



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 12, 2021 – 11:26 AM EDT

PDB ID : 1YKN  
Title : Protocatechuate 3,4-dioxygenase Y408E mutant bound to DHB  
Authors : Brown, C.K.; Ohlendorf, D.H.  
Deposited on : 2005-01-18  
Resolution : 2.06 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.23.2  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.23.2

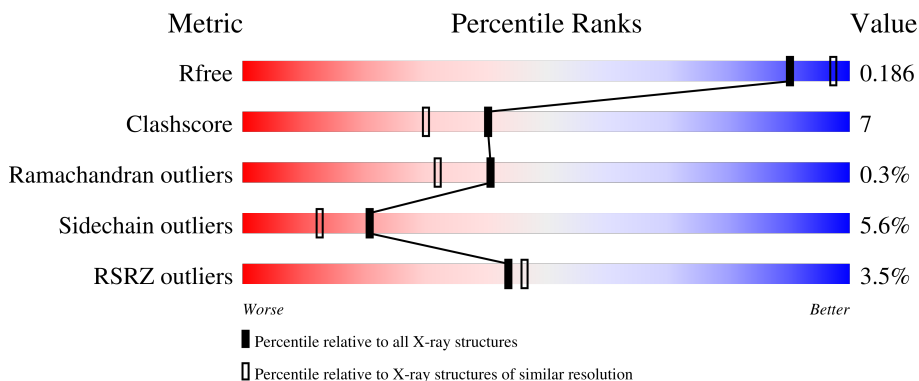
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.06 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))
$R_{free}$	130704	2684 (2.08-2.04)
Clashscore	141614	2801 (2.08-2.04)
Ramachandran outliers	138981	2768 (2.08-2.04)
Sidechain outliers	138945	2768 (2.08-2.04)
RSRZ outliers	127900	2646 (2.08-2.04)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	200	
1	C	200	
1	E	200	
1	G	200	
1	I	200	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	200	
2	B	238	
2	D	238	
2	F	238	
2	H	238	
2	J	238	
2	L	238	

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
2	CME	H	429	-	X	-	-
4	DHB	B	550	-	-	X	-
4	DHB	F	2550	-	-	X	-
4	DHB	H	3550	-	-	X	-
4	DHB	L	5550	-	-	X	-

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 22266 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 Protocatechuate 3,4-dioxygenase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	200	1571	993	276	299	3	0	0	0
1	C	200	1571	993	276	299	3	0	0	0
1	E	200	1571	993	276	299	3	0	0	0
1	G	200	1571	993	276	299	3	0	0	0
1	I	200	1571	993	276	299	3	0	0	0
1	K	200	1571	993	276	299	3	0	0	0

- Molecule 2 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	238	1876	1186	342	339	9	0	0	0
2	D	238	1876	1186	342	339	9	0	0	0
2	F	238	1876	1186	342	339	9	0	0	0
2	H	238	1876	1186	342	339	9	0	0	0
2	J	238	1876	1186	342	339	9	0	0	0
2	L	238	1876	1186	342	339	9	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	408	GLU	TYR	engineered mutation	UNP P00437

*Continued on next page...*

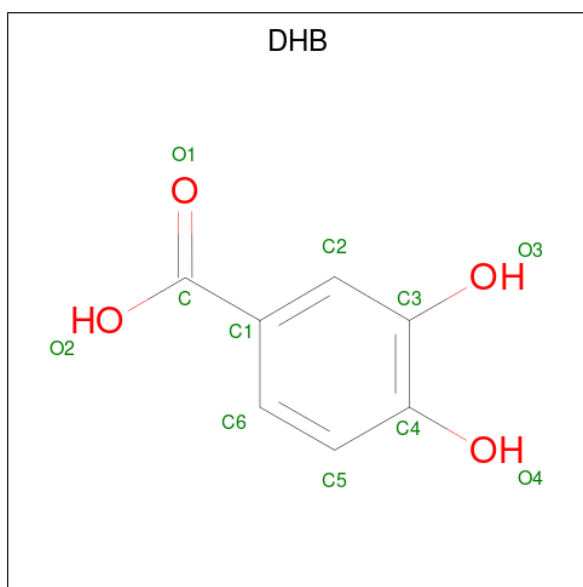
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	429	CME	CYS	modified residue	UNP P00437
D	408	GLU	TYR	engineered mutation	UNP P00437
D	429	CME	CYS	modified residue	UNP P00437
F	408	GLU	TYR	engineered mutation	UNP P00437
F	429	CME	CYS	modified residue	UNP P00437
H	408	GLU	TYR	engineered mutation	UNP P00437
H	429	CME	CYS	modified residue	UNP P00437
J	408	GLU	TYR	engineered mutation	UNP P00437
J	429	CME	CYS	modified residue	UNP P00437
L	408	GLU	TYR	engineered mutation	UNP P00437
L	429	CME	CYS	modified residue	UNP P00437

- Molecule 3 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Fe 1 1	0	0
3	D	1	Total Fe 1 1	0	0
3	F	1	Total Fe 1 1	0	0
3	H	1	Total Fe 1 1	0	0
3	J	1	Total Fe 1 1	0	0
3	L	1	Total Fe 1 1	0	0

- Molecule 4 is 3,4-DIHYDROXYBENZOIC ACID (three-letter code: DHB) (formula: C<sub>7</sub>H<sub>6</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	C	O	0	0
			11	7	4		
4	D	1	Total	C	O	0	0
			11	7	4		
4	F	1	Total	C	O	0	0
			11	7	4		
4	H	1	Total	C	O	0	0
			11	7	4		
4	J	1	Total	C	O	0	0
			11	7	4		
4	L	1	Total	C	O	0	0
			11	7	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	88	Total	O	0	0
			88	88		
5	B	162	Total	O	0	0
			162	162		
5	C	88	Total	O	0	0
			88	88		
5	D	168	Total	O	0	0
			168	168		
5	E	86	Total	O	0	0
			86	86		
5	F	163	Total	O	0	0
			163	163		

*Continued on next page...*

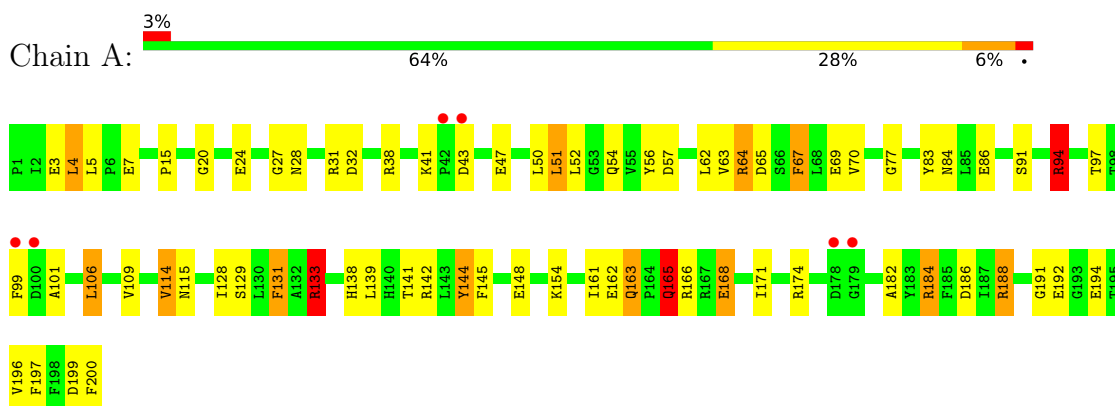
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
5	G	92	Total 92	O 92	0	0
5	H	157	Total 157	O 157	0	0
5	I	87	Total 87	O 87	0	0
5	J	172	Total 172	O 172	0	0
5	K	81	Total 81	O 81	0	0
5	L	168	Total 168	O 168	0	0

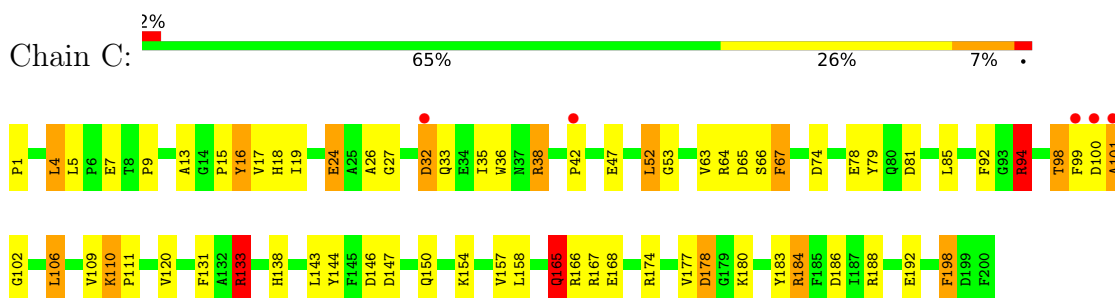
### 3 Residue-property plots [i](#)

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

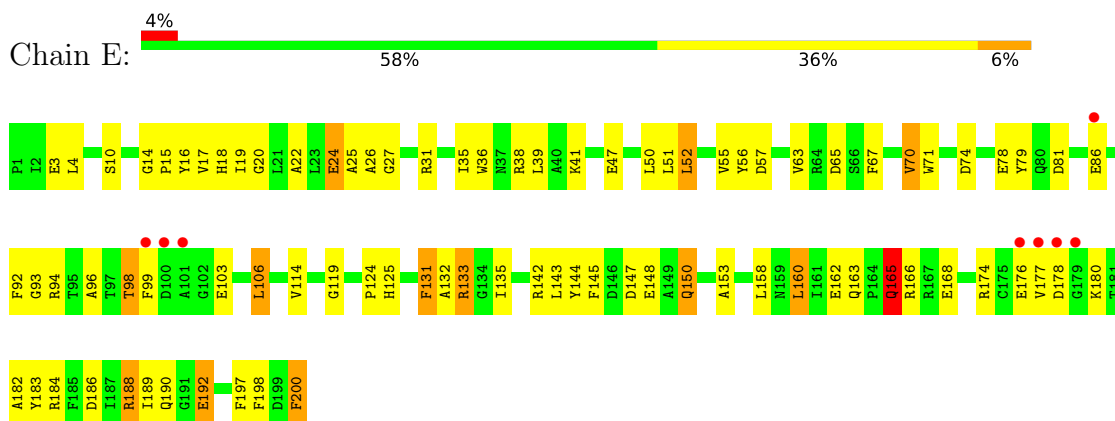
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain

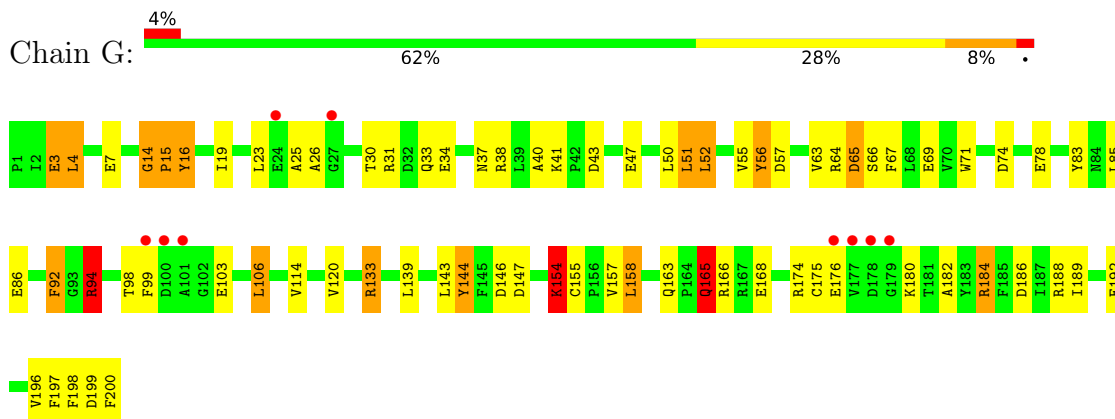


- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain

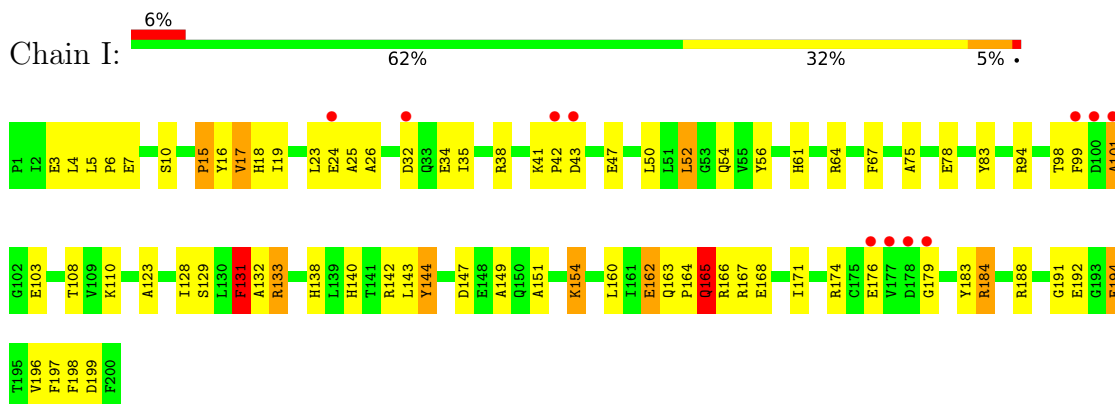




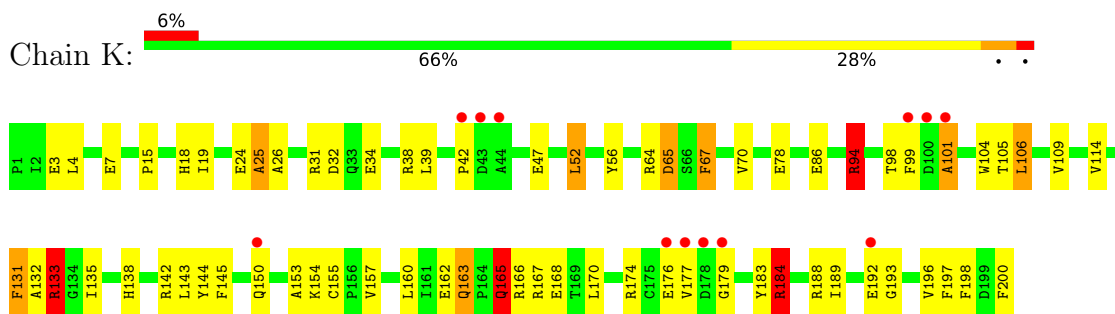
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



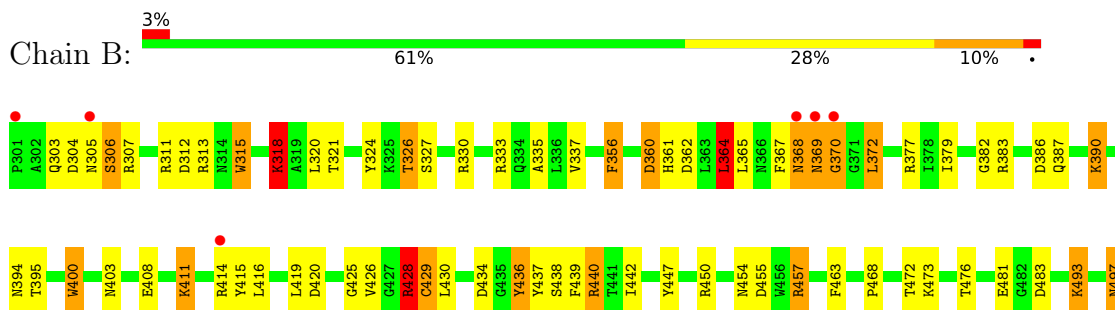
- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain



- Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain

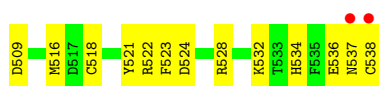


- Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

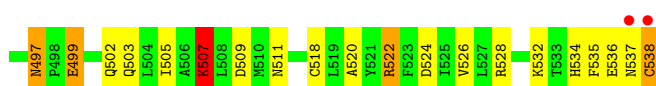
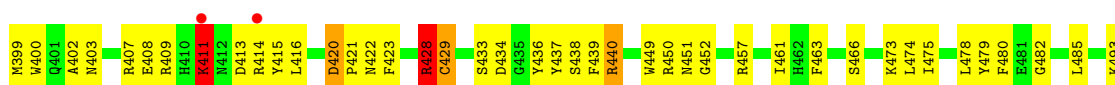




- Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



- Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

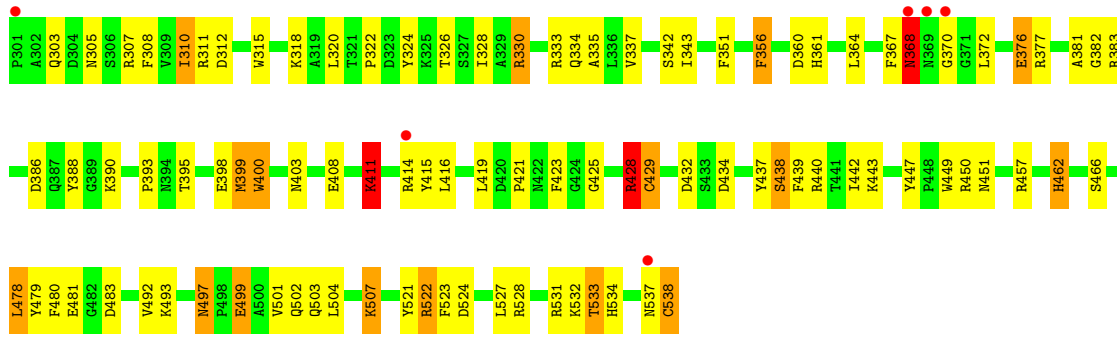


- Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

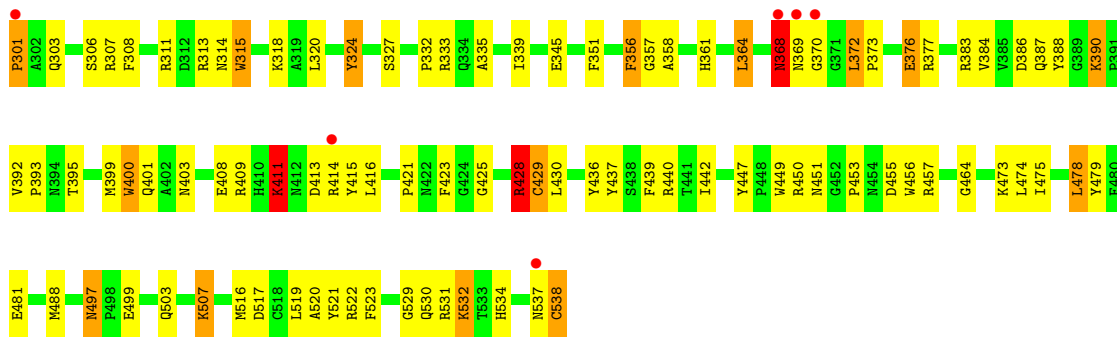


- Molecule 2: Protocatechuate 3,4-dioxygenase beta chain





• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	195.82Å 128.57Å 134.36Å 90.00° 97.98° 90.00°	Depositor
Resolution (Å)	8.51 – 2.06 20.06 – 1.94	Depositor EDS
% Data completeness (in resolution range)	69.6 (8.51-2.06) 68.8 (20.06-1.94)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.00 (at 1.94Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.149 , 0.192 0.145 , 0.186	Depositor DCC
$R_{free}$ test set	1950 reflections (0.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.3	Xtrriage
Anisotropy	0.221	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.46 , 66.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	22266	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.76% 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, DHB, CME

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	2.25	50/1611 (3.1%)	1.68	30/2195 (1.4%)
1	C	2.25	61/1611 (3.8%)	1.90	39/2195 (1.8%)
1	E	2.26	70/1611 (4.3%)	1.56	19/2195 (0.9%)
1	G	2.20	60/1611 (3.7%)	1.74	31/2195 (1.4%)
1	I	2.24	59/1611 (3.7%)	1.71	23/2195 (1.0%)
1	K	2.23	61/1611 (3.8%)	1.72	20/2195 (0.9%)
2	B	2.07	56/1920 (2.9%)	1.81	48/2612 (1.8%)
2	D	2.11	58/1920 (3.0%)	1.75	36/2612 (1.4%)
2	F	2.16	62/1920 (3.2%)	1.91	44/2612 (1.7%)
2	H	2.12	61/1920 (3.2%)	1.75	37/2612 (1.4%)
2	J	2.15	61/1920 (3.2%)	1.76	41/2612 (1.6%)
2	L	2.13	66/1920 (3.4%)	1.79	33/2612 (1.3%)
All	All	2.18	725/21186 (3.4%)	1.76	401/28842 (1.4%)

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	C	0	4
1	E	0	5
1	G	0	5
1	I	0	6
1	K	0	3
2	B	0	5
2	D	0	3
2	F	0	4
2	H	0	5
2	J	0	3

*Continued on next page...*

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
2	L	0	5
All	All	0	52

The worst 5 of 725 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	168	GLU	CD-OE1	15.12	1.42	1.25
1	A	192	GLU	CG-CD	14.26	1.73	1.51
2	L	532	LYS	CA-CB	13.83	1.84	1.53
2	D	449	TRP	CG-CD1	13.78	1.56	1.36
1	E	168	GLU	CD-OE1	12.07	1.39	1.25

The worst 5 of 401 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	440	ARG	NE-CZ-NH1	23.77	132.19	120.30
1	K	94	ARG	NE-CZ-NH1	22.83	131.72	120.30
1	K	94	ARG	NE-CZ-NH2	-22.73	108.94	120.30
1	C	184	ARG	NE-CZ-NH2	-22.09	109.26	120.30
2	H	440	ARG	NE-CZ-NH2	-21.08	109.76	120.30

There are no chirality outliers.

5 of 52 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	115	ASN	Mainchain
1	A	144	TYR	Sidechain
1	A	188	ARG	Sidechain
1	A	64	ARG	Mainchain
2	B	318	LYS	Mainchain

## 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	1571	0	1499	20	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1571	0	1499	18	0
1	E	1571	0	1499	26	0
1	G	1571	0	1499	20	0
1	I	1571	0	1499	21	0
1	K	1571	0	1499	20	0
2	B	1876	0	1822	37	0
2	D	1876	0	1822	27	0
2	F	1876	0	1821	31	0
2	H	1876	0	1822	27	0
2	J	1876	0	1822	32	0
2	L	1876	0	1822	32	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0
3	J	1	0	0	0	0
3	L	1	0	0	0	0
4	B	11	0	4	4	0
4	D	11	0	3	3	0
4	F	11	0	4	8	0
4	H	11	0	4	5	0
4	J	11	0	3	3	0
4	L	11	0	3	4	0
5	A	88	0	0	1	0
5	B	162	0	0	2	0
5	C	88	0	0	2	0
5	D	168	0	0	1	0
5	E	86	0	0	1	0
5	F	163	0	0	2	0
5	G	92	0	0	0	0
5	H	157	0	0	6	0
5	I	87	0	0	1	0
5	J	172	0	0	3	0
5	K	81	0	0	1	0
5	L	168	0	0	3	0
All	All	22266	0	19946	291	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 7.

The worst 5 of 291 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)
2:L:390:LYS:CG	2:L:390:LYS:CD	1.76	1.63
2:J:343:ILE:CD1	2:J:343:ILE:CG1	1.74	1.60
2:D:390:LYS:CD	2:D:390:LYS:CG	1.80	1.59
2:B:493:LYS:CG	2:B:493:LYS:CD	1.77	1.58
1:G:154:LYS:CE	1:G:154:LYS:CD	1.77	1.58

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	198/200 (99%)	188 (95%)	10 (5%)	0	100	100
1	C	198/200 (99%)	189 (96%)	9 (4%)	0	100	100
1	E	198/200 (99%)	192 (97%)	6 (3%)	0	100	100
1	G	198/200 (99%)	190 (96%)	8 (4%)	0	100	100
1	I	198/200 (99%)	191 (96%)	6 (3%)	1 (0%)	29	19
1	K	198/200 (99%)	184 (93%)	14 (7%)	0	100	100
2	B	235/238 (99%)	227 (97%)	7 (3%)	1 (0%)	34	25
2	D	235/238 (99%)	227 (97%)	7 (3%)	1 (0%)	34	25
2	F	235/238 (99%)	227 (97%)	7 (3%)	1 (0%)	34	25
2	H	235/238 (99%)	227 (97%)	7 (3%)	1 (0%)	34	25
2	J	235/238 (99%)	226 (96%)	8 (3%)	1 (0%)	34	25
2	L	235/238 (99%)	226 (96%)	8 (3%)	1 (0%)	34	25
All	All	2598/2628 (99%)	2494 (96%)	97 (4%)	7 (0%)	41	32

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	368	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	D	368	ASN
2	F	368	ASN
2	J	368	ASN
2	L	368	ASN

### 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	162/163 (99%)	154 (95%)	8 (5%)	25 17
1	C	162/163 (99%)	149 (92%)	13 (8%)	12 5
1	E	162/163 (99%)	154 (95%)	8 (5%)	25 17
1	G	162/163 (99%)	153 (94%)	9 (6%)	21 13
1	I	162/163 (99%)	155 (96%)	7 (4%)	29 22
1	K	162/163 (99%)	154 (95%)	8 (5%)	25 17
2	B	199/201 (99%)	189 (95%)	10 (5%)	24 16
2	D	199/201 (99%)	187 (94%)	12 (6%)	19 11
2	F	199/201 (99%)	188 (94%)	11 (6%)	21 13
2	H	199/201 (99%)	186 (94%)	13 (6%)	17 9
2	J	199/201 (99%)	188 (94%)	11 (6%)	21 13
2	L	199/201 (99%)	188 (94%)	11 (6%)	21 13
All	All	2166/2184 (99%)	2045 (94%)	121 (6%)	21 13

5 of 121 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	497	ASN
2	L	306	SER
2	H	352	SER
1	K	184	ARG
2	L	507	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 40 such sidechains are listed below:

Mol	Chain	Res	Type
2	J	361	HIS
2	L	369	ASN
2	J	368	ASN
1	K	165	GLN
2	L	497	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

6 non-standard protein/DNA/RNA residues 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	CME	B	429	2	8,9,10	3.39	3 (37%)	5,9,11	2.95	3 (60%)
2	CME	H	429	2	8,9,10	3.57	4 (50%)	5,9,11	2.91	5 (100%)
2	CME	J	429	2	8,9,10	2.66	2 (25%)	5,9,11	1.73	2 (40%)
2	CME	D	429	2	8,9,10	2.80	3 (37%)	5,9,11	2.56	3 (60%)
2	CME	F	429	2	8,9,10	3.19	3 (37%)	5,9,11	1.82	2 (40%)
2	CME	L	429	2	8,9,10	2.85	3 (37%)	5,9,11	2.98	3 (60%)

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	CME	B	429	2	-	3/5/8/10	-
2	CME	H	429	2	-	3/5/8/10	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CME	J	429	2	-	3/5/8/10	-
2	CME	D	429	2	-	3/5/8/10	-
2	CME	F	429	2	-	3/5/8/10	-
2	CME	L	429	2	-	3/5/8/10	-

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	429	CME	CB-SG	-7.93	1.56	1.81
2	F	429	CME	CB-SG	-7.17	1.58	1.81
2	H	429	CME	CB-SG	-6.92	1.59	1.81
2	D	429	CME	CB-SG	-6.49	1.60	1.81
2	J	429	CME	CB-SG	-6.44	1.60	1.81

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	429	CME	CZ-CE-SD	-4.91	96.32	113.37
2	L	429	CME	CB-SG-SD	-4.30	92.68	103.82
2	H	429	CME	CB-SG-SD	-4.07	93.29	103.82
2	D	429	CME	CE-SD-SG	3.31	118.72	103.45
2	L	429	CME	CZ-CE-SD	-3.23	102.15	113.37

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	429	CME	N-CA-CB-SG
2	H	429	CME	N-CA-CB-SG
2	L	429	CME	N-CA-CB-SG
2	B	429	CME	SD-CE-CZ-OH
2	D	429	CME	SD-CE-CZ-OH

There are no ring outliers.

6 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	429	CME	2	0
2	H	429	CME	2	0
2	J	429	CME	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	429	CME	2	0
2	F	429	CME	2	0
2	L	429	CME	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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$
4	DHB	L	5550	3	9,11,11	0.88	1 (11%)	12,15,15	0.54	0
4	DHB	J	4550	3	9,11,11	0.91	0	12,15,15	0.44	0
4	DHB	F	2550	3	9,11,11	1.22	2 (22%)	12,15,15	0.54	0
4	DHB	B	550	3	9,11,11	1.52	2 (22%)	12,15,15	0.69	0
4	DHB	H	3550	3	9,11,11	1.88	4 (44%)	12,15,15	0.78	0
4	DHB	D	1550	3	9,11,11	1.33	2 (22%)	12,15,15	0.57	0

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
4	DHB	L	5550	3	-	0/0/4/4	0/1/1/1
4	DHB	J	4550	3	-	0/0/4/4	0/1/1/1
4	DHB	F	2550	3	-	0/0/4/4	0/1/1/1
4	DHB	B	550	3	-	0/0/4/4	0/1/1/1
4	DHB	H	3550	3	-	0/0/4/4	0/1/1/1
4	DHB	D	1550	3	-	0/0/4/4	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1550	DHB	C1-C	-2.98	1.44	1.47
4	B	550	DHB	C4-C3	-2.79	1.35	1.40
4	H	3550	DHB	C2-C3	-2.73	1.35	1.38
4	F	2550	DHB	C4-C3	-2.53	1.36	1.40
4	H	3550	DHB	C5-C4	-2.53	1.35	1.39

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 27 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	5550	DHB	4	0
4	J	4550	DHB	3	0
4	F	2550	DHB	8	0
4	B	550	DHB	4	0
4	H	3550	DHB	5	0
4	D	1550	DHB	3	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	200/200 (100%)	-0.43	6 (3%) 50 53	6, 18, 47, 61	0
1	C	200/200 (100%)	-0.40	5 (2%) 57 60	7, 19, 45, 62	0
1	E	200/200 (100%)	-0.35	8 (4%) 38 40	7, 20, 48, 61	0
1	G	200/200 (100%)	-0.33	9 (4%) 33 34	7, 20, 49, 62	0
1	I	200/200 (100%)	-0.27	11 (5%) 25 26	8, 21, 49, 63	0
1	K	200/200 (100%)	-0.17	12 (6%) 21 22	9, 23, 49, 62	0
2	B	237/238 (99%)	-0.61	8 (3%) 45 47	7, 13, 38, 58	0
2	D	237/238 (99%)	-0.60	7 (2%) 50 53	6, 13, 39, 60	0
2	F	237/238 (99%)	-0.65	8 (3%) 45 47	5, 14, 38, 58	0
2	H	237/238 (99%)	-0.64	6 (2%) 57 60	7, 14, 39, 54	0
2	J	237/238 (99%)	-0.64	6 (2%) 57 60	7, 15, 39, 59	0
2	L	237/238 (99%)	-0.65	6 (2%) 57 60	9, 15, 39, 59	0
All	All	2622/2628 (99%)	-0.49	92 (3%) 44 46	5, 17, 46, 63	0

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	99	PHE	6.0
2	D	369	ASN	5.6
2	L	301	PRO	5.1
1	C	100	ASP	5.1
1	K	178	ASP	5.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
2	CME	D	429	10/11	0.92	0.12	17,25,51,51	0
2	CME	F	429	10/11	0.93	0.12	16,23,49,50	0
2	CME	H	429	10/11	0.93	0.12	14,23,49,49	0
2	CME	L	429	10/11	0.94	0.10	19,27,49,50	0
2	CME	J	429	10/11	0.95	0.10	20,26,50,51	0
2	CME	B	429	10/11	0.95	0.11	14,24,46,47	0

### 6.3 Carbohydrates [i](#)

There are no monosaccharides 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(Å <sup>2</sup> )	Q<0.9
3	FE	H	3600	1/1	0.68	0.09	82,82,82,82	0
3	FE	L	5600	1/1	0.70	0.07	75,75,75,75	0
3	FE	F	2600	1/1	0.79	0.07	77,77,77,77	0
4	DHB	L	5550	11/11	0.79	0.19	54,56,58,60	0
4	DHB	B	550	11/11	0.80	0.19	46,49,50,51	0
3	FE	B	600	1/1	0.81	0.08	83,83,83,83	0
4	DHB	J	4550	11/11	0.82	0.18	50,51,52,53	0
4	DHB	F	2550	11/11	0.82	0.21	55,57,59,60	0
4	DHB	D	1550	11/11	0.84	0.19	51,52,55,55	0
3	FE	D	1600	1/1	0.89	0.06	69,69,69,69	0
4	DHB	H	3550	11/11	0.89	0.18	54,55,57,57	0
3	FE	J	4600	1/1	0.91	0.07	80,80,80,80	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.