



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:43 am GMT

PDB ID : 1YWK  
Title : Crystal structure of 4-deoxy-1-threo-5-hexosulose-uronate ketol-isomerase from *Enterococcus faecalis*  
Authors : Fedorov, A.A.; Fedorov, E.V.; Almo, S.C.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)  
Deposited on : 2005-02-18  
Resolution : 2.95 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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

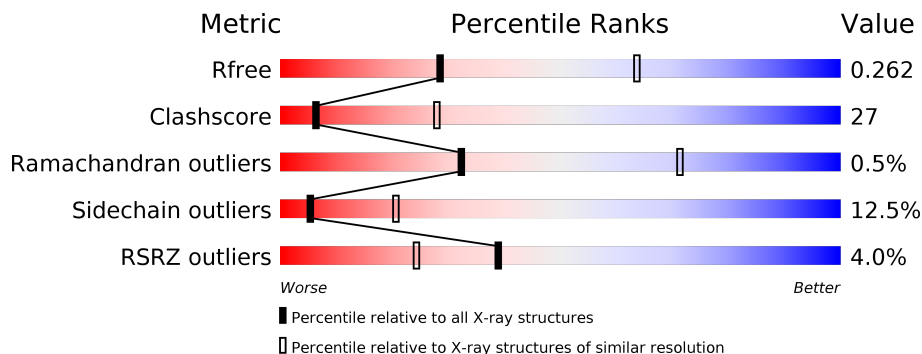
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.95 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))
$R_{free}$	100719	2395 (3.00-2.92)
Clashscore	112137	2773 (3.00-2.92)
Ramachandran outliers	110173	2680 (3.00-2.92)
Sidechain outliers	110143	2683 (3.00-2.92)
RSRZ outliers	101464	2421 (3.00-2.92)

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	289	<div> <div>5%</div> <div> <div>51%</div> <div>31%</div> <div>6%</div> <div>12%</div> </div> </div>
1	B	289	<div> <div>3%</div> <div> <div>48%</div> <div>33%</div> <div>7%</div> <div>12%</div> </div> </div>
1	C	289	<div> <div>2%</div> <div> <div>49%</div> <div>33%</div> <div>6%</div> <div>12%</div> </div> </div>
1	D	289	<div> <div>2%</div> <div> <div>49%</div> <div>32%</div> <div>7%</div> <div>12%</div> </div> </div>
1	E	289	<div> <div>5%</div> <div> <div>48%</div> <div>34%</div> <div>6%</div> <div>12%</div> </div> </div>
1	F	289	<div> <div>4%</div> <div> <div>46%</div> <div>35%</div> <div>7%</div> <div>12%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 12252 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 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	255	Total	C	N	O	S	0	0	0
			2053	1302	340	396	15			
1	B	254	Total	C	N	O	S	0	0	0
			2042	1296	336	395	15			
1	C	253	Total	C	N	O	S	0	0	0
			2036	1291	338	392	15			
1	D	254	Total	C	N	O	S	0	0	0
			2045	1296	339	395	15			
1	E	254	Total	C	N	O	S	0	0	0
			2042	1296	336	395	15			
1	F	253	Total	C	N	O	S	0	0	0
			2034	1290	335	394	15			

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	996	MET	-	CLONING ARTIFACT	UNP Q838L9
A	997	SER	-	CLONING ARTIFACT	UNP Q838L9
A	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
A	999	GLN	-	CLONING ARTIFACT	UNP Q838L9
A	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
A	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
A	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9
A	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
A	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
A	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
A	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
A	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
A	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	996	MET	-	CLONING ARTIFACT	UNP Q838L9
B	997	SER	-	CLONING ARTIFACT	UNP Q838L9
B	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
B	999	GLN	-	CLONING ARTIFACT	UNP Q838L9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
B	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
B	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9
B	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
B	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	996	MET	-	CLONING ARTIFACT	UNP Q838L9
C	997	SER	-	CLONING ARTIFACT	UNP Q838L9
C	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
C	999	GLN	-	CLONING ARTIFACT	UNP Q838L9
C	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
C	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
C	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9
C	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
C	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	996	MET	-	CLONING ARTIFACT	UNP Q838L9
D	997	SER	-	CLONING ARTIFACT	UNP Q838L9
D	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
D	999	GLN	-	CLONING ARTIFACT	UNP Q838L9
D	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
D	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
D	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9
D	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
D	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	996	MET	-	CLONING ARTIFACT	UNP Q838L9
E	997	SER	-	CLONING ARTIFACT	UNP Q838L9
E	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
E	999	GLN	-	CLONING ARTIFACT	UNP Q838L9
E	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
E	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
E	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9

*Continued on next page...*

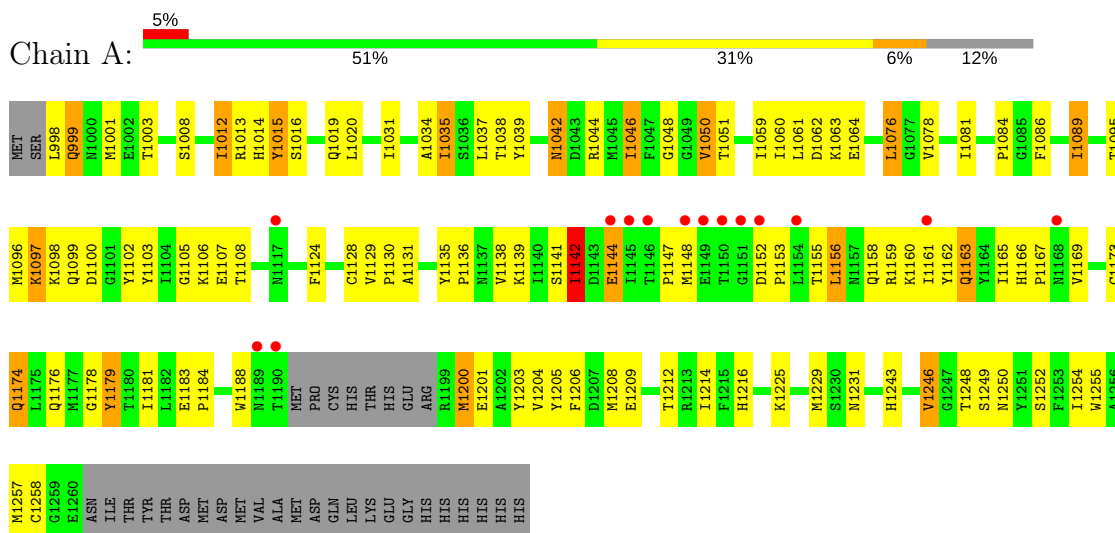
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
E	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	996	MET	-	CLONING ARTIFACT	UNP Q838L9
F	997	SER	-	CLONING ARTIFACT	UNP Q838L9
F	998	LEU	-	CLONING ARTIFACT	UNP Q838L9
F	999	GLN	-	CLONING ARTIFACT	UNP Q838L9
F	1000	ASN	-	CLONING ARTIFACT	UNP Q838L9
F	1277	GLU	-	CLONING ARTIFACT	UNP Q838L9
F	1278	GLY	-	CLONING ARTIFACT	UNP Q838L9
F	1279	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	1280	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	1281	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	1282	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	1283	HIS	-	CLONING ARTIFACT	UNP Q838L9
F	1284	HIS	-	CLONING ARTIFACT	UNP Q838L9

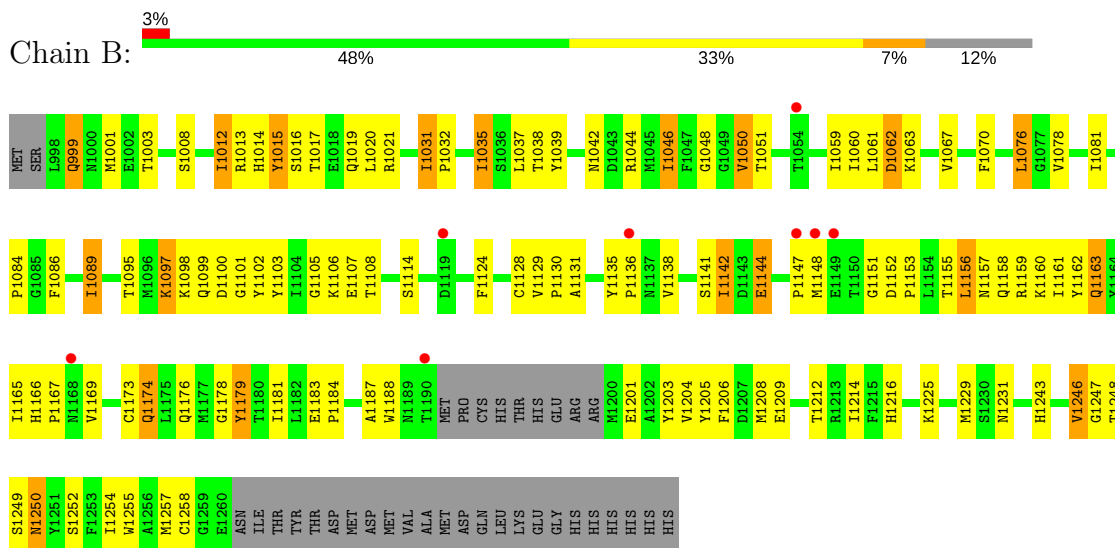
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1

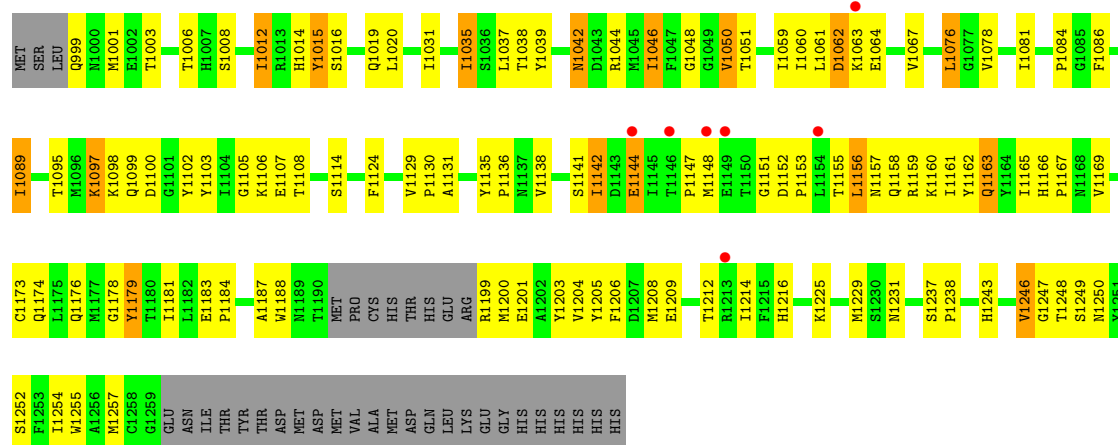


- Molecule 1: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1

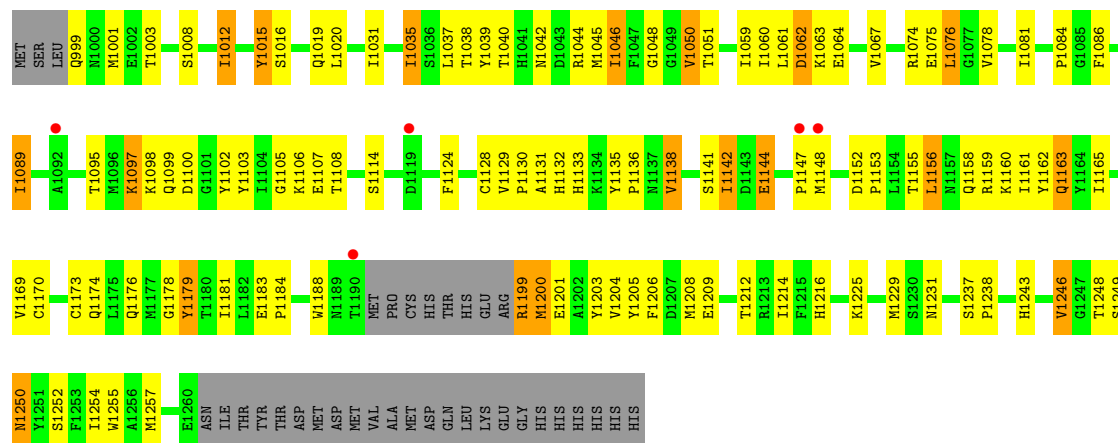


- Molecule 1: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1

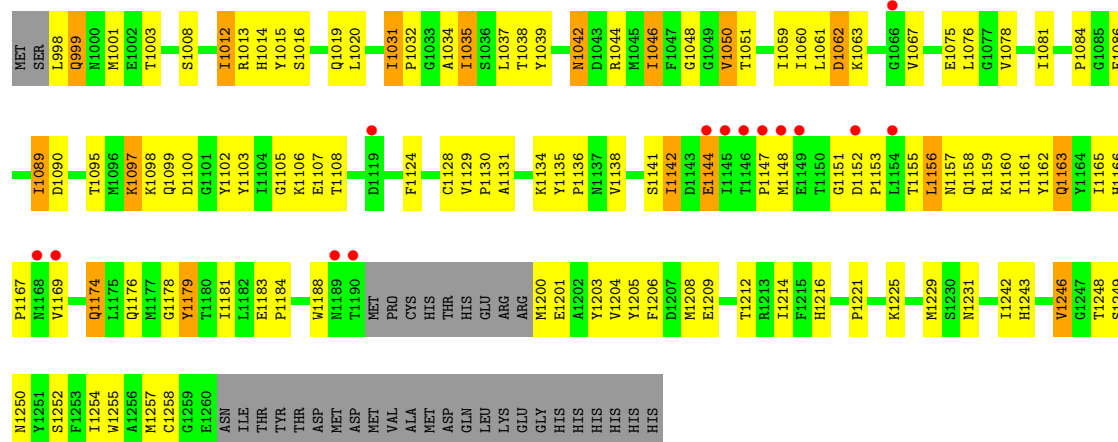




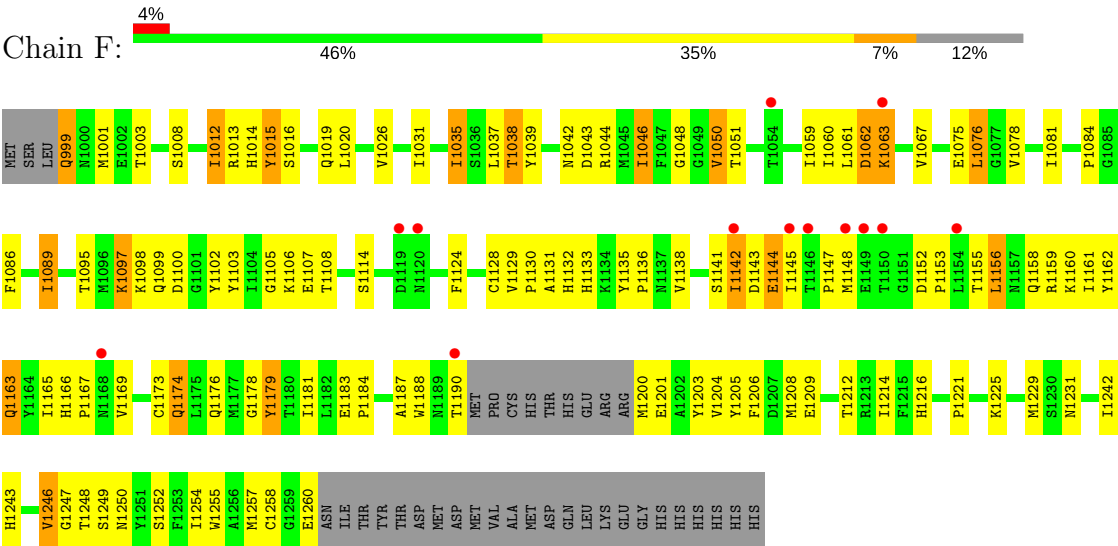
- Molecule 1: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1



- Molecule 1: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1



- Molecule 1: 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase 1





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.01Å 107.69Å 191.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	24.90 – 2.95 29.69 – 2.95	Depositor EDS
% Data completeness (in resolution range)	96.9 (24.90-2.95) 96.8 (29.69-2.95)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 2.95Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.241 , 0.265 0.240 , 0.262	Depositor DCC
$R_{free}$ test set	1722 reflections (4.95%)	DCC
Wilson B-factor (Å <sup>2</sup> )	46.8	Xtriage
Anisotropy	0.351	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 43.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	12252	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	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 37.88 % 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 4.0135e-04. 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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	0.42	0/2105	0.65	0/2850
1	B	0.42	0/2094	0.66	0/2836
1	C	0.41	0/2088	0.65	0/2827
1	D	0.41	0/2097	0.66	0/2839
1	E	0.40	0/2094	0.65	0/2836
1	F	0.41	0/2086	0.65	0/2825
All	All	0.41	0/12564	0.65	0/17013

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2053	0	1970	100	0
1	B	2042	0	1957	100	0
1	C	2036	0	1953	111	0
1	D	2045	0	1959	121	0
1	E	2042	0	1957	111	0
1	F	2034	0	1946	123	0
All	All	12252	0	11742	640	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 27.

The worst 5 of 640 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:E:999:GLN:HB2	1:E:1035:ILE:HG12	1.33	1.10
1:D:1199:ARG:HH11	1:D:1199:ARG:HB2	1.29	0.96
1:A:1016:SER:H	1:A:1019:GLN:HE21	1.20	0.90
1:C:1016:SER:H	1:C:1019:GLN:HE21	1.16	0.90
1:E:1016:SER:H	1:E:1019:GLN:HE21	1.19	0.88

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	251/289 (87%)	232 (92%)	17 (7%)	2 (1%)	22	61
1	B	250/289 (86%)	229 (92%)	19 (8%)	2 (1%)	22	61
1	C	249/289 (86%)	233 (94%)	15 (6%)	1 (0%)	38	75
1	D	250/289 (86%)	230 (92%)	19 (8%)	1 (0%)	38	75
1	E	250/289 (86%)	232 (93%)	17 (7%)	1 (0%)	38	75
1	F	249/289 (86%)	229 (92%)	19 (8%)	1 (0%)	38	75
All	All	1499/1734 (86%)	1385 (92%)	106 (7%)	8 (0%)	32	71

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1144	GLU
1	B	1144	GLU
1	C	1144	GLU
1	E	1144	GLU
1	F	1144	GLU

### 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	227/259 (88%)	198 (87%)	29 (13%)	5	20
1	B	226/259 (87%)	197 (87%)	29 (13%)	5	20
1	C	225/259 (87%)	198 (88%)	27 (12%)	6	23
1	D	226/259 (87%)	196 (87%)	30 (13%)	4	18
1	E	226/259 (87%)	199 (88%)	27 (12%)	6	23
1	F	225/259 (87%)	197 (88%)	28 (12%)	5	21
All	All	1355/1554 (87%)	1185 (88%)	170 (12%)	5	20

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	1142	ILE
1	D	1061	LEU
1	F	1089	ILE
1	C	1163	GLN
1	D	1012	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	1216	HIS
1	D	1132	HIS
1	F	1174	GLN
1	C	1243	HIS
1	D	1019	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](#)

There are no ligands in this entry.

## 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	255/289 (88%)	0.18	14 (5%) 26 15	13, 37, 69, 99	0
1	B	254/289 (87%)	0.12	8 (3%) 49 31	15, 36, 64, 98	0
1	C	253/289 (87%)	0.10	7 (2%) 53 35	19, 41, 67, 98	0
1	D	254/289 (87%)	0.12	5 (1%) 65 46	13, 39, 64, 95	0
1	E	254/289 (87%)	0.07	14 (5%) 26 15	11, 36, 73, 95	0
1	F	253/289 (87%)	0.22	13 (5%) 29 18	19, 42, 70, 96	0
All	All	1523/1734 (87%)	0.14	61 (4%) 39 24	11, 39, 68, 99	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	1148	MET	7.2
1	A	1146	THR	6.3
1	C	1148	MET	6.2
1	A	1144	GLU	6.1
1	E	1146	THR	6.0

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.