



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 10:53 PM GMT

PDB ID : 1VKD  
Title : Crystal structure of a predicted glycosidase (tm1225) from thermotoga maritima msb8 at 2.10 Å resolution  
Authors : Joint Center for Structural Genomics (JCSG)  
Deposited on : 2004-05-12  
Resolution : 2.10 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

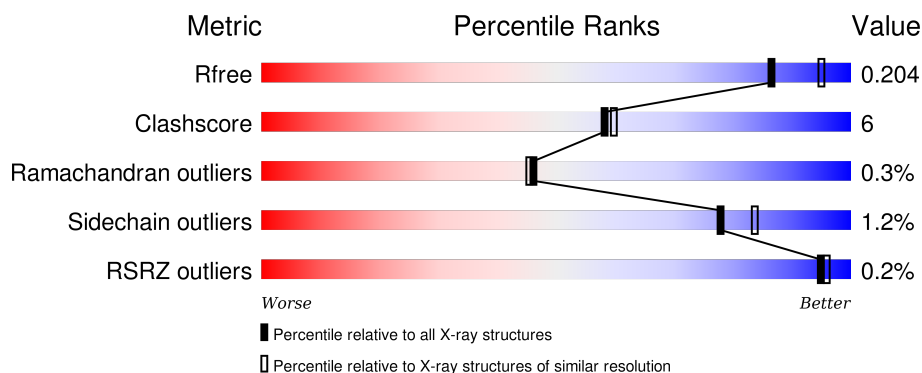
# 1 Overall quality at a glance

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))
$R_{free}$	91344	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (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	338	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 15%, green 80%, grey 14%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>80%</span> <span>15%</span> </div> </div>
1	B	338	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, yellow 10%, green 85%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>85%</span> <span>10%</span> </div> </div>
1	C	338	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, yellow 16%, green 80%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>80%</span> <span>16%</span> </div> </div>
1	D	338	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, yellow 14%, green 81%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>81%</span> <span>14%</span> </div> </div>
1	E	338	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, yellow 17%, green 78%, grey 5%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>78%</span> <span>17%</span> </div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	338	<div><div></div><div>84%</div><div>11%<div></div><div></div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 16823 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 conserved hypothetical protein TM1225.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	327	Total	C	N	O	S	Se	0	0	0
			2652	1720	443	480	4	5			
1	B	326	Total	C	N	O	S	Se	0	0	0
			2642	1714	440	479	4	5			
1	C	326	Total	C	N	O	S	Se	0	0	0
			2642	1714	440	479	4	5			
1	D	326	Total	C	N	O	S	Se	0	0	0
			2642	1714	440	479	4	5			
1	E	325	Total	C	N	O	S	Se	0	0	0
			2633	1709	439	477	4	4			
1	F	326	Total	C	N	O	S	Se	0	0	0
			2642	1714	440	479	4	5			

There are 102 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
A	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
A	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
A	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
A	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
A	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
A	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
A	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
A	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
A	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
A	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
B	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
B	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
B	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
B	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
B	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
B	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
B	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
B	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
B	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
B	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
C	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
C	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
C	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
C	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
C	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
C	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
C	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
C	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
C	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
C	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
C	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
D	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
D	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
D	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
D	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
D	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
D	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
D	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
D	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2

*Continued on next page...*

*Continued from previous page...*

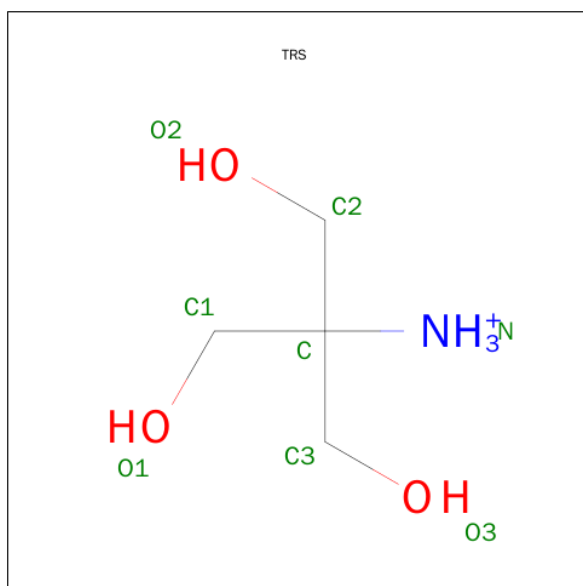
Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
D	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
D	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
D	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
D	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
D	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
D	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
D	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
E	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
E	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
E	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
E	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
E	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
E	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
E	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
E	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
E	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
E	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
E	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
F	-11	MET	-	LEADER SEQUENCE	UNP Q9X0V2
F	-10	GLY	-	LEADER SEQUENCE	UNP Q9X0V2
F	-9	SER	-	LEADER SEQUENCE	UNP Q9X0V2
F	-8	ASP	-	LEADER SEQUENCE	UNP Q9X0V2
F	-7	LYS	-	LEADER SEQUENCE	UNP Q9X0V2
F	-6	ILE	-	LEADER SEQUENCE	UNP Q9X0V2
F	-5	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	-4	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	-3	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	-2	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	-1	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	0	HIS	-	LEADER SEQUENCE	UNP Q9X0V2
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
F	133	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
F	167	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2
F	190	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	326	MSE	MET	MODIFIED RESIDUE	UNP Q9X0V2

- Molecule 2 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula:  $C_4H_{12}NO_3$ ).



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

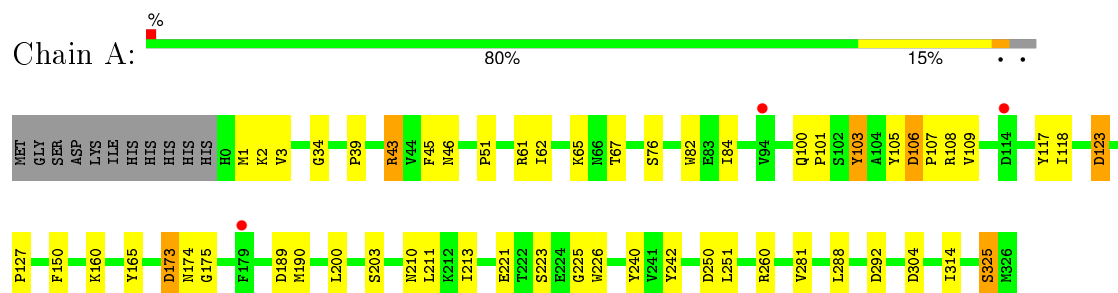
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	144	Total	O	0	0
			144	144		
3	B	154	Total	O	0	0
			154	154		
3	C	142	Total	O	0	0
			142	142		
3	D	184	Total	O	0	0
			184	184		
3	E	128	Total	O	0	0
			128	128		
3	F	202	Total	O	0	0
			202	202		

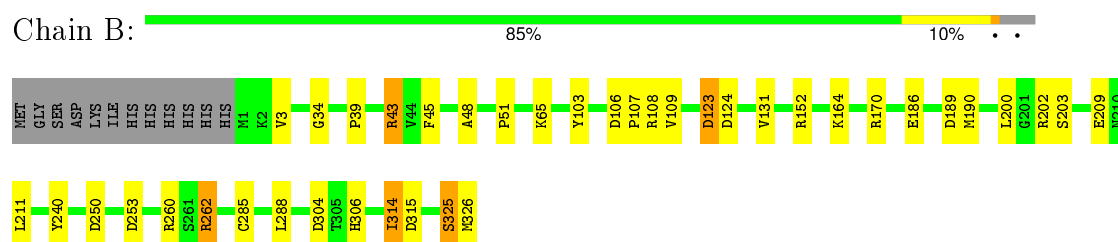
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

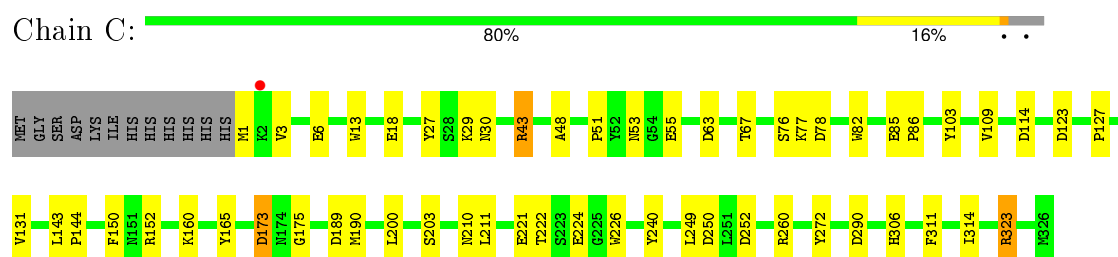
- Molecule 1: conserved hypothetical protein TM1225



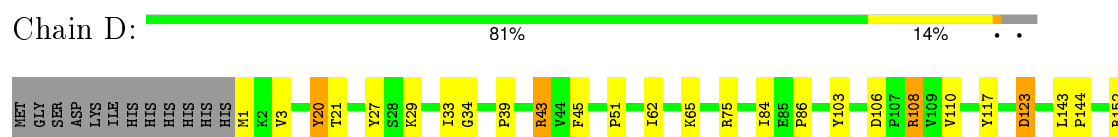
- Molecule 1: conserved hypothetical protein TM1225



- Molecule 1: conserved hypothetical protein TM1225



- Molecule 1: conserved hypothetical protein TM1225

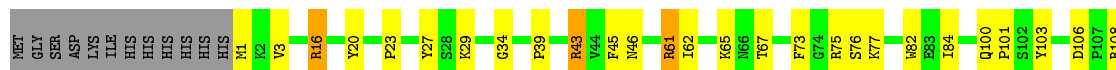






- Molecule 1: conserved hypothetical protein TM1225

Chain E: 78% 17% ..



- Molecule 1: conserved hypothetical protein TM1225

Chain F: 84% 11% ..



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	86.87Å 100.91Å 253.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	82.17 – 2.10 82.17 – 2.10	Depositor EDS
% Data completeness (in resolution range)	91.3 (82.17-2.10) 91.1 (82.17-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.79 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.2.0001	Depositor
R, $R_{free}$	0.162 , 0.204 0.162 , 0.204	Depositor DCC
$R_{free}$ test set	6056 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.1	Xtriage
Anisotropy	0.348	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 119239 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	16823	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.63	0/2738	0.80	6/3727 (0.2%)
1	B	0.70	0/2728	0.84	7/3715 (0.2%)
1	C	0.65	0/2728	0.83	8/3715 (0.2%)
1	D	0.73	0/2728	0.88	9/3715 (0.2%)
1	E	0.65	0/2720	0.81	10/3708 (0.3%)
1	F	0.77	0/2728	0.86	7/3715 (0.2%)
All	All	0.69	0/16370	0.84	47/22295 (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	D	0	1

There are no bond length outliers.

All (47) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	43	ARG	NE-CZ-NH2	-11.26	114.67	120.30
1	F	43	ARG	NE-CZ-NH2	-10.99	114.81	120.30
1	B	43	ARG	NE-CZ-NH1	9.95	125.27	120.30
1	E	43	ARG	NE-CZ-NH2	-9.18	115.71	120.30
1	B	43	ARG	NE-CZ-NH2	-9.07	115.76	120.30
1	F	43	ARG	NE-CZ-NH1	8.43	124.51	120.30
1	C	43	ARG	NE-CZ-NH2	-8.05	116.28	120.30
1	D	20	TYR	C-N-CA	8.02	141.76	121.70
1	A	250	ASP	CB-CG-OD2	7.27	124.84	118.30
1	E	252	ASP	CB-CG-OD2	6.68	124.31	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	108	ARG	NE-CZ-NH1	-6.51	117.05	120.30
1	D	250	ASP	CB-CG-OD2	6.50	124.15	118.30
1	C	43	ARG	NE-CZ-NH1	6.48	123.54	120.30
1	B	123	ASP	CB-CG-OD2	6.44	124.10	118.30
1	D	252	ASP	CB-CG-OD2	6.42	124.08	118.30
1	E	43	ARG	NE-CZ-NH1	6.38	123.49	120.30
1	D	43	ARG	NE-CZ-NH1	6.35	123.47	120.30
1	F	197	ARG	NE-CZ-NH2	-5.88	117.36	120.30
1	D	173	ASP	CB-CA-C	-5.78	98.84	110.40
1	F	123	ASP	CB-CG-OD2	5.78	123.50	118.30
1	B	315	ASP	CB-CG-OD2	5.69	123.42	118.30
1	A	106	ASP	CB-CG-OD2	5.66	123.40	118.30
1	E	123	ASP	CB-CG-OD2	5.63	123.36	118.30
1	C	290	ASP	CB-CG-OD2	5.62	123.36	118.30
1	C	252	ASP	CB-CG-OD2	5.58	123.32	118.30
1	A	43	ARG	NE-CZ-NH2	-5.58	117.51	120.30
1	E	250	ASP	CB-CG-OD2	5.55	123.29	118.30
1	C	173	ASP	CB-CA-C	-5.49	99.42	110.40
1	B	253	ASP	CB-CG-OD2	5.44	123.19	118.30
1	C	114	ASP	CB-CG-OD2	5.42	123.17	118.30
1	B	250	ASP	CB-CG-OD2	5.39	123.16	118.30
1	A	292	ASP	CB-CG-OD2	5.37	123.13	118.30
1	C	250	ASP	CB-CG-OD2	5.37	123.13	118.30
1	A	123	ASP	CB-CG-OD2	5.32	123.09	118.30
1	A	173	ASP	CB-CA-C	-5.28	99.85	110.40
1	E	315	ASP	CB-CG-OD2	5.24	123.02	118.30
1	F	200	LEU	CB-CG-CD2	-5.21	102.14	111.00
1	F	78	ASP	CB-CG-OD2	5.18	122.96	118.30
1	E	189	ASP	CB-CG-OD2	5.17	122.95	118.30
1	E	136	ASP	CB-CG-OD2	5.12	122.91	118.30
1	F	93	ASP	CB-CG-OD2	5.12	122.91	118.30
1	D	315	ASP	CB-CG-OD2	5.11	122.90	118.30
1	E	173	ASP	CB-CA-C	-5.04	100.33	110.40
1	B	124	ASP	CB-CG-OD2	5.02	122.82	118.30
1	D	123	ASP	CB-CG-OD2	5.02	122.81	118.30
1	C	78	ASP	CB-CG-OD2	5.01	122.81	118.30
1	E	61	ARG	NE-CZ-NH1	-5.00	117.80	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	20	TYR	Peptide

## 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	2652	0	2536	36	0
1	B	2642	0	2529	25	0
1	C	2642	0	2529	36	0
1	D	2642	0	2529	35	0
1	E	2633	0	2520	34	0
1	F	2642	0	2529	22	0
2	A	8	0	12	0	0
2	D	8	0	12	0	0
3	A	144	0	0	1	0
3	B	154	0	0	4	0
3	C	142	0	0	4	0
3	D	184	0	0	7	0
3	E	128	0	0	2	0
3	F	202	0	0	3	0
All	All	16823	0	15196	183	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 6.

All (183) 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:1:MSE:HE2	1:E:222:THR:HB	1.54	0.90
1:E:1:MSE:HE1	1:E:318:VAL:HG21	1.54	0.89
1:D:106:ASP:OD1	1:D:108:ARG:NH1	2.12	0.82
1:B:262:ARG:NH1	1:B:326:MSE:HG2	1.98	0.77
1:C:1:MSE:HE3	1:C:224:GLU:HG2	1.65	0.77
1:D:3:VAL:HG21	1:D:314:ILE:HG21	1.71	0.71
1:F:105:TYR:OH	1:F:152:ARG:NH2	2.26	0.68
1:D:173:ASP:HB2	1:D:177:THR:HG23	1.75	0.68
1:F:189:ASP:O	1:F:190:MSE:HB2	1.93	0.67

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:182:ILE:HD12	1:D:213:ILE:HD12	1.77	0.67
1:E:3:VAL:HG21	1:E:314:ILE:HG21	1.74	0.66
1:D:39:PRO:HG2	1:F:95:ASN:O	1.95	0.66
1:C:221:GLU:HB2	1:C:226:TRP:CZ3	2.31	0.65
1:E:1:MSE:HE2	1:E:222:THR:CB	2.25	0.64
1:C:3:VAL:HG21	1:C:314:ILE:HG21	1.78	0.64
1:B:3:VAL:HG21	1:B:314:ILE:HG21	1.80	0.63
1:B:211:LEU:HD22	1:B:240:TYR:CE1	2.34	0.63
1:B:48:ALA:HB1	1:B:109:VAL:HG23	1.82	0.62
1:B:285:CYS:SG	3:B:475:HOH:O	2.55	0.61
1:D:51:PRO:HD3	1:D:288:LEU:CD2	2.30	0.61
1:D:51:PRO:HD3	1:D:288:LEU:HD22	1.83	0.61
1:E:106:ASP:OD1	1:E:108:ARG:NH1	2.32	0.61
1:C:48:ALA:HB1	1:C:109:VAL:HG23	1.82	0.61
1:C:211:LEU:HD22	1:C:240:TYR:CE1	2.36	0.61
1:C:211:LEU:HD22	1:C:240:TYR:HE1	1.66	0.60
1:A:189:ASP:O	1:A:190:MSE:HB2	2.02	0.60
1:B:48:ALA:CB	1:B:109:VAL:HG23	2.31	0.60
1:C:48:ALA:CB	1:C:109:VAL:HG23	2.31	0.60
1:B:43:ARG:HD3	3:B:399:HOH:O	2.02	0.60
1:E:43:ARG:HD3	3:E:409:HOH:O	2.03	0.59
1:D:117:TYR:CD1	1:D:190:MSE:HE2	2.37	0.59
1:C:306:HIS:HE1	3:C:409:HOH:O	1.85	0.58
1:E:103:TYR:CE2	1:E:123:ASP:HB2	2.38	0.58
1:F:1:MSE:HE3	1:F:222:THR:HB	1.85	0.58
1:E:200:LEU:HD11	1:E:260:ARG:HD2	1.86	0.58
1:C:200:LEU:HD11	1:C:260:ARG:HD2	1.86	0.57
1:C:103:TYR:CE2	1:C:123:ASP:HB2	2.39	0.57
1:D:86:PRO:HG3	1:F:94:VAL:O	2.04	0.57
1:B:152:ARG:HD3	3:B:411:HOH:O	2.04	0.56
1:A:3:VAL:HG21	1:A:314:ILE:HG21	1.88	0.56
1:D:323:ARG:HD2	3:D:455:HOH:O	2.05	0.56
1:C:6:GLU:OE1	1:C:51:PRO:HG2	2.07	0.55
1:F:43:ARG:HD3	3:F:377:HOH:O	2.07	0.55
1:D:103:TYR:CE2	1:D:123:ASP:HB2	2.43	0.54
1:F:306:HIS:HE1	3:F:350:HOH:O	1.89	0.54
1:F:200:LEU:HD13	1:F:257:VAL:HG21	1.89	0.54
1:B:106:ASP:OD1	1:B:108:ARG:NH1	2.39	0.54
1:A:51:PRO:HD3	1:A:288:LEU:CD2	2.38	0.54
1:D:1:MSE:HE2	1:D:222:THR:HB	1.88	0.54
1:A:117:TYR:CD1	1:A:190:MSE:HE2	2.43	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:160:LYS:HE2	1:F:165:TYR:CZ	2.43	0.54
1:B:262:ARG:HH12	1:B:326:MSE:HG2	1.71	0.53
1:A:107:PRO:O	1:A:108:ARG:HD2	2.08	0.53
1:D:1:MSE:CE	1:D:222:THR:HB	2.39	0.53
1:A:106:ASP:OD1	1:A:108:ARG:NH1	2.42	0.53
1:F:114:ASP:HB2	3:F:397:HOH:O	2.08	0.52
1:B:164:LYS:HE2	1:B:186:GLU:OE2	2.09	0.52
1:D:33:ILE:HD11	3:D:502:HOH:O	2.08	0.52
1:A:211:LEU:HD22	1:A:240:TYR:CE1	2.44	0.52
1:A:106:ASP:CG	1:A:108:ARG:HH12	2.13	0.52
1:D:306:HIS:HE1	3:D:387:HOH:O	1.92	0.52
1:D:43:ARG:HD3	3:D:343:HOH:O	2.08	0.52
1:D:117:TYR:CE1	1:D:190:MSE:HE2	2.45	0.51
1:A:46:ASN:HB2	1:A:61:ARG:HB3	1.92	0.51
1:B:103:TYR:CE2	1:B:123:ASP:HB2	2.44	0.51
1:F:45:PHE:CD2	1:F:304:ASP:HA	2.46	0.51
1:F:103:TYR:CE2	1:F:123:ASP:HB2	2.45	0.51
1:A:109:VAL:HG22	1:A:118:ILE:HG12	1.93	0.50
1:C:211:LEU:CD2	1:C:240:TYR:HE1	2.25	0.50
1:E:67:THR:HG22	1:E:67:THR:O	2.10	0.50
1:D:164:LYS:HE2	1:D:186:GLU:OE2	2.11	0.50
1:F:173:ASP:HB2	1:F:177:THR:OG1	2.12	0.50
1:E:160:LYS:HE2	1:E:165:TYR:CZ	2.46	0.50
1:E:46:ASN:HB2	1:E:61:ARG:HB3	1.94	0.50
1:E:164:LYS:HE2	1:E:186:GLU:OE2	2.11	0.49
1:C:323:ARG:HD3	3:C:453:HOH:O	2.12	0.49
1:B:51:PRO:HD3	1:B:288:LEU:CD2	2.43	0.49
1:C:43:ARG:NH2	1:C:63:ASP:OD2	2.43	0.49
1:D:221:GLU:HB2	1:D:226:TRP:CZ3	2.48	0.48
1:D:211:LEU:HD22	1:D:240:TYR:CE1	2.47	0.48
1:C:189:ASP:O	1:C:190:MSE:HB2	2.13	0.48
1:C:127:PRO:HD2	1:C:150:PHE:HD2	1.78	0.48
1:E:221:GLU:HB2	1:E:226:TRP:CZ3	2.49	0.48
1:E:260:ARG:O	1:E:325:SER:HA	2.14	0.48
1:A:45:PHE:CD2	1:A:304:ASP:HA	2.49	0.48
1:B:131:VAL:HG11	1:B:190:MSE:CE	2.44	0.48
1:F:3:VAL:HG21	1:F:314:ILE:HG21	1.96	0.47
1:A:62:ILE:HD11	1:A:84:ILE:HG21	1.96	0.47
1:E:189:ASP:O	1:E:190:MSE:HB2	2.15	0.47
1:A:260:ARG:O	1:A:325:SER:HA	2.15	0.47
1:C:131:VAL:HG11	1:C:190:MSE:CE	2.45	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:75:ARG:NH1	3:D:467:HOH:O	2.33	0.47
1:C:1:MSE:HE2	1:C:222:THR:HB	1.96	0.47
1:E:211:LEU:HD22	1:E:240:TYR:CE1	2.50	0.47
1:A:43:ARG:HD3	3:A:391:HOH:O	2.15	0.47
1:D:189:ASP:O	1:D:190:MSE:HB2	2.15	0.47
1:C:226:TRP:HB2	1:C:249:LEU:HB2	1.96	0.46
1:A:103:TYR:C	1:A:103:TYR:CD1	2.88	0.46
1:C:152:ARG:HD3	3:C:419:HOH:O	2.13	0.46
1:E:1:MSE:CE	1:E:222:THR:CB	2.93	0.46
1:F:221:GLU:HB2	1:F:226:TRP:CZ3	2.50	0.46
1:B:200:LEU:HD11	1:B:260:ARG:HD2	1.97	0.46
1:E:62:ILE:HD11	1:E:84:ILE:HG21	1.97	0.46
1:A:67:THR:O	1:A:67:THR:HG22	2.16	0.46
1:D:323:ARG:HD3	3:D:510:HOH:O	2.15	0.46
1:D:110:VAL:HG11	1:D:156:LEU:HB2	1.98	0.46
1:C:27:TYR:CZ	1:C:29:LYS:HB2	2.50	0.46
1:C:160:LYS:HE2	1:C:165:TYR:CZ	2.51	0.46
1:D:39:PRO:O	1:D:65:LYS:HE2	2.15	0.46
1:A:127:PRO:HD2	1:A:150:PHE:HD2	1.81	0.45
1:A:61:ARG:HD2	1:A:105:TYR:HB2	1.98	0.45
1:C:143:LEU:HB3	1:C:144:PRO:HD2	1.98	0.45
1:A:39:PRO:O	1:A:65:LYS:HE2	2.17	0.45
1:A:1:MSE:HE1	1:A:314:ILE:HG12	1.99	0.45
1:A:211:LEU:CD2	1:A:240:TYR:CE1	3.00	0.45
1:B:189:ASP:O	1:B:190:MSE:HB2	2.16	0.45
1:D:62:ILE:HD11	1:D:84:ILE:HG21	1.98	0.45
1:F:260:ARG:O	1:F:325:SER:HA	2.17	0.45
1:A:200:LEU:HD23	1:A:213:ILE:HD12	1.98	0.45
1:A:221:GLU:HB2	1:A:226:TRP:CZ3	2.52	0.45
1:E:161:ILE:HG22	1:E:162:ASN:ND2	2.32	0.45
1:A:160:LYS:HE2	1:A:165:TYR:CZ	2.52	0.45
1:F:173:ASP:HB3	1:F:175:GLY:H	1.82	0.44
1:F:211:LEU:HD22	1:F:240:TYR:CE1	2.52	0.44
1:C:211:LEU:CD2	1:C:240:TYR:CE1	3.00	0.44
1:E:45:PHE:CD2	1:E:304:ASP:HA	2.53	0.44
1:A:51:PRO:HD3	1:A:288:LEU:HD22	2.00	0.44
1:A:223:SER:HA	3:D:506:HOH:O	2.17	0.44
1:A:2:LYS:HB2	1:D:325:SER:O	2.19	0.43
1:B:202:ARG:HD3	1:B:211:LEU:O	2.19	0.43
1:E:39:PRO:O	1:E:65:LYS:HE2	2.18	0.43
1:D:190:MSE:HB3	1:D:190:MSE:HE3	1.90	0.43

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:23:PRO:HG3	1:E:320:PHE:CD1	2.54	0.43
1:B:306:HIS:HE1	3:B:366:HOH:O	1.99	0.43
1:B:106:ASP:N	1:B:107:PRO:HD3	2.34	0.43
1:D:1:MSE:HE3	1:D:1:MSE:HB2	1.79	0.43
1:E:1:MSE:CE	1:E:318:VAL:HG21	2.37	0.43
1:A:103:TYR:CE2	1:A:123:ASP:HB2	2.54	0.43
1:C:173:ASP:HB3	1:C:175:GLY:H	1.84	0.42
1:F:112:ILE:HD12	1:F:117:TYR:CD2	2.53	0.42
1:B:45:PHE:CD2	1:B:304:ASP:HA	2.54	0.42
1:B:131:VAL:HG11	1:B:190:MSE:HE3	2.00	0.42
1:B:39:PRO:O	1:B:65:LYS:HE2	2.19	0.42
1:E:27:TYR:CE2	1:E:29:LYS:HB2	2.55	0.42
1:C:203:SER:O	1:C:210:ASN:HA	2.20	0.42
1:E:73:PHE:HE2	1:E:75:ARG:CZ	2.33	0.42
1:C:306:HIS:HD2	3:C:400:HOH:O	2.01	0.42
1:A:173:ASP:HB3	1:A:175:GLY:H	1.84	0.42
1:C:131:VAL:HG11	1:C:190:MSE:HE3	2.01	0.42
1:E:211:LEU:HD22	1:E:240:TYR:HE1	1.83	0.42
1:D:143:LEU:HB3	1:D:144:PRO:HD2	2.01	0.42
1:F:177:THR:HA	1:F:178:PRO:HD3	1.94	0.42
1:D:45:PHE:CD2	1:D:304:ASP:HA	2.54	0.42
1:E:203:SER:HB3	1:E:209:GLU:HG3	2.02	0.41
1:D:27:TYR:CZ	1:D:29:LYS:HB2	2.55	0.41
1:C:55:GLU:OE2	1:C:77:LYS:HG2	2.20	0.41
1:D:1:MSE:HE3	1:D:224:GLU:HG2	2.03	0.41
1:A:100:GLN:HA	1:A:101:PRO:HD3	1.95	0.41
1:F:187:SER:HB2	1:F:193:TRP:CE3	2.55	0.41
1:C:1:MSE:CE	1:C:222:THR:HB	2.51	0.41
1:C:13:TRP:CZ3	1:C:311:PHE:HB3	2.55	0.41
1:A:76:SER:HB2	1:A:82:TRP:CD2	2.56	0.41
1:E:154:GLY:HA2	1:E:168:LEU:O	2.21	0.41
1:B:203:SER:HB3	1:B:209:GLU:HG3	2.03	0.41
1:E:16:ARG:HG3	1:E:20:TYR:HB3	2.02	0.41
1:E:306:HIS:HD2	3:E:412:HOH:O	2.03	0.41
1:A:242:TYR:HB2	1:A:281:VAL:HA	2.03	0.41
1:B:260:ARG:O	1:B:325:SER:HA	2.20	0.41
1:C:67:THR:HG22	1:C:67:THR:O	2.21	0.41
1:A:2:LYS:HE3	1:D:326:MSE:OXT	2.21	0.41
1:E:143:LEU:HB3	1:E:144:PRO:HD2	2.02	0.41
1:E:76:SER:HB2	1:E:82:TRP:CD2	2.56	0.41
1:C:48:ALA:HB3	1:C:109:VAL:HG23	2.02	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:30:ASN:HB3	1:C:272:TYR:CZ	2.57	0.40
1:C:76:SER:HB2	1:C:82:TRP:CD2	2.56	0.40
1:C:85:GLU:HA	1:C:86:PRO:HD3	1.85	0.40
1:E:100:GLN:HA	1:E:101:PRO:HD3	1.94	0.40
1:D:323:ARG:HB2	1:D:323:ARG:HE	1.76	0.40
1:E:171:PRO:HB3	1:F:149:PRO:O	2.21	0.40
1:B:211:LEU:CD2	1:B:240:TYR:CE1	3.03	0.40
1:A:106:ASP:N	1:A:107:PRO:CD	2.84	0.40
1:A:225:GLY:HA2	1:A:251:LEU:HG	2.04	0.40
1:A:203:SER:O	1:A:210:ASN:HA	2.22	0.40

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	325/338 (96%)	311 (96%)	13 (4%)	1 (0%)	46	45
1	B	324/338 (96%)	309 (95%)	14 (4%)	1 (0%)	46	45
1	C	324/338 (96%)	307 (95%)	17 (5%)	0	100	100
1	D	324/338 (96%)	309 (95%)	13 (4%)	2 (1%)	30	24
1	E	323/338 (96%)	307 (95%)	15 (5%)	1 (0%)	46	45
1	F	324/338 (96%)	309 (95%)	15 (5%)	0	100	100
All	All	1944/2028 (96%)	1852 (95%)	87 (4%)	5 (0%)	46	45

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	21	THR
1	A	34	GLY

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	34	GLY
1	D	34	GLY
1	E	34	GLY

### 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	284/289 (98%)	281 (99%)	3 (1%)	80	85
1	B	283/289 (98%)	279 (99%)	4 (1%)	74	80
1	C	283/289 (98%)	280 (99%)	3 (1%)	80	85
1	D	283/289 (98%)	280 (99%)	3 (1%)	80	85
1	E	282/289 (98%)	278 (99%)	4 (1%)	74	80
1	F	283/289 (98%)	279 (99%)	4 (1%)	74	80
All	All	1698/1734 (98%)	1677 (99%)	21 (1%)	78	84

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	103	TYR
1	A	174	ASN
1	A	325	SER
1	B	170	ARG
1	B	262	ARG
1	B	314	ILE
1	B	325	SER
1	C	18	GLU
1	C	53	ASN
1	C	323	ARG
1	D	152	ARG
1	D	170	ARG
1	D	173	ASP
1	E	16	ARG
1	E	77	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	319	ASP
1	E	323	ARG
1	F	83	GLU
1	F	195	ASN
1	F	319	ASP
1	F	325	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	174	ASN
1	B	306	HIS
1	C	306	HIS
1	D	306	HIS
1	E	162	ASN
1	E	306	HIS
1	F	306	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 ⓘ

2 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	TRS	A	327	-	7,7,7	0.83	0	9,9,9	0.92	0
2	TRS	D	327	-	7,7,7	0.57	0	9,9,9	0.93	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
2	TRS	A	327	-	-	0/9/9/9	0/0/0/0
2	TRS	D	327	-	-	0/9/9/9	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	322/338 (95%)	-0.27	3 (0%) 85 88	24, 30, 44, 66	0
1	B	321/338 (94%)	-0.35	0 100 100	25, 30, 44, 66	0
1	C	321/338 (94%)	-0.28	1 (0%) 94 95	24, 30, 44, 66	0
1	D	321/338 (94%)	-0.39	0 100 100	24, 30, 44, 66	0
1	E	321/338 (94%)	-0.35	0 100 100	24, 30, 44, 66	0
1	F	321/338 (94%)	-0.34	0 100 100	24, 30, 44, 66	0
All	All	1927/2028 (95%)	-0.33	4 (0%) 95 96	24, 30, 44, 66	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2	LYS	2.7
1	A	114	ASP	2.5
1	A	179	PHE	2.5
1	A	94	VAL	2.1

### 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. 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	TRS	A	327	8/8	0.99	0.08	-1.74	23,28,29,29	0
2	TRS	D	327	8/8	0.98	0.09	-2.36	17,20,23,23	0

## 6.5 Other polymers

There are no such residues in this entry.