146	0	0	22	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	3134	0	0	17	0
1	C	3118	0	0	25	0
1	D	3157	0	0	27	0
2	A	13	0	5	1	0
2	B	13	0	5	0	0
2	C	26	0	10	6	0
3	A	247	0	0	5	0
3	B	314	0	0	6	0
3	C	270	0	0	6	0
3	D	289	0	0	13	0
All	All	13727	0	20	90	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

The worst 5 of 90 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:501:CIT:H42	3:B:587:HOH:O	1.66	0.95
1:A:394:PRO:O	1:A:397:HIS:ND1	2.11	0.83
1:A:397:HIS:O	1:A:397:HIS:CD2	2.38	0.76
1:A:137:LEU:CB	1:B:317:GLN:CD	2.63	0.67
1:C:292:LEU:O	3:C:967:HOH:O	2.12	0.66

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	381/403 (94%)	367 (96%)	13 (3%)	1 (0%)	50	31
1	B	380/403 (94%)	368 (97%)	11 (3%)	1 (0%)	50	31
1	C	381/403 (94%)	362 (95%)	19 (5%)	0	100	100
1	D	379/403 (94%)	369 (97%)	9 (2%)	1 (0%)	50	31

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1521/1612 (94%)	1466 (96%)	52 (3%)	3 (0%)	56 38

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	138	GLY
1	D	235	ILE
1	B	235	ILE

### 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	331/341 (97%)	316 (96%)	15 (4%)	38 17
1	B	331/341 (97%)	307 (93%)	24 (7%)	20 6
1	C	328/341 (96%)	308 (94%)	20 (6%)	26 9
1	D	332/341 (97%)	315 (95%)	17 (5%)	33 14
All	All	1322/1364 (97%)	1246 (94%)	76 (6%)	28 11

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	333	GLU
1	C	57	THR
1	D	313	ARG
1	B	362	ILE
1	B	397	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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	CIT	A	501	-	12,12,12	1.04	0	17,17,17	1.91	5 (29%)
2	CIT	B	501	-	12,12,12	1.96	4 (33%)	17,17,17	1.92	6 (35%)
2	CIT	C	501	-	12,12,12	1.69	4 (33%)	17,17,17	2.78	7 (41%)
2	CIT	C	502	-	12,12,12	1.70	3 (25%)	17,17,17	2.60	8 (47%)

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	CIT	A	501	-	-	0/16/16/16	0/0/0/0
2	CIT	B	501	-	-	0/16/16/16	0/0/0/0
2	CIT	C	501	-	-	0/16/16/16	0/0/0/0
2	CIT	C	502	-	-	1/16/16/16	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	501	CIT	C3-C6	-4.56	1.48	1.53
2	C	501	CIT	O2-C1	-3.27	1.18	1.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	502	CIT	O7-C3	2.99	1.49	1.43
2	B	501	CIT	C2-C3	-2.88	1.49	1.53
2	C	501	CIT	C3-C6	-2.70	1.50	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	501	CIT	O2-C1-O1	-7.51	104.20	123.30
2	C	502	CIT	O6-C6-C3	5.71	121.20	112.89
2	C	502	CIT	C2-C3-C6	-4.83	98.93	110.12
2	C	501	CIT	O6-C6-C3	4.59	119.56	112.89
2	A	501	CIT	O6-C6-C3	4.22	119.03	112.89

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	502	CIT	O7-C3-C6-O6

There are no ring outliers.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	378/403 (93%)	0.03	10 (2%) 53 46	18, 35, 73, 167	0
1	B	379/403 (94%)	-0.00	13 (3%) 43 35	15, 30, 72, 114	0
1	C	379/403 (94%)	-0.03	7 (1%) 65 60	18, 34, 71, 102	0
1	D	378/403 (93%)	0.01	7 (1%) 64 58	18, 36, 71, 98	0
All	All	1514/1612 (93%)	0.00	37 (2%) 56 49	15, 34, 71, 167	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	137	LEU	14.9
1	B	398	ASP	6.0
1	B	105	PRO	6.0
1	D	137	LEU	5.2
1	A	136	GLY	4.9

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	CIT	C	501	13/13	0.23	9.00	64,68,71,75	0
2	CIT	B	501	13/13	0.15	8.23	87,92,94,96	0
2	CIT	C	502	13/13	0.21	7.33	83,88,93,94	0
2	CIT	A	501	13/13	0.15	3.44	38,54,72,76	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.