



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 10:52 pm BST

PDB ID : 1W91  
Title : crystal structure of 1,4-BETA-D-XYLAN XYLOHYDROLASE solve using anomalous signal from Seleniomethionine  
Authors : Jakoncic, J.; Shoham, G.; Stojanoff, V.  
Deposited on : 2004-10-05  
Resolution : 2.20 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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

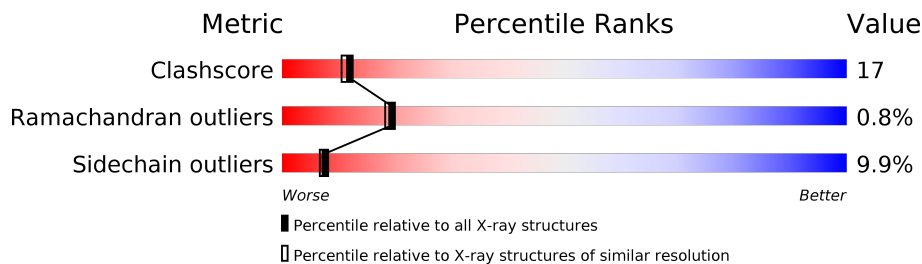
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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 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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	503	
1	B	503	
1	C	503	
1	D	503	
1	E	503	
1	F	503	
1	G	503	
1	H	503	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 35545 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 BETA-XYLOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	499	4082	2640	691	740	11	0	0	0
1	B	499	4082	2640	691	740	11	0	0	0
1	C	499	4082	2640	691	740	11	0	0	0
1	D	499	4082	2640	691	740	11	0	0	0
1	E	499	4082	2640	691	740	11	0	0	0
1	F	499	4082	2640	691	740	11	0	0	0
1	G	499	4082	2640	691	740	11	0	0	0
1	H	499	4082	2640	691	740	11	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	445	ARG	PRO	conflict	UNP Q9ZFM2
A	446	GLN	SER	conflict	UNP Q9ZFM2
B	445	ARG	PRO	conflict	UNP Q9ZFM2
B	446	GLN	SER	conflict	UNP Q9ZFM2
C	445	ARG	PRO	conflict	UNP Q9ZFM2
C	446	GLN	SER	conflict	UNP Q9ZFM2
D	445	ARG	PRO	conflict	UNP Q9ZFM2
D	446	GLN	SER	conflict	UNP Q9ZFM2
E	445	ARG	PRO	conflict	UNP Q9ZFM2
E	446	GLN	SER	conflict	UNP Q9ZFM2
F	445	ARG	PRO	conflict	UNP Q9ZFM2
F	446	GLN	SER	conflict	UNP Q9ZFM2
G	445	ARG	PRO	conflict	UNP Q9ZFM2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
G	446	GLN	SER	conflict	UNP Q9ZFM2
H	445	ARG	PRO	conflict	UNP Q9ZFM2
H	446	GLN	SER	conflict	UNP Q9ZFM2

- Molecule 2 is water.

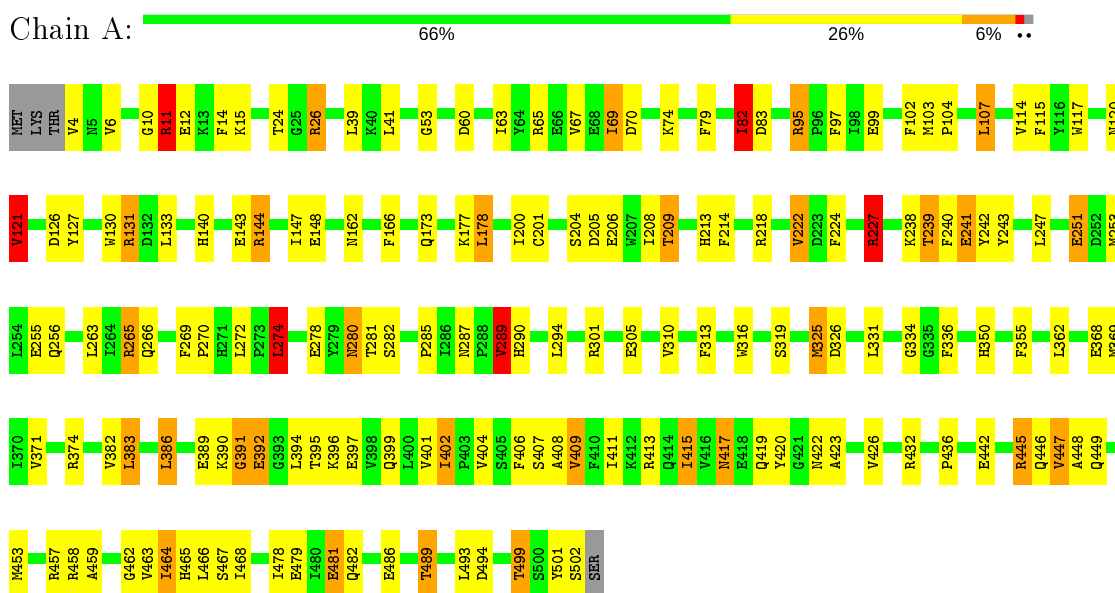
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	363	Total O 363 363	0	0
2	B	317	Total O 317 317	0	0
2	C	371	Total O 371 371	0	0
2	D	340	Total O 340 340	0	0
2	E	347	Total O 347 347	0	0
2	F	414	Total O 414 414	0	0
2	G	340	Total O 340 340	0	0
2	H	397	Total O 397 397	0	0

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

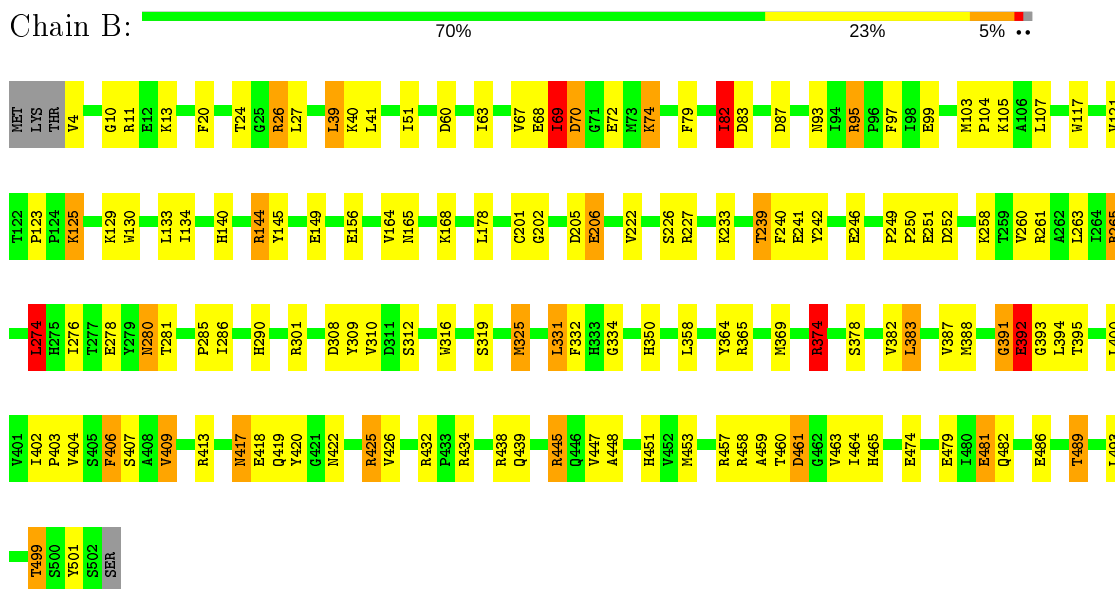
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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. 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.

Note EDS was not executed.

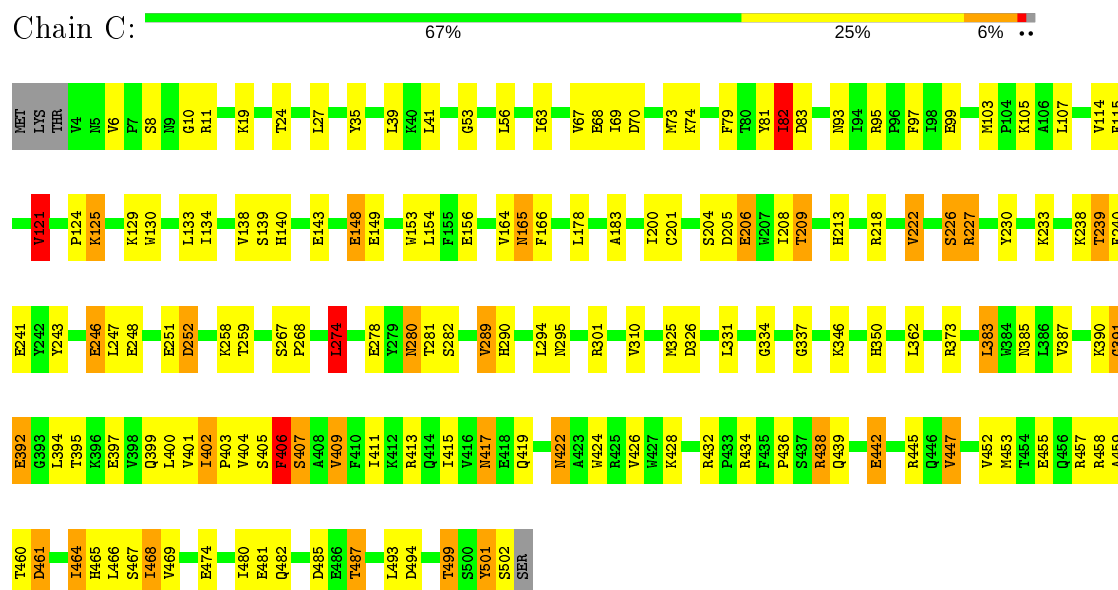
- Molecule 1: BETA-XYLOSIDASE



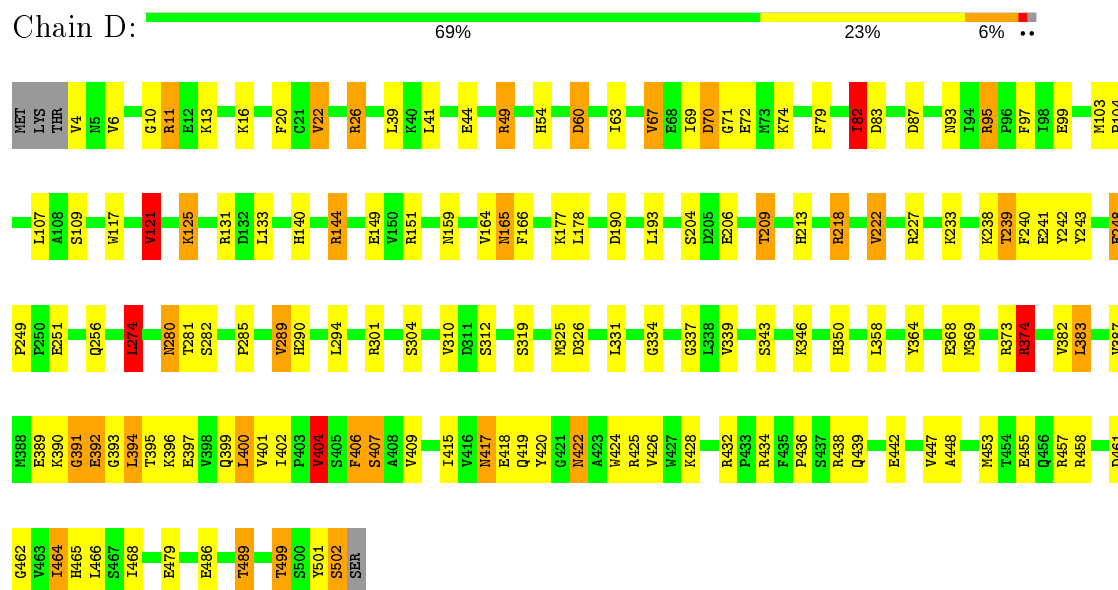
- Molecule 1: BETA-XYLOSIDASE



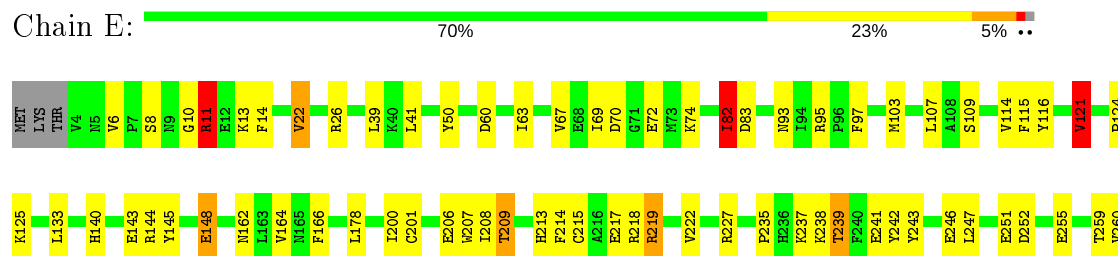
- Molecule 1: BETA-XYLOSIDASE

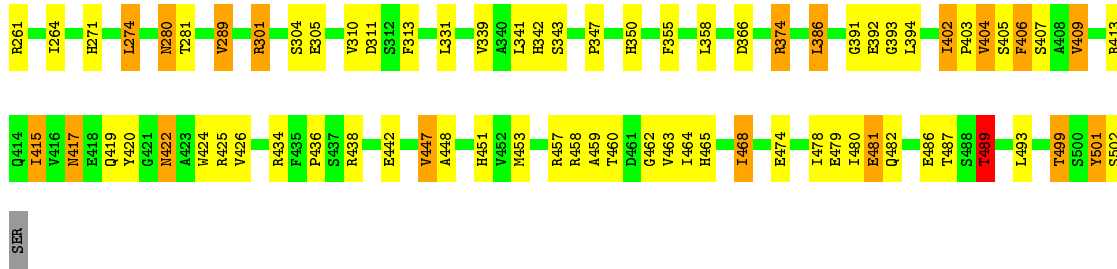


- Molecule 1: BETA-XYLOSIDASE

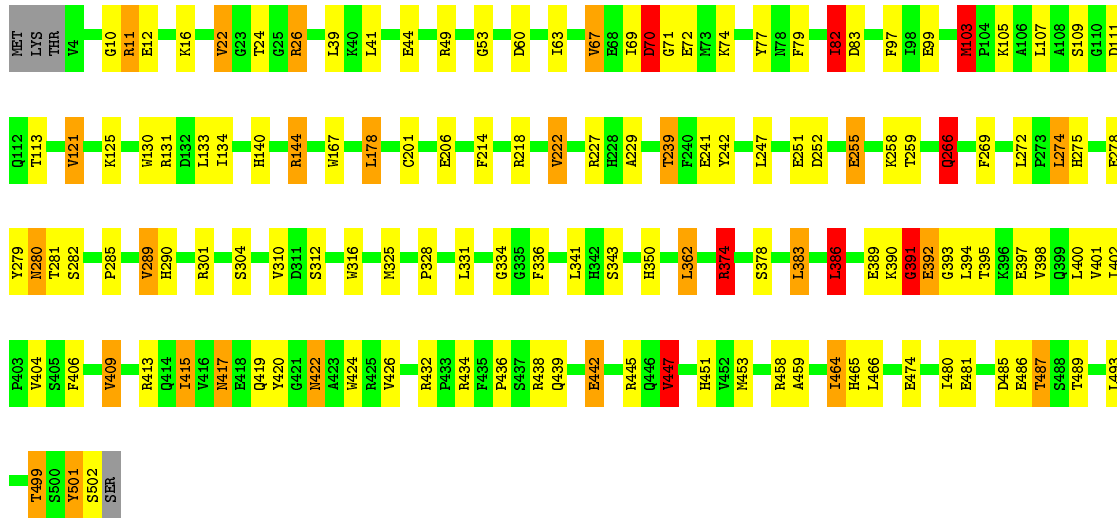


- Molecule 1: BETA-XYLOSIDASE

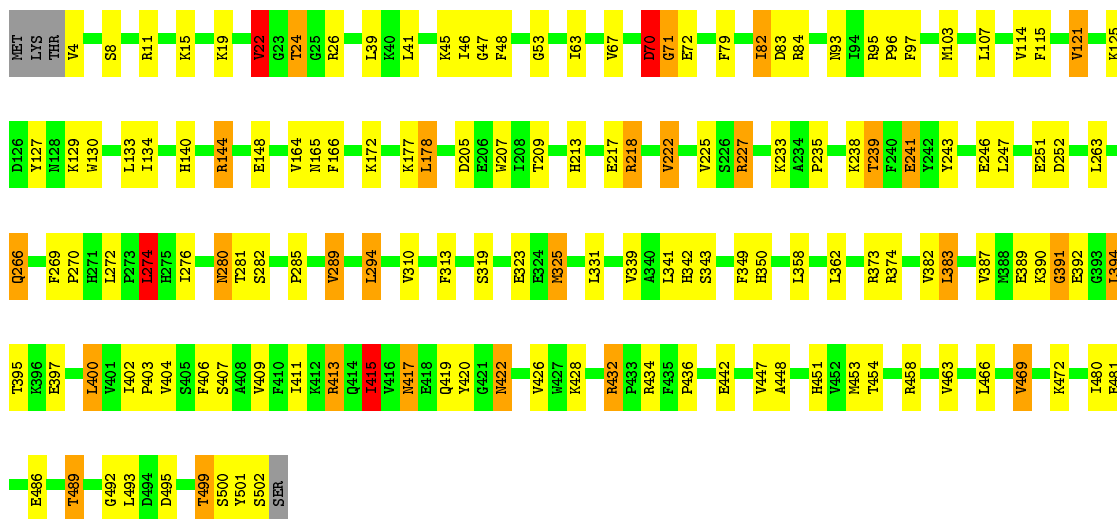




• Molecule 1: BETA-XYLOSIDASE

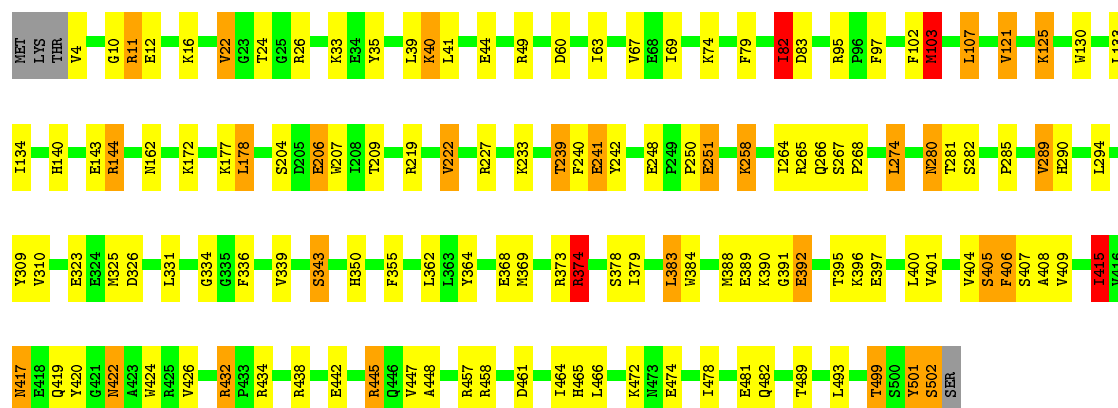


• Molecule 1: BETA-XYLOSIDASE



• Molecule 1: BETA-XYLOSIDASE

Chain H:  72% 21% 6% ..





## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.70Å 166.02Å 313.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.20	Depositor
% Data completeness (in resolution range)	100.0 (8.00-2.20)	Depositor
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.174 , 0.242	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	35545	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	1.05	6/4198 (0.1%)	1.04	23/5699 (0.4%)
1	B	1.04	4/4198 (0.1%)	1.00	18/5699 (0.3%)
1	C	1.09	9/4198 (0.2%)	1.03	19/5699 (0.3%)
1	D	1.07	1/4197 (0.0%)	1.01	19/5696 (0.3%)
1	E	1.10	9/4198 (0.2%)	1.06	14/5699 (0.2%)
1	F	1.11	8/4198 (0.2%)	1.09	27/5699 (0.5%)
1	G	1.03	3/4198 (0.1%)	1.05	22/5699 (0.4%)
1	H	1.13	6/4198 (0.1%)	1.08	24/5699 (0.4%)
All	All	1.08	46/33583 (0.1%)	1.05	166/45589 (0.4%)

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	B	0	1
1	F	0	3
1	H	0	2
All	All	0	6

The worst 5 of 46 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	501	TYR	C-N	15.40	1.69	1.34
1	E	501	TYR	C-N	15.16	1.69	1.34
1	F	501	TYR	C-N	11.59	1.60	1.34
1	F	442	GLU	CG-CD	10.00	1.67	1.51
1	A	501	TYR	C-N	9.61	1.56	1.34

The worst 5 of 166 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	501	TYR	O-C-N	-20.19	90.40	122.70
1	F	501	TYR	O-C-N	-18.79	92.63	122.70
1	C	501	TYR	O-C-N	-18.28	93.45	122.70
1	B	501	TYR	C-N-CA	17.61	165.72	121.70
1	H	501	TYR	C-N-CA	17.18	164.65	121.70

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	69	ILE	Peptide
1	F	391	GLY	Peptide
1	F	501	TYR	Mainchain
1	F	71	GLY	Peptide
1	H	390	LYS	Peptide

## 5.2 Too-close contacts [i](#)

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	4082	0	3992	157	0
1	B	4082	0	3992	138	0
1	C	4082	0	3992	165	0
1	D	4082	0	3991	159	0
1	E	4082	0	3992	142	0
1	F	4082	0	3992	139	0
1	G	4082	0	3991	115	0
1	H	4082	0	3992	140	0
2	A	363	0	0	43	1
2	B	317	0	0	26	0
2	C	371	0	0	37	0
2	D	340	0	0	37	1
2	E	347	0	0	41	0
2	F	414	0	0	48	1
2	G	340	0	0	25	1
2	H	397	0	0	41	0
All	All	35545	0	31934	1109	2

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 17.

The worst 5 of 1109 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:F:103:MET:CB	1:F:103:MET:CG	1.76	1.60
1:H:103:MET:HE3	1:H:103:MET:CB	1.11	1.57
1:H:103:MET:CG	1:H:103:MET:CB	1.76	1.54
1:H:103:MET:HB3	1:H:103:MET:CE	1.02	1.47
1:E:501:TYR:C	1:E:502:SER:N	1.68	1.46

All (2) 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)
2:D:2193:HOH:O	2:G:2069:HOH:O[3_645]	1.99	0.21
2:A:2007:HOH:O	2:F:2015:HOH:O[4_556]	2.18	0.02

## 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	497/503 (99%)	467 (94%)	27 (5%)	3 (1%)	25 26
1	B	497/503 (99%)	473 (95%)	18 (4%)	6 (1%)	13 10
1	C	497/503 (99%)	469 (94%)	24 (5%)	4 (1%)	19 19
1	D	496/503 (99%)	468 (94%)	22 (4%)	6 (1%)	13 10
1	E	497/503 (99%)	475 (96%)	21 (4%)	1 (0%)	47 55
1	F	497/503 (99%)	468 (94%)	25 (5%)	4 (1%)	19 19
1	G	497/503 (99%)	474 (95%)	20 (4%)	3 (1%)	25 26
1	H	497/503 (99%)	476 (96%)	18 (4%)	3 (1%)	25 26

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	3975/4024 (99%)	3770 (95%)	175 (4%)	30 (1%)	19 19

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	391	GLY
1	B	406	PHE
1	B	461	ASP
1	C	391	GLY
1	C	406	PHE

### 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	439/443 (99%)	392 (89%)	47 (11%)	6 6
1	B	439/443 (99%)	402 (92%)	37 (8%)	11 11
1	C	439/443 (99%)	394 (90%)	45 (10%)	7 6
1	D	439/443 (99%)	392 (89%)	47 (11%)	6 6
1	E	439/443 (99%)	399 (91%)	40 (9%)	9 9
1	F	439/443 (99%)	397 (90%)	42 (10%)	8 8
1	G	439/443 (99%)	392 (89%)	47 (11%)	6 6
1	H	439/443 (99%)	397 (90%)	42 (10%)	8 8
All	All	3512/3544 (99%)	3165 (90%)	347 (10%)	8 7

5 of 347 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	319	SER
1	E	289	VAL
1	H	219	ARG
1	D	394	LEU
1	E	8	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 105 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	350	HIS
1	E	385	ASN
1	H	295	ASN
1	D	385	ASN
1	E	140	HIS

### 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1
1	C	1
1	F	1
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	501:TYR	C	502:SER	N	2.24
1	C	501:TYR	C	502:SER	N	1.69
1	E	501:TYR	C	502:SER	N	1.68
1	F	501:TYR	C	502:SER	N	1.60

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.