



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 02:55 am BST

PDB ID : 5V3P  
Title : Human A20 OTU domain (I325N) with acetamidylated C103  
Authors : Langley, D.B.; Christ, D.; Grey, S.  
Deposited on : 2017-03-07  
Resolution : 2.50 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

<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 : 2.11  
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) : 2.11

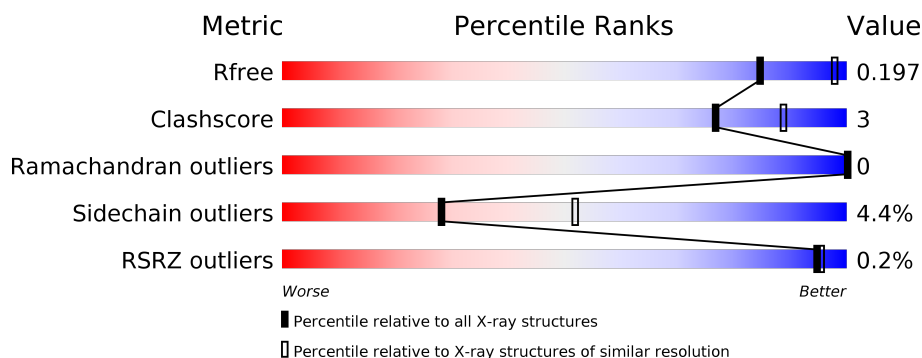
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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  $\leq 5\%$ . 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	366	<div> <div>77%</div> <div>9%</div> <div>14%</div> </div>
1	B	366	<div> <div>71%</div> <div>9%</div> <div>•</div> <div>18%</div> </div>
1	C	366	<div> <div>75%</div> <div>7%</div> <div>•</div> <div>17%</div> </div>
1	D	366	<div> <div>75%</div> <div>8%</div> <div>•</div> <div>16%</div> </div>
1	E	366	<div> <div>69%</div> <div>8%</div> <div>•</div> <div>22%</div> </div>
1	F	366	<div> <div>63%</div> <div>7%</div> <div>•</div> <div>28%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 14786 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 Tumor necrosis factor alpha-induced protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2593	1663	450	466	14			
1	B	300	Total	C	N	O	S	0	0	0
			2505	1611	432	448	14			
1	C	305	Total	C	N	O	S	0	0	0
			2548	1642	440	452	14			
1	D	308	Total	C	N	O	S	0	0	0
			2562	1647	442	459	14			
1	E	286	Total	C	N	O	S	0	0	0
			2385	1535	411	425	14			
1	F	262	Total	C	N	O	S	0	0	0
			2172	1410	376	372	14			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	ASN	ILE	engineered mutation	UNP P21580
B	325	ASN	ILE	engineered mutation	UNP P21580
C	325	ASN	ILE	engineered mutation	UNP P21580
D	325	ASN	ILE	engineered mutation	UNP P21580
E	325	ASN	ILE	engineered mutation	UNP P21580
F	325	ASN	ILE	engineered mutation	UNP P21580

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	O	0	0
			3	3		
2	B	8	Total	O	0	0
			8	8		
2	C	2	Total	O	0	0
			2	2		

*Continued on next page...*

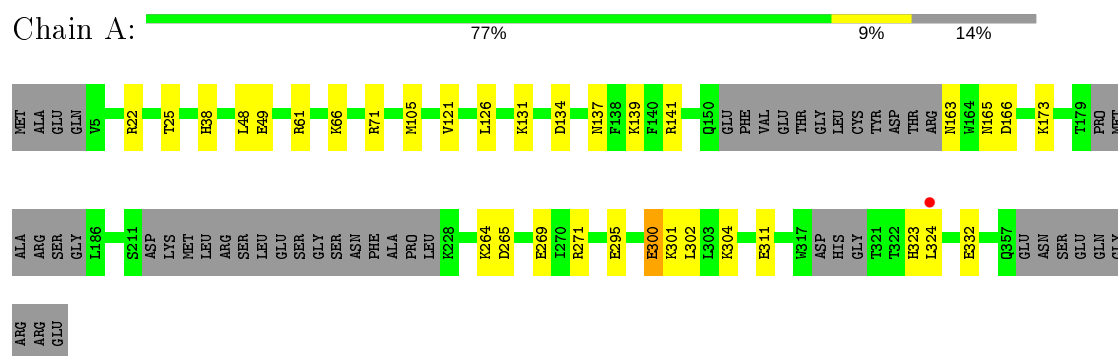
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	D	6	Total	O	0	0
			6	6		
2	F	2	Total	O	0	0
			2	2		

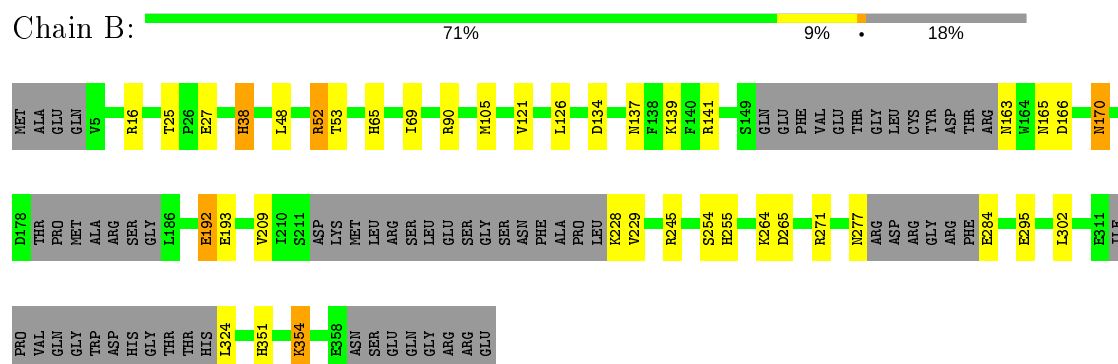
### 3 Residue-property plots [i](#)

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 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.

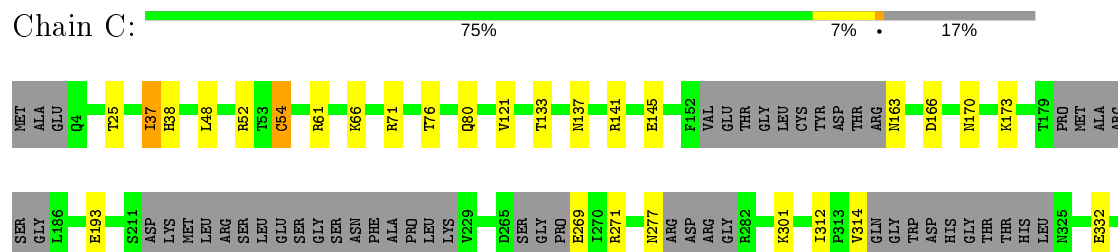
- Molecule 1: Tumor necrosis factor alpha-induced protein 3

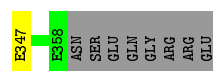


- Molecule 1: Tumor necrosis factor alpha-induced protein 3



- Molecule 1: Tumor necrosis factor alpha-induced protein 3





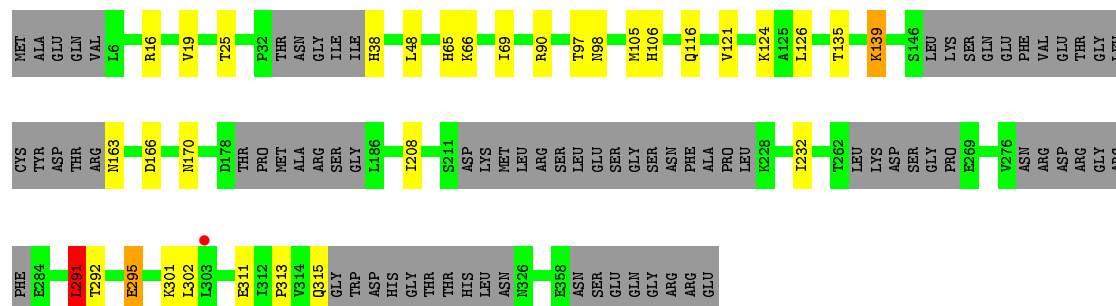
- Molecule 1: Tumor necrosis factor alpha-induced protein 3

Chain D: 75% 8% 16%



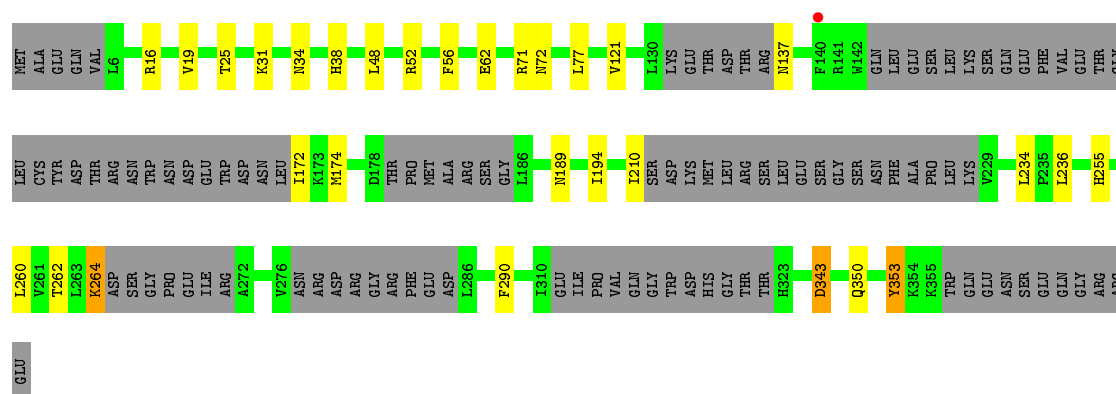
- Molecule 1: Tumor necrosis factor alpha-induced protein 3

Chain E: 69% 8% 22%



- Molecule 1: Tumor necrosis factor alpha-induced protein 3

Chain F: 63% 7% 28%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	81.87Å 81.87Å 297.34Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.47 – 2.50 39.47 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.6 (39.47-2.50) 99.7 (39.47-2.50)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.48 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.161 , 0.200 0.163 , 0.197	Depositor DCC
$R_{free}$ test set	3843 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	42.0	Xtriage
Anisotropy	0.605	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 13.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.37$ , $\langle L^2 \rangle = 0.20$	Xtriage
Estimated twinning fraction	0.457 for -h,-k,l 0.438 for h,-h-k,-l 0.437 for -k,-h,-l	Xtriage
Reported twinning fraction	0.263 for H, K, L 0.214 for -K, -H, -L 0.306 for -h,-k,l 0.217 for K, H, -L	Depositor
Outliers	0 of 76879 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	14786	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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: YCM

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	0.60	0/2640	0.77	1/3572 (0.0%)
1	B	0.62	0/2549	0.80	5/3441 (0.1%)
1	C	0.62	1/2593 (0.0%)	0.77	2/3500 (0.1%)
1	D	0.59	0/2606	0.82	6/3520 (0.2%)
1	E	0.58	1/2427 (0.0%)	0.77	3/3278 (0.1%)
1	F	0.55	0/2207	0.78	1/2974 (0.0%)
All	All	0.60	2/15022 (0.0%)	0.79	18/20285 (0.1%)

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
1	E	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	193	GLU	CD-OE2	8.31	1.34	1.25
1	E	295	GLU	CD-OE1	-6.65	1.18	1.25

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	286	LEU	CB-CG-CD1	10.47	128.80	111.00
1	E	295	GLU	OE1-CD-OE2	-7.92	113.80	123.30
1	D	245	ARG	NE-CZ-NH1	6.87	123.73	120.30

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	192	GLU	CA-CB-CG	6.35	127.37	113.40
1	B	245	ARG	NE-CZ-NH1	6.30	123.45	120.30
1	D	343	ASP	CB-CG-OD1	-6.21	112.71	118.30
1	A	22	ARG	NE-CZ-NH2	-6.16	117.22	120.30
1	E	291	LEU	CA-CB-CG	5.63	128.25	115.30
1	E	139	LYS	CA-CB-CG	5.51	125.53	113.40
1	B	134	ASP	CB-CG-OD2	-5.47	113.37	118.30
1	F	343	ASP	CB-CG-OD1	-5.42	113.42	118.30
1	D	298	MET	CG-SD-CE	5.30	108.68	100.20
1	D	62	GLU	CA-CB-CG	5.29	125.03	113.40
1	D	355	LYS	CD-CE-NZ	5.28	123.85	111.70
1	C	145	GLU	OE1-CD-OE2	-5.21	117.05	123.30
1	C	193	GLU	OE1-CD-OE2	-5.20	117.06	123.30
1	B	271	ARG	NE-CZ-NH1	5.02	122.81	120.30
1	B	52	ARG	NE-CZ-NH2	-5.01	117.80	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	313	PRO	Peptide
1	E	313	PRO	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	2593	0	2564	12	0
1	B	2505	0	2514	22	0
1	C	2548	0	2538	10	1
1	D	2562	0	2561	14	0
1	E	2385	0	2354	17	0
1	F	2172	0	2209	14	1
2	A	3	0	0	0	0
2	B	8	0	0	0	0
2	C	2	0	0	0	0
2	D	6	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	2	0	0	0	0
All	All	14786	0	14740	88	1

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 3.

All (88) 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:174:MET:SD	1:F:189:ASN:OD1	2.25	0.94
1:E:291:LEU:HA	1:E:295:GLU:OE2	1.70	0.91
1:D:52:ARG:HG3	1:D:264:LYS:O	1.79	0.82
1:B:351:HIS:O	1:B:354:LYS:HG3	1.81	0.80
1:F:174:MET:SD	1:F:189:ASN:CG	2.61	0.79
1:D:137:ASN:OD1	1:D:141:ARG:NH1	2.22	0.73
1:C:137:ASN:OD1	1:C:141:ARG:NH1	2.22	0.73
1:C:25:THR:HG22	1:C:121:VAL:HG11	1.72	0.72
1:A:137:ASN:OD1	1:A:141:ARG:NH1	2.22	0.71
1:D:25:THR:HG22	1:D:121:VAL:HG11	1.72	0.71
1:A:49:GLU:OE1	1:A:264:LYS:HD2	1.91	0.71
1:B:137:ASN:OD1	1:B:141:ARG:NH1	2.22	0.71
1:F:25:THR:HG22	1:F:121:VAL:HG11	1.71	0.71
1:A:25:THR:HG22	1:A:121:VAL:HG11	1.71	0.71
1:B:25:THR:HG22	1:B:121:VAL:HG11	1.73	0.69
1:E:25:THR:HG22	1:E:121:VAL:HG11	1.72	0.69
1:F:16:ARG:O	1:F:19:VAL:HG12	1.92	0.69
1:E:291:LEU:HG	1:E:295:GLU:HG3	1.72	0.69
1:E:135:THR:O	1:E:139:LYS:HG3	1.96	0.66
1:F:52:ARG:CG	1:F:264:LYS:O	2.51	0.58
1:B:52:ARG:CG	1:B:264:LYS:O	2.52	0.58
1:A:311:GLU:HB3	1:A:324:LEU:HB2	1.86	0.57
1:B:255:HIS:CG	1:B:255:HIS:O	2.57	0.57
1:B:65:HIS:CD2	1:B:69:ILE:HD12	2.39	0.57
1:F:172:ILE:N	1:F:172:ILE:HD12	2.19	0.57
1:D:52:ARG:CG	1:D:264:LYS:O	2.52	0.57
1:E:291:LEU:HD12	1:E:295:GLU:OE2	2.04	0.57
1:D:80:GLN:HE22	1:D:212:ASP:HB3	1.71	0.56
1:F:52:ARG:HG3	1:F:264:LYS:O	2.05	0.56
1:B:38:HIS:HD1	1:B:38:HIS:C	2.08	0.56
1:E:208:ILE:HG23	1:E:232:ILE:HD13	1.87	0.56
1:D:275:LEU:C	1:D:286:LEU:HD13	2.27	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:139:LYS:NZ	1:D:165:ASN:OD1	2.36	0.54
1:F:210:ILE:HG12	1:F:260:LEU:CD1	2.37	0.54
1:B:52:ARG:HG2	1:B:264:LYS:O	2.08	0.53
1:B:170:ASN:ND2	1:B:170:ASN:C	2.63	0.52
1:D:312:ILE:HD12	1:D:313:PRO:HD2	1.91	0.52
1:F:77:LEU:HD13	1:F:210:ILE:HD11	1.92	0.51
1:B:192:GLU:HG2	1:B:193:GLU:OE1	2.11	0.51
1:E:16:ARG:O	1:E:19:VAL:HG22	2.10	0.51
1:B:139:LYS:NZ	1:B:165:ASN:OD1	2.36	0.51
1:C:170:ASN:O	1:C:173:LYS:HG2	2.10	0.50
1:E:291:LEU:CG	1:E:295:GLU:HG3	2.40	0.50
1:E:292:THR:O	1:E:295:GLU:HG2	2.11	0.50
1:D:73:ILE:O	1:D:76:THR:HG22	2.11	0.50
1:E:292:THR:H	1:E:295:GLU:CD	2.15	0.50
1:B:16:ARG:NE	1:C:347:GLU:OE1	2.45	0.49
1:E:65:HIS:CD2	1:E:69:ILE:HD12	2.47	0.49
1:B:193:GLU:OE2	1:B:228:LYS:HG3	2.12	0.49
1:F:62:GLU:HG2	1:F:71:ARG:HH22	1.78	0.48
1:A:139:LYS:NZ	1:A:165:ASN:OD1	2.38	0.48
1:D:162:ARG:NE	1:D:162:ARG:HA	2.29	0.48
1:A:66:LYS:HA	1:A:71:ARG:NH1	2.29	0.48
1:C:66:LYS:HA	1:C:71:ARG:NH1	2.28	0.48
1:A:271:ARG:CZ	1:A:332:GLU:HG3	2.45	0.47
1:F:234:LEU:HB2	1:F:236:LEU:CD2	2.44	0.47
1:B:192:GLU:CG	1:B:193:GLU:OE1	2.63	0.47
1:B:192:GLU:HG2	1:B:193:GLU:H	1.80	0.46
1:E:98:ASN:OD1	1:E:106:HIS:CD2	2.68	0.46
1:D:73:ILE:HA	1:D:76:THR:HG22	1.96	0.46
1:B:52:ARG:HG3	1:B:264:LYS:O	2.15	0.45
1:C:37:ILE:HA	1:C:133:THR:HG21	1.98	0.45
1:B:170:ASN:C	1:B:170:ASN:HD22	2.19	0.45
1:D:37:ILE:HA	1:D:133:THR:HG21	1.99	0.45
1:E:295:GLU:HB2	1:E:302:LEU:HD11	1.97	0.45
1:B:53:THR:HG21	1:B:65:HIS:NE2	2.32	0.45
1:F:52:ARG:HG2	1:F:264:LYS:O	2.17	0.45
1:B:277:ASN:ND2	1:B:284:GLU:OE1	2.48	0.44
1:F:56:PHE:HD1	1:F:353:TYR:CD2	2.36	0.44
1:A:49:GLU:OE1	1:A:264:LYS:CD	2.65	0.44
1:E:291:LEU:HA	1:E:295:GLU:CD	2.37	0.44
1:A:300:GLU:HG2	1:A:301:LYS:HD2	2.00	0.43
1:D:229:VAL:HG21	1:D:252:TYR:CD1	2.53	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:97:THR:HA	1:E:116:GLN:HE22	1.84	0.43
1:B:105:MET:CE	1:B:126:LEU:HD22	2.48	0.43
1:B:295:GLU:HB3	1:B:302:LEU:HD11	2.01	0.42
1:E:105:MET:CE	1:E:126:LEU:HD22	2.49	0.42
1:A:295:GLU:HB3	1:A:302:LEU:HD11	2.02	0.42
1:C:271:ARG:CZ	1:C:332:GLU:HG3	2.50	0.41
1:A:105:MET:CE	1:A:126:LEU:HD22	2.49	0.41
1:E:16:ARG:HA	1:E:19:VAL:HG22	2.02	0.41
1:C:52:ARG:NH2	1:C:54:CYS:HB2	2.35	0.41
1:D:212:ASP:HB2	1:D:227:LEU:HD23	2.03	0.41
1:A:300:GLU:CD	1:A:300:GLU:H	2.24	0.40
1:B:209:VAL:HG11	1:B:229:VAL:HG12	2.04	0.40
1:C:312:ILE:O	1:C:314:VAL:N	2.54	0.40
1:C:66:LYS:HB2	1:C:66:LYS:HE2	1.97	0.40
1:F:194:ILE:HG22	1:F:290:PHE:CZ	2.56	0.40

All (1) 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)
1:C:76:THR:OG1	1:F:34:ASN:ND2[1_455]	1.52	0.68

## 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	305/366 (83%)	293 (96%)	12 (4%)	0	100	100
1	B	287/366 (78%)	280 (98%)	7 (2%)	0	100	100
1	C	290/366 (79%)	279 (96%)	11 (4%)	0	100	100
1	D	293/366 (80%)	283 (97%)	10 (3%)	0	100	100

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	E	269/366 (74%)	261 (97%)	8 (3%)	0	100	100
1	F	245/366 (67%)	238 (97%)	7 (3%)	0	100	100
All	All	1689/2196 (77%)	1634 (97%)	55 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	282/334 (84%)	269 (95%)	13 (5%)	27	50
1	B	278/334 (83%)	267 (96%)	11 (4%)	31	56
1	C	279/334 (84%)	268 (96%)	11 (4%)	32	57
1	D	283/334 (85%)	270 (95%)	13 (5%)	27	50
1	E	260/334 (78%)	248 (95%)	12 (5%)	27	50
1	F	238/334 (71%)	227 (95%)	11 (5%)	27	50
All	All	1620/2004 (81%)	1549 (96%)	71 (4%)	28	52

All (71) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	HIS
1	A	48	LEU
1	A	61	ARG
1	A	131	LYS
1	A	134	ASP
1	A	163	ASN
1	A	166	ASP
1	A	173	LYS
1	A	265	ASP
1	A	269	GLU
1	A	300	GLU
1	A	304	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	323	HIS
1	B	27	GLU
1	B	38	HIS
1	B	48	LEU
1	B	90	ARG
1	B	163	ASN
1	B	166	ASP
1	B	170	ASN
1	B	254	SER
1	B	265	ASP
1	B	324	LEU
1	B	354	LYS
1	C	37	ILE
1	C	38	HIS
1	C	48	LEU
1	C	54	CYS
1	C	61	ARG
1	C	80	GLN
1	C	163	ASN
1	C	166	ASP
1	C	269	GLU
1	C	277	ASN
1	C	301	LYS
1	D	37	ILE
1	D	38	HIS
1	D	48	LEU
1	D	62	GLU
1	D	66	LYS
1	D	163	ASN
1	D	212	ASP
1	D	269	GLU
1	D	298	MET
1	D	300	GLU
1	D	332	GLU
1	D	337	LYS
1	D	343	ASP
1	E	38	HIS
1	E	48	LEU
1	E	66	LYS
1	E	90	ARG
1	E	124	LYS
1	E	163	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	166	ASP
1	E	170	ASN
1	E	291	LEU
1	E	301	LYS
1	E	311	GLU
1	E	315	GLN
1	F	31	LYS
1	F	38	HIS
1	F	48	LEU
1	F	72	ASN
1	F	137	ASN
1	F	255	HIS
1	F	262	THR
1	F	264	LYS
1	F	343	ASP
1	F	350	GLN
1	F	353	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	237	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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	$\# Z  > 2$
1	YCM	A	103	1	7,9,10	0.85	0	4,10,12	0.54	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
1	YCM	E	103	1	7,9,10	1.07	1 (14%)	4,10,12	1.41	1 (25%)
1	YCM	F	103	1	7,9,10	0.89	0	4,10,12	0.48	0
1	YCM	D	103	1	7,9,10	0.84	0	4,10,12	0.48	0
1	YCM	B	103	1	7,9,10	0.92	0	4,10,12	0.45	0
1	YCM	C	103	1	7,9,10	0.92	0	4,10,12	0.61	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
1	YCM	A	103	1	-	1/6/8/10	-
1	YCM	E	103	1	-	2/6/8/10	-
1	YCM	F	103	1	-	1/6/8/10	-
1	YCM	D	103	1	-	1/6/8/10	-
1	YCM	B	103	1	-	1/6/8/10	-
1	YCM	C	103	1	-	1/6/8/10	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	103	YCM	CD-SG	-2.07	1.76	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	103	YCM	OZ1-CE-NZ2	-2.08	116.84	122.50

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	103	YCM	SG-CD-CE-NZ2
1	E	103	YCM	CE-CD-SG-CB
1	E	103	YCM	SG-CD-CE-NZ2
1	F	103	YCM	SG-CD-CE-NZ2
1	D	103	YCM	SG-CD-CE-NZ2
1	B	103	YCM	SG-CD-CE-NZ2
1	C	103	YCM	SG-CD-CE-NZ2



There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

There are no ligands in this entry.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data [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	315/366 (86%)	-0.56	1 (0%) 94 94	21, 49, 77, 96	0
1	B	299/366 (81%)	-0.56	0 100 100	22, 49, 75, 87	0
1	C	304/366 (83%)	-0.53	0 100 100	33, 55, 75, 103	0
1	D	307/366 (83%)	-0.54	1 (0%) 94 94	30, 54, 80, 96	0
1	E	285/366 (77%)	-0.43	1 (0%) 92 93	40, 66, 91, 108	0
1	F	261/366 (71%)	-0.39	1 (0%) 92 93	41, 65, 93, 104	0
All	All	1771/2196 (80%)	-0.51	4 (0%) 95 95	21, 56, 85, 108	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	314	VAL	2.6
1	A	324	LEU	2.3
1	F	140	PHE	2.2
1	E	303	LEU	2.0

### 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
1	YCM	E	103	10/11	0.94	0.12	61,65,78,84	0
1	YCM	A	103	10/11	0.97	0.12	37,43,52,53	0
1	YCM	F	103	10/11	0.97	0.09	57,61,64,66	0
1	YCM	D	103	10/11	0.97	0.12	37,43,66,67	0
1	YCM	B	103	10/11	0.97	0.11	45,50,70,73	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	YCM	C	103	10/11	0.98	0.08	40,48,64,65	0

### 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.