



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 11:42 am BST

PDB ID : 3FH6
Title : Crystal structure of the resting state maltose transporter from E. coli
Authors : Khare, D.; Oldham, M.L.; Orelle, C.; Davidson, A.L.; Chen, J.
Deposited on : 2008-12-08
Resolution : 4.50 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

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:

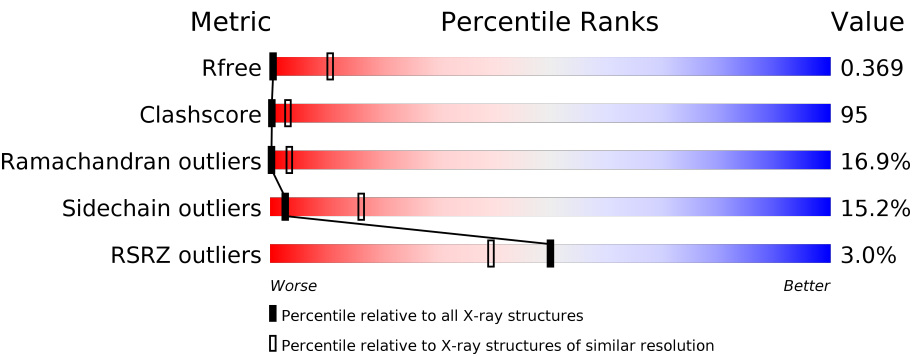
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 4.50 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	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)
RSRZ outliers	127900	1101 (5.30-3.70)

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 ≥ 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	F	480	<div><div>9%</div><div>41%</div><div>14%</div><div>34%</div></div>
1	H	480	<div><div>4%</div><div>9%</div><div>42%</div><div>14%</div><div>34%</div></div>
2	G	296	<div><div>2%</div><div>17%</div><div>52%</div><div>14%</div><div>14%</div></div>
2	I	296	<div><div>2%</div><div>17%</div><div>52%</div><div>14%</div><div>14%</div></div>
3	A	381	<div><div>2%</div><div>12%</div><div>60%</div><div>22%</div><div>4%</div></div>
3	B	381	<div><div>3%</div><div>14%</div><div>61%</div><div>20%</div><div>2%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	381	<div><div></div><div>6%</div><div>13%</div><div>58%</div><div>23%</div><div></div><div></div></div>
3	D	381	<div><div></div><div>2%</div><div>15%</div><div>59%</div><div>21%</div><div></div><div></div></div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20236 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 Maltose transport system permease protein malF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	F	316	Total	C	N	O	S	0	0	0
			2418	1607	378	418	15			
1	H	316	Total	C	N	O	S	0	0	0
			2418	1607	378	418	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	35	MET	-	EXPRESSION TAG	UNP P02916
H	35	MET	-	EXPRESSION TAG	UNP P02916

- Molecule 2 is a protein called Maltose transport system permease protein malG.

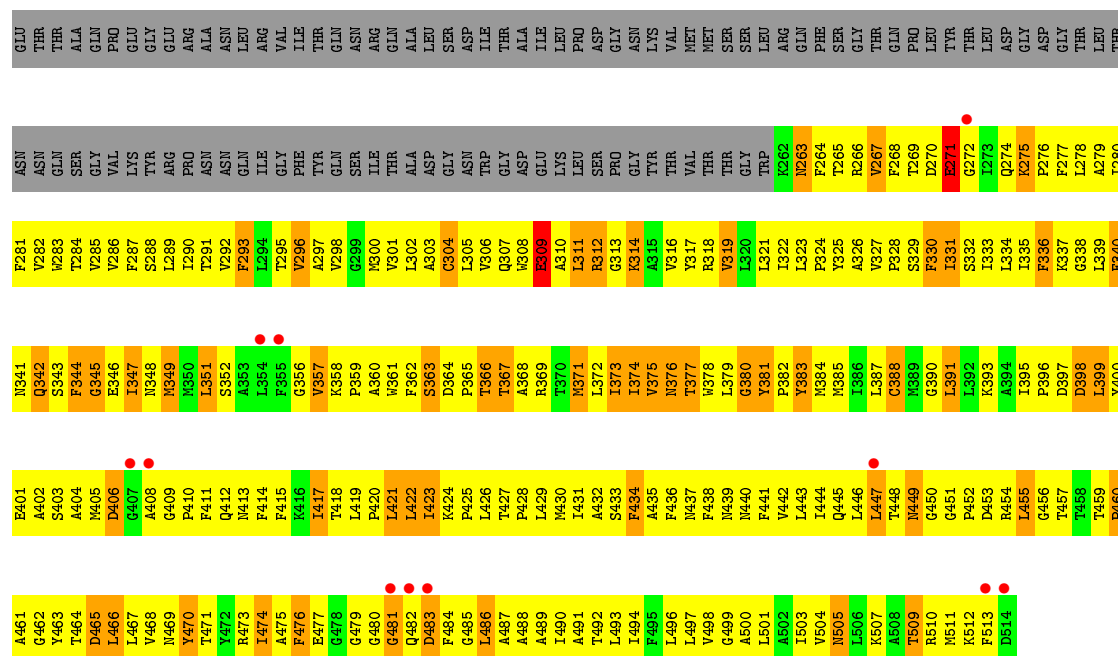
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	254	Total	C	N	O	S	0	0	0
			1942	1308	306	319	9			
2	I	254	Total	C	N	O	S	0	0	0
			1942	1308	306	319	9			

- Molecule 3 is a protein called Maltose/maltodextrin import ATP-binding protein malK.

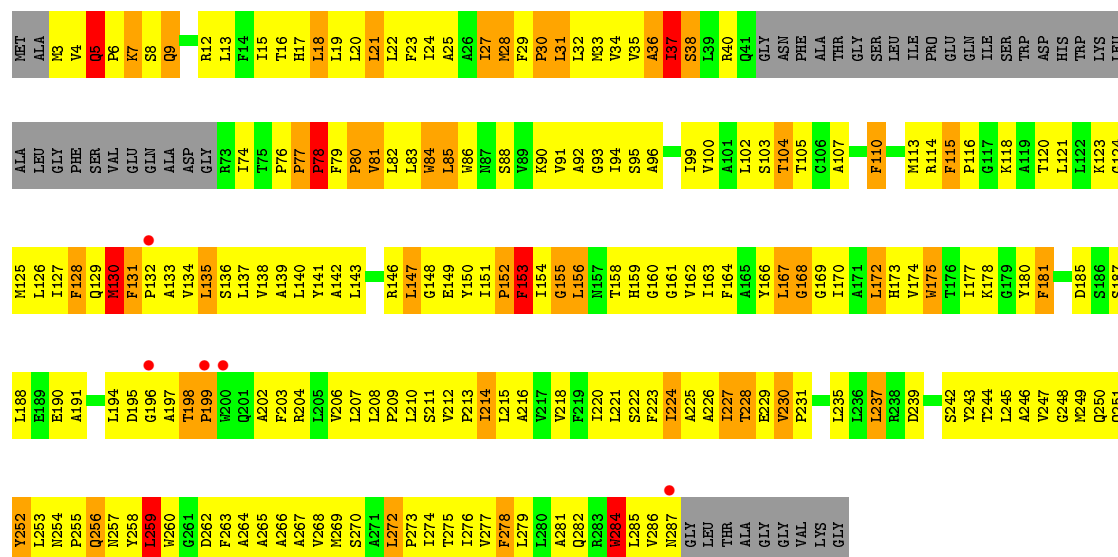
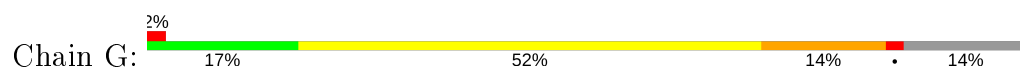
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	371	Total	C	N	O	S	0	0	0
			2876	1819	515	529	13			
3	B	372	Total	C	N	O	S	0	0	0
			2882	1822	516	531	13			
3	C	371	Total	C	N	O	S	0	0	0
			2876	1819	515	529	13			
3	D	372	Total	C	N	O	S	0	0	0
			2882	1822	516	531	13			

There are 40 discrepancies between the modelled and reference sequences:

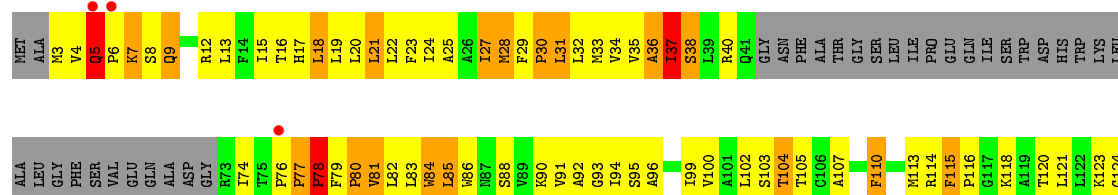
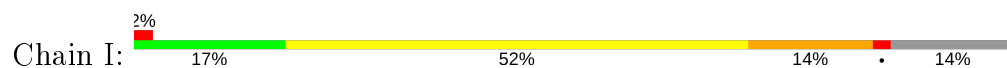
Chain	Residue	Modelled	Actual	Comment	Reference
A	372	ALA	-	EXPRESSION TAG	UNP P68187
A	373	SER	-	EXPRESSION TAG	UNP P68187
A	374	ALA	-	EXPRESSION TAG	UNP P68187
A	375	SER	-	EXPRESSION TAG	UNP P68187
A	376	HIS	-	EXPRESSION TAG	UNP P68187
A	377	HIS	-	EXPRESSION TAG	UNP P68187
A	378	HIS	-	EXPRESSION TAG	UNP P68187
A	379	HIS	-	EXPRESSION TAG	UNP P68187
A	380	HIS	-	EXPRESSION TAG	UNP P68187
A	381	HIS	-	EXPRESSION TAG	UNP P68187
B	372	ALA	-	EXPRESSION TAG	UNP P68187
B	373	SER	-	EXPRESSION TAG	UNP P68187
B	374	ALA	-	EXPRESSION TAG	UNP P68187
B	375	SER	-	EXPRESSION TAG	UNP P68187
B	376	HIS	-	EXPRESSION TAG	UNP P68187
B	377	HIS	-	EXPRESSION TAG	UNP P68187
B	378	HIS	-	EXPRESSION TAG	UNP P68187
B	379	HIS	-	EXPRESSION TAG	UNP P68187
B	380	HIS	-	EXPRESSION TAG	UNP P68187
B	381	HIS	-	EXPRESSION TAG	UNP P68187
C	372	ALA	-	EXPRESSION TAG	UNP P68187
C	373	SER	-	EXPRESSION TAG	UNP P68187
C	374	ALA	-	EXPRESSION TAG	UNP P68187
C	375	SER	-	EXPRESSION TAG	UNP P68187
C	376	HIS	-	EXPRESSION TAG	UNP P68187
C	377	HIS	-	EXPRESSION TAG	UNP P68187
C	378	HIS	-	EXPRESSION TAG	UNP P68187
C	379	HIS	-	EXPRESSION TAG	UNP P68187
C	380	HIS	-	EXPRESSION TAG	UNP P68187
C	381	HIS	-	EXPRESSION TAG	UNP P68187
D	372	ALA	-	EXPRESSION TAG	UNP P68187
D	373	SER	-	EXPRESSION TAG	UNP P68187
D	374	ALA	-	EXPRESSION TAG	UNP P68187
D	375	SER	-	EXPRESSION TAG	UNP P68187
D	376	HIS	-	EXPRESSION TAG	UNP P68187
D	377	HIS	-	EXPRESSION TAG	UNP P68187
D	378	HIS	-	EXPRESSION TAG	UNP P68187
D	379	HIS	-	EXPRESSION TAG	UNP P68187
D	380	HIS	-	EXPRESSION TAG	UNP P68187
D	381	HIS	-	EXPRESSION TAG	UNP P68187

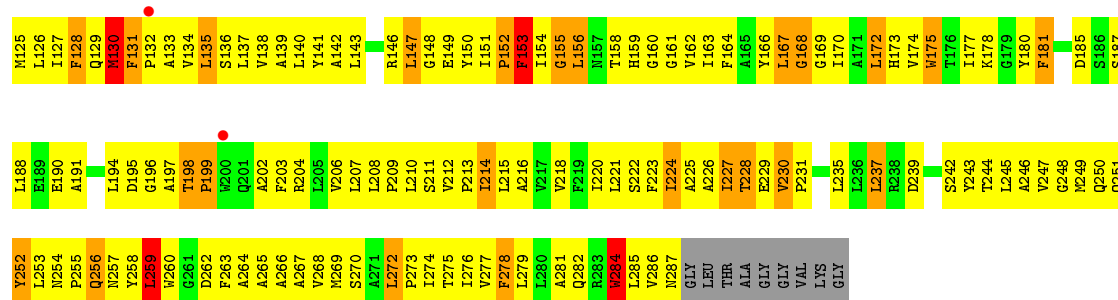


• Molecule 2: Maltose transport system permease protein malG

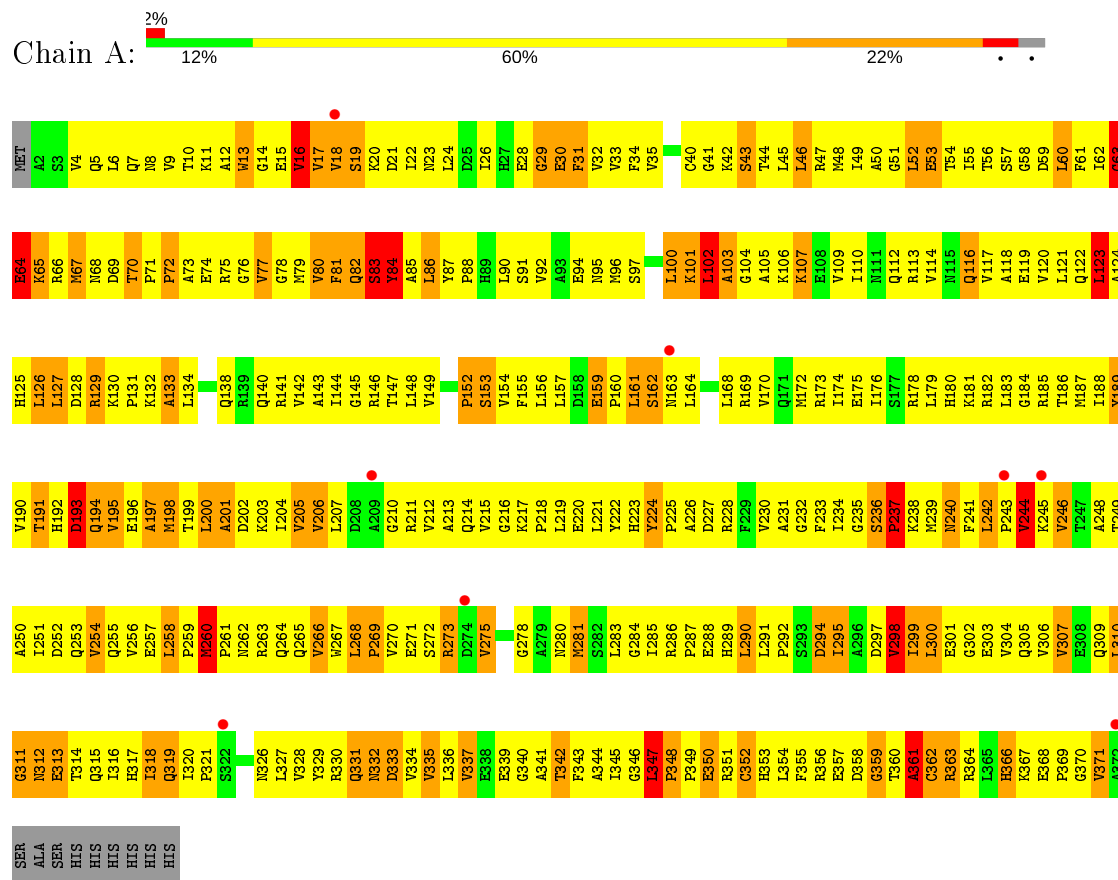


• Molecule 2: Maltose transport system permease protein malG

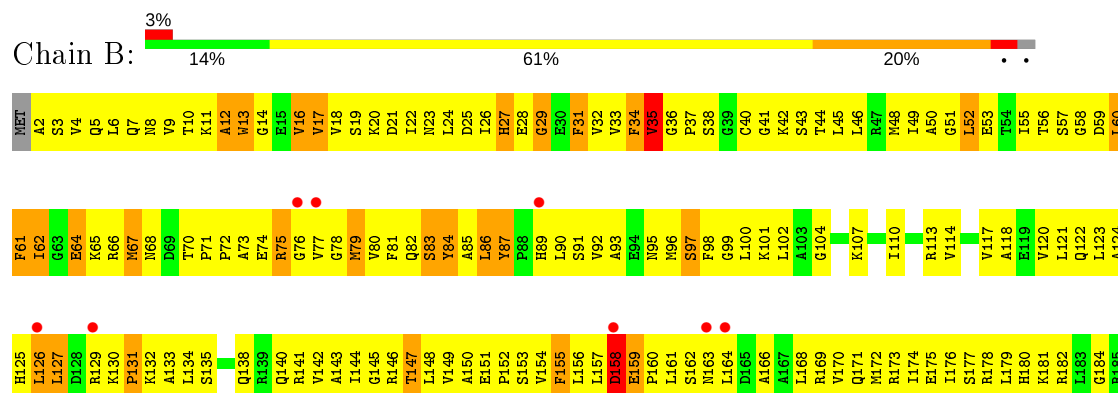


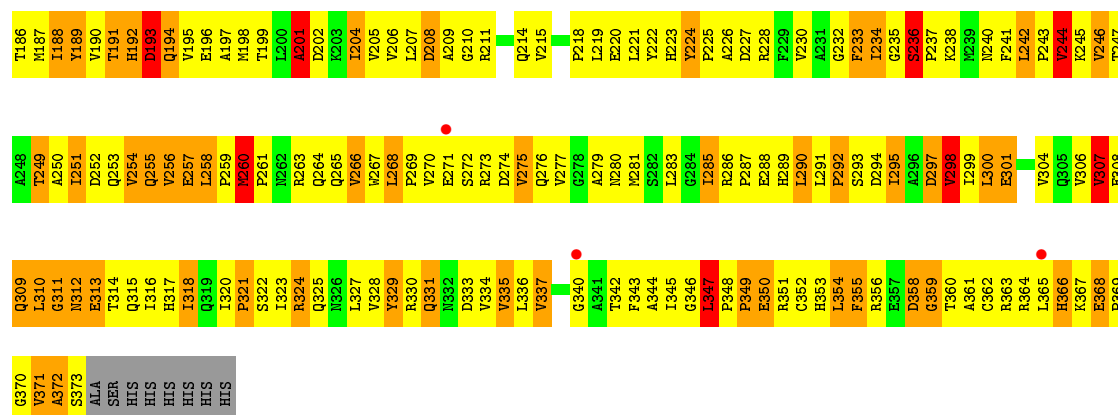


• Molecule 3: Maltose/maltodextrin import ATP-binding protein malK



• Molecule 3: Maltose/maltodextrin import ATP-binding protein malK

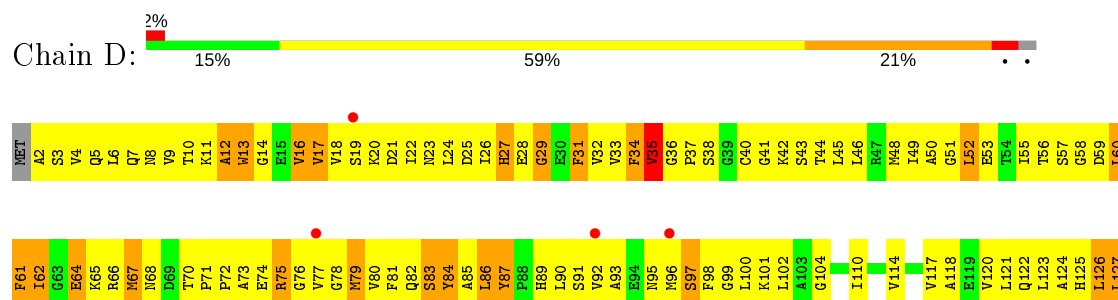




• Molecule 3: Maltose/maltodextrin import ATP-binding protein malK



• Molecule 3: Maltose/maltodextrin import ATP-binding protein malK



HIS	T314	Q253	T191	D128
HIS	Q315	V254	H192	R129
HIS	I316	V255	D193	K130
	H317	V256	Q194	P131
	Q318	E257	V195	K132
	Q319	L258	E196	A133
	I320	P259	A197	L134
	P321	D260	M198	S135
	S322	P261	T199	
	I323	N262	L200	Q138
	R324	R263	A201	R139
	Q325	N264	D202	Q140
		Q265	V203	R141
	V328	V266	I204	V142
	Y329	V267	V205	A143
	R330	L268	V206	I144
	Q331	P269	L207	G145
	N332	V270	D208	R146
	D333	E271	A209	T147
	V334	S272	G210	L148
	V335	R273	R211	V149
	L336	D274	Q214	A150
	V337	V275	V215	E151
		Q276		P152
	T342	Q277		S153
	F343	G278	P218	V154
	A344	A279	I219	F155
	I345	N280	E220	L156
	G346	N281	L221	L157
	L347	S282	V222	D158
	P348	L283	H223	E159
	P349	G284	V224	P160
	E350	I285	P225	L161
	R351	R286	A226	S162
	C352	P287	D227	N163
	H353	E288	R228	L164
	L354	H289	F229	D165
	F355	L290	V230	A166
	R356	L291	A231	A167
	E357	P292	G232	L168
	D358	S293	F233	R169
	G359	D294	I234	V170
	T360	I295	G235	Q171
	A361	A296	S236	M172
	C362	D297	F237	R173
	R363	V298	R238	I174
	R364	I299	N239	E175
	L365	L300	N240	I176
	H366	E301	F241	S177
	K367		L242	R178
	E368	V304	F243	L179
	P369	Q305	V244	H180
	G370	V306	K245	K181
	V371	V307	V246	R182
	A372	E308	T247	
	S373	Q309	I248	T186
	ALA	I310	T249	M187
	SER	G311	A250	I188
	HIS	N312	I251	Y189
	HIS	E313	D252	V190

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	171.10 Å 209.48 Å 438.74 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.50 49.19 – 4.50	Depositor EDS
% Data completeness (in resolution range)	85.2 (50.00-4.50) 85.2 (49.19-4.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 4.45 Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.340 , 0.363 0.354 , 0.369	Depositor DCC
R_{free} test set	2038 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	216.6	Xtriage
Anisotropy	0.633	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.18 , 150.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	20236	wwPDB-VP
Average B, all atoms (Å ²)	300.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.84% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	F	0.51	0/2473	0.85	3/3365 (0.1%)
1	H	0.51	0/2473	0.85	2/3365 (0.1%)
2	G	0.49	0/1992	0.87	5/2724 (0.2%)
2	I	0.50	0/1992	0.87	5/2724 (0.2%)
3	A	0.46	1/2926 (0.0%)	0.90	6/3968 (0.2%)
3	B	0.49	1/2932 (0.0%)	0.92	2/3976 (0.1%)
3	C	0.46	1/2926 (0.0%)	0.90	6/3968 (0.2%)
3	D	0.49	1/2932 (0.0%)	0.92	2/3976 (0.1%)
All	All	0.49	4/20646 (0.0%)	0.89	31/28066 (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
3	A	0	1
3	C	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	159	GLU	CD-OE2	6.81	1.33	1.25
3	A	159	GLU	CD-OE2	6.79	1.33	1.25
3	D	159	GLU	CD-OE2	6.50	1.32	1.25
3	B	159	GLU	CD-OE2	6.49	1.32	1.25

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	259	LEU	N-CA-C	-7.22	91.50	111.00
2	I	259	LEU	N-CA-C	-7.22	91.51	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	196	GLY	N-CA-C	6.62	129.66	113.10
2	I	196	GLY	N-CA-C	6.62	129.66	113.10
1	H	380	GLY	N-CA-C	6.46	129.24	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	84	TYR	Sidechain
3	C	84	TYR	Sidechain

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	F	2418	0	2476	555	0
1	H	2418	0	2476	557	0
2	G	1942	0	2008	362	0
2	I	1942	0	2008	365	0
3	A	2876	0	2942	582	15
3	B	2882	0	2947	557	12
3	C	2876	0	2942	571	9
3	D	2882	0	2947	549	5
All	All	20236	0	20746	3892	31

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 95.

The worst 5 of 3892 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:347:ILE:HG23	1:F:348:ASN:H	1.03	1.15
3:D:157:LEU:HD23	3:D:160:PRO:HG3	1.30	1.13
1:F:275:LYS:H	1:F:276:PRO:HD2	1.07	1.13
3:C:79:MET:HG2	3:C:80:VAL:H	1.14	1.12
1:F:387:LEU:HD21	1:F:429:LEU:HD13	1.29	1.12

The worst 5 of 31 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)
3:A:182:ARG:O	3:B:324:ARG:NH1[2_565]	0.90	1.30
3:C:267:TRP:NE1	3:C:267:TRP:NE1[3_555]	1.15	1.05
3:B:340:GLY:O	3:B:340:GLY:O[2_565]	1.19	1.01
3:B:276:GLN:OE1	3:D:274:ASP:N[2_565]	1.26	0.94
3:A:119:GLU:CG	3:A:119:GLU:CD[2_565]	1.27	0.93

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	F	310/480 (65%)	209 (67%)	61 (20%)	40 (13%)	0	5
1	H	310/480 (65%)	209 (67%)	61 (20%)	40 (13%)	0	5
2	G	250/296 (84%)	179 (72%)	38 (15%)	33 (13%)	0	4
2	I	250/296 (84%)	179 (72%)	38 (15%)	33 (13%)	0	4
3	A	369/381 (97%)	198 (54%)	97 (26%)	74 (20%)	0	2
3	B	370/381 (97%)	203 (55%)	95 (26%)	72 (20%)	0	2
3	C	369/381 (97%)	198 (54%)	96 (26%)	75 (20%)	0	2
3	D	370/381 (97%)	203 (55%)	95 (26%)	72 (20%)	0	2
All	All	2598/3076 (84%)	1578 (61%)	581 (22%)	439 (17%)	0	3

5 of 439 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	90	ALA
1	F	92	THR
1	F	94	TYR
1	F	96	SER
1	F	98	ASN

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	F	253/394 (64%)	201 (79%)	52 (21%)	1	7
1	H	253/394 (64%)	201 (79%)	52 (21%)	1	7
2	G	198/237 (84%)	166 (84%)	32 (16%)	2	15
2	I	198/237 (84%)	166 (84%)	32 (16%)	2	15
3	A	314/323 (97%)	273 (87%)	41 (13%)	4	20
3	B	315/323 (98%)	276 (88%)	39 (12%)	4	22
3	C	314/323 (97%)	273 (87%)	41 (13%)	4	20
3	D	315/323 (98%)	276 (88%)	39 (12%)	4	22
All	All	2160/2554 (85%)	1832 (85%)	328 (15%)	3	16

5 of 328 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	244	VAL
1	H	331	ILE
3	D	158	ASP
3	B	258	LEU
1	H	55	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	264	GLN
1	H	341	ASN
3	D	264	GLN
3	B	276	GLN
3	B	315	GLN

5.3.3 RNA ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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, 95th 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 ²)	Q<0.9
1	F	316/480 (65%)	-0.31	3 (0%) 84 77	187, 310, 422, 493	0
1	H	316/480 (65%)	0.04	19 (6%) 21 18	187, 310, 422, 493	0
2	G	254/296 (85%)	-0.19	5 (1%) 65 56	204, 296, 401, 510	0
2	I	254/296 (85%)	-0.29	5 (1%) 65 56	204, 296, 401, 510	0
3	A	371/381 (97%)	-0.08	8 (2%) 62 52	195, 296, 384, 474	0
3	B	372/381 (97%)	-0.06	11 (2%) 50 39	150, 281, 361, 454	0
3	C	371/381 (97%)	0.15	22 (5%) 22 19	195, 296, 384, 474	0
3	D	372/381 (97%)	-0.15	7 (1%) 66 58	150, 281, 361, 454	0
All	All	2626/3076 (85%)	-0.10	80 (3%) 50 39	150, 294, 398, 510	0

The worst 5 of 80 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	280	ASN	6.5
1	H	56	ILE	6.2
1	H	483	ASP	5.7
3	C	281	MET	5.5
3	C	244	VAL	5.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

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.