wwPDB X-ray Structure Validation Summary Report (i)

Oct 16, 2023 – 12:39 PM EDT

PDB ID	:	1ULQ
Title	:	Crystal structure of tt0182 from Thermus thermophilus HB8
Authors	:	Ago, H.; Hamada, K.; Ida, K.; Kanda, H.; Sugahara, M.; Yamamoto, M.; Kuroishi, C.; Kuramitsu, S.; Yokoyama, S.; Miyano, M.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2003-09-16
Resolution	:	3.00 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

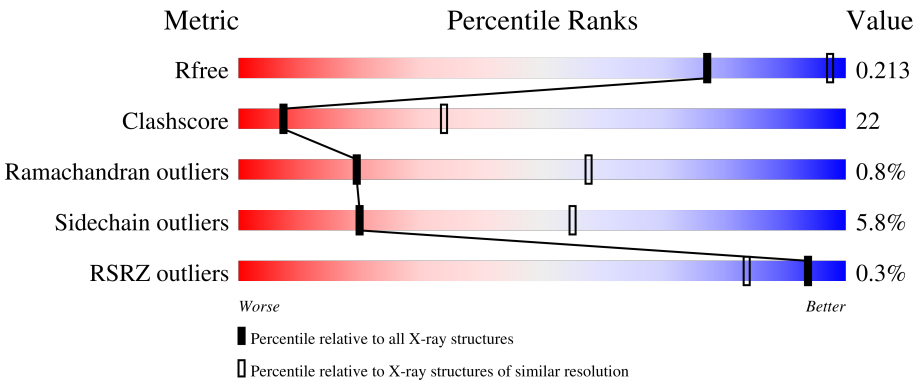
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 of 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 <=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	401	61% 34% • •
1	B	401	60% 35% • •
1	C	401	59% 37% •
1	D	401	60% 36% •
1	E	401	62% 34% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	401	62% 35%
1	G	401	58% 38%
1	H	401	59% 37%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 23907 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 putative acetyl-CoA acetyltransferase.

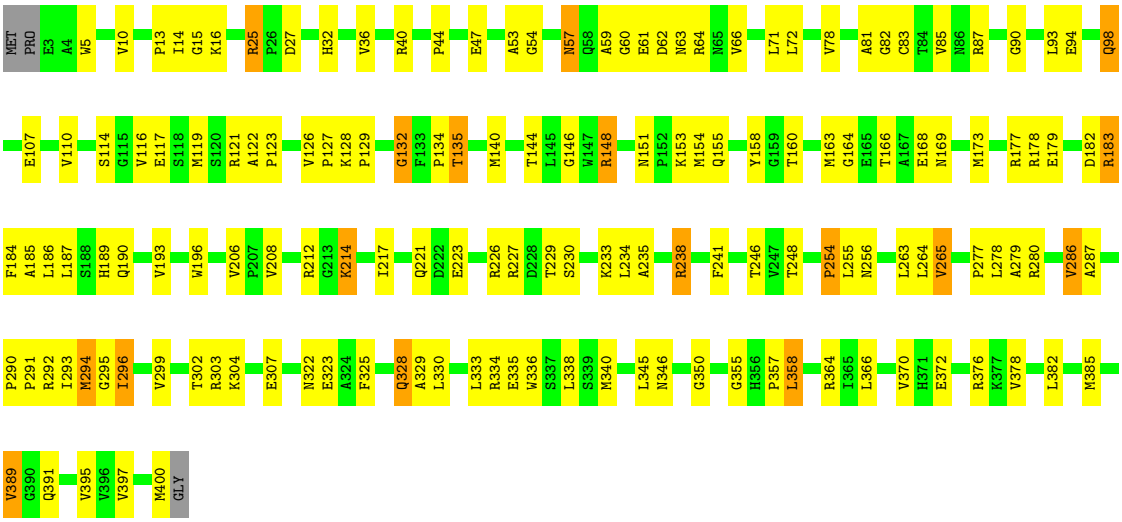
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	398	2984	1867	546	556	15	0	0	0
1	B	398	2984	1867	546	556	15	0	0	0
1	C	399	2991	1872	547	557	15	0	0	0
1	D	399	2991	1872	547	557	15	0	0	0
1	E	399	2991	1872	547	557	15	0	0	0
1	F	398	2984	1867	546	556	15	0	0	0
1	G	399	2991	1872	547	557	15	0	0	0
1	H	399	2991	1872	547	557	15	0	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

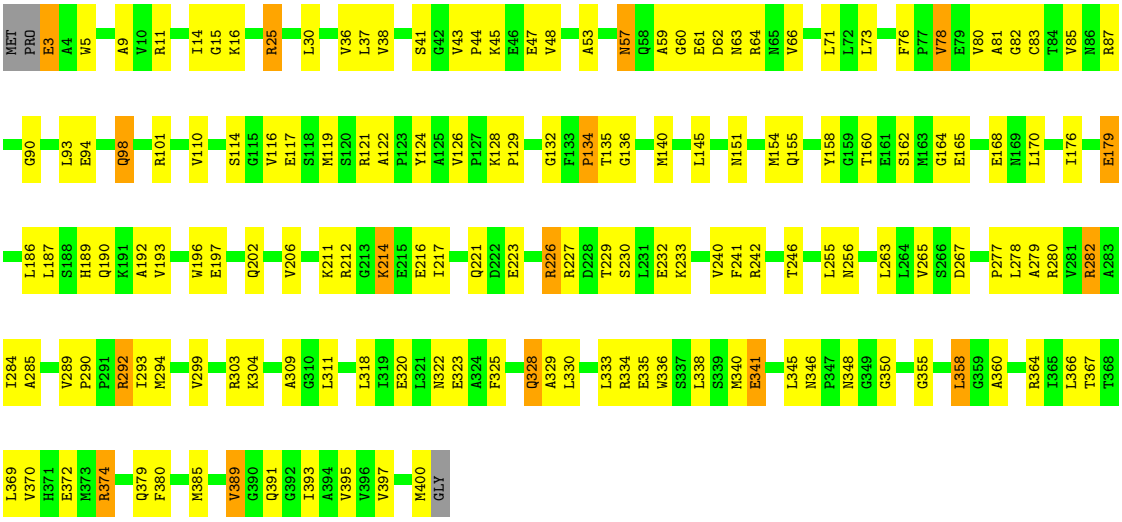
• Molecule 1: putative acetyl-CoA acetyltransferase

Chain A: 61% 34%



• Molecule 1: putative acetyl-CoA acetyltransferase

Chain B: 60% 35%



• Molecule 1: putative acetyl-CoA acetyltransferase

Chain C: 59% 37% •

MET P2 E3 A4 W5 I6 V7 E8 A9 V10 R11 G15 K16 H17 S23 D27 D28 L29 L30 A31 H32 V36 L37 V38 D39 R40 P44 K45 E46 E47 V48 V51 Y52 A53 N57 Q58 A59 D62 N63 R64 N65 V66 A70 L71 L72 F76 P77 A81 G82

C83 R87 L88 C89 G90 S91 E94 A95 V96 A97 Q98 W104 E107 I112 G113 S114 G115 A122 S120 R121 V126 P127 K128 P129 E130 R131 G132 F133 P134 T135 L138 V139 M140 Y141 D142 T143 T144 L145 G146 W147 N151 P152 K153 M154 Q155 Y158 G159 T160

T166 A167 E168 N169 L170 R178 Q181 F184 A185 L186 H189 Q190 V193 E197 V206 K211 R212 G213 K214 E215 E216 I217 E223 R227 E232 K233 L234 T246 V247 T248 S252 N256 V262 L263 L264 V265 L275 R276 P277 L278 A279 R280

I284 A285 V286 V289 P290 P291 R292 I293 M294 G295 I296 G297 P298 V299 R303 K304 E307 R308 A309 G310 L311 L316 N322 E323 A324 F325 A326 A327 Q328 A329 L330 A331 V332 L333 R334 E335 L338 S339 M340 L345 N346 G350 P357 L358 G359 L366 T367 V370

R374 L382 M385 C386 V389 G390 Q391 V397 E398 G399 M400 GLY

• Molecule 1: putative acetyl-CoA acetyltransferase

Chain D: 60% 36% •

MET P2 E3 A4 W5 A9 V10 P13 I14 G15 R25 H32 V36 L37 P44 K45 E46 E47 A53 N57 Q58 A59 G60 E61 D62 N63 R64 N65 V66 A67 R68 L71 L72 L73 V78 E79 V80 A81 G82 C83 R87 L93 A97 Q98 E107

G113 S114 G115 V116 E117 S118 R121 A122 K128 P129 E130 R131 G132 F133 P134 T135 L138 V139 M140 L145 R148 M154 Q155 Y158 G159 T160 E161 S162 M163 G164 E165 T166 N169 E172 M173 Y174 R177 R178 E179 R183 L186 H189 Q190 K191 A192 V193

W196 E197 E204 V205 V206 P207 V208 P209 V210 K211 R212 G213 K214 E215 E216 Q221 R226 R227 D228 T229 S230 L231 L234 R238 F241 R242 T246 V247 T248 N251 S252 S253 P254 L255 N256 A261 V262 L263 L264 V265 S266 Y269 P277 L278 A279 R280 V281 R282 A283

I284 V289 P290 I293 M294 V299 P300 A301 T302 L306 E307 R308 F313 L318 N322 E323 A324 F325 Q328 A329 L330 R334 E335 L338 S339 M340 E341 L345 N346 P347 N348 G349 G350 G355 H356 P357 L358 G359 A360 R364 T367 T368 E372 M373 R374

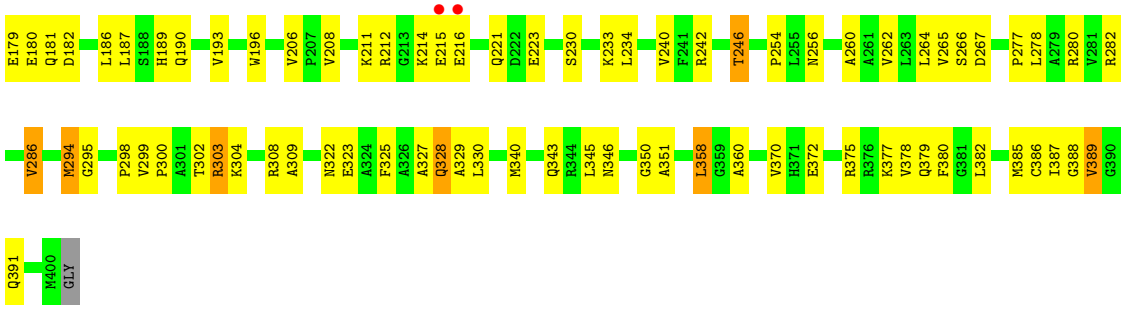
R375 R376 K377 V378 Q379 F380 G381 L382 M385 V389 G390 Q391 G392 I393 V397 E398 G399 M400 GLY

• Molecule 1: putative acetyl-CoA acetyltransferase

Chain E: 62% 34% •

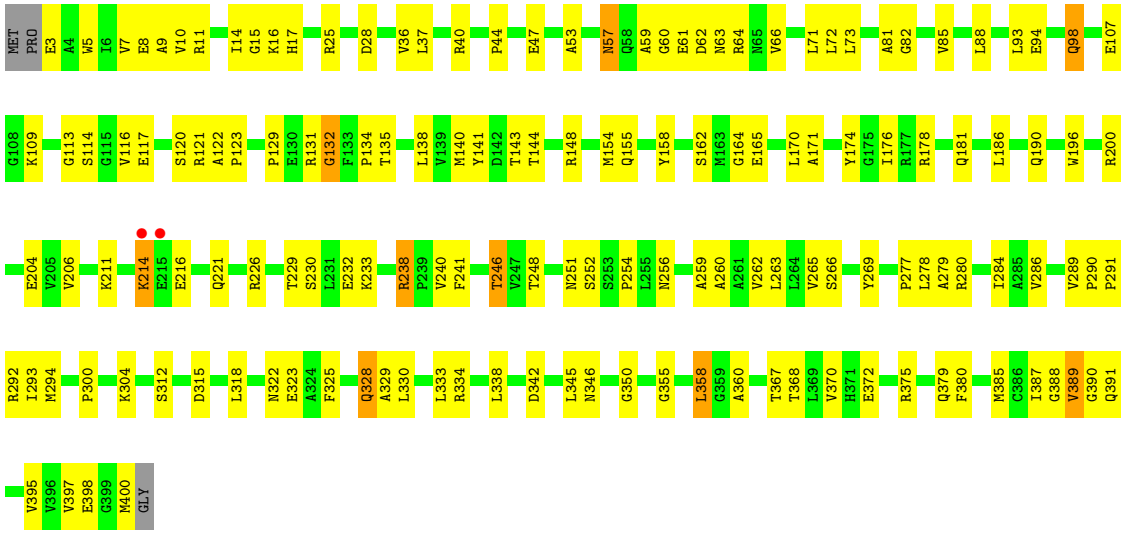
MET P2 E3 A4 W5 I6 V10 P13 I14 G15 K16 L30 A31 H32 S35 V36 L37 V38 S41 G42 V43 P44 K45 E46 E47 V48 A53 N57 Q58 A59 G60 E61 R64 N65 V66 A67 R68 L71 F76 P77 V78 A81 G82 V85 N86 R87 G90

E94 Q98 R101 V110 S114 G115 V116 E117 S118 M119 S120 R121 A122 P123 V126 P127 K128 P129 G132 F133 P134 T135 G136 N137 M140 Y141 D142 T143 T144 L145 G146 W147 R148 M154 Q155 Y158 G159 T160 E161 S162 M163 G164 E165 E168 N169 L170 R177 R178



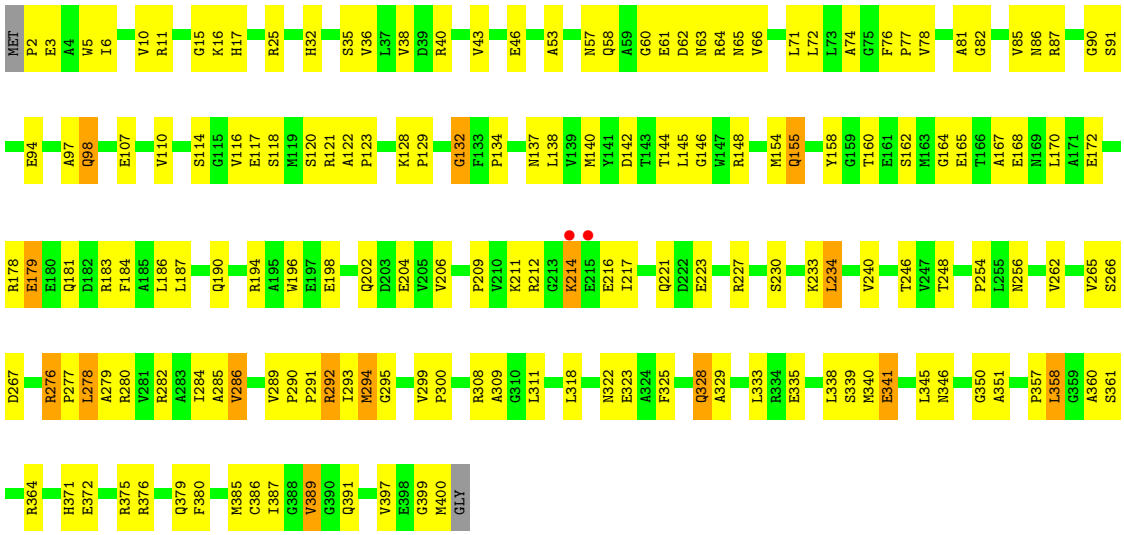
- Molecule 1: putative acetyl-CoA acetyltransferase

Chain F:



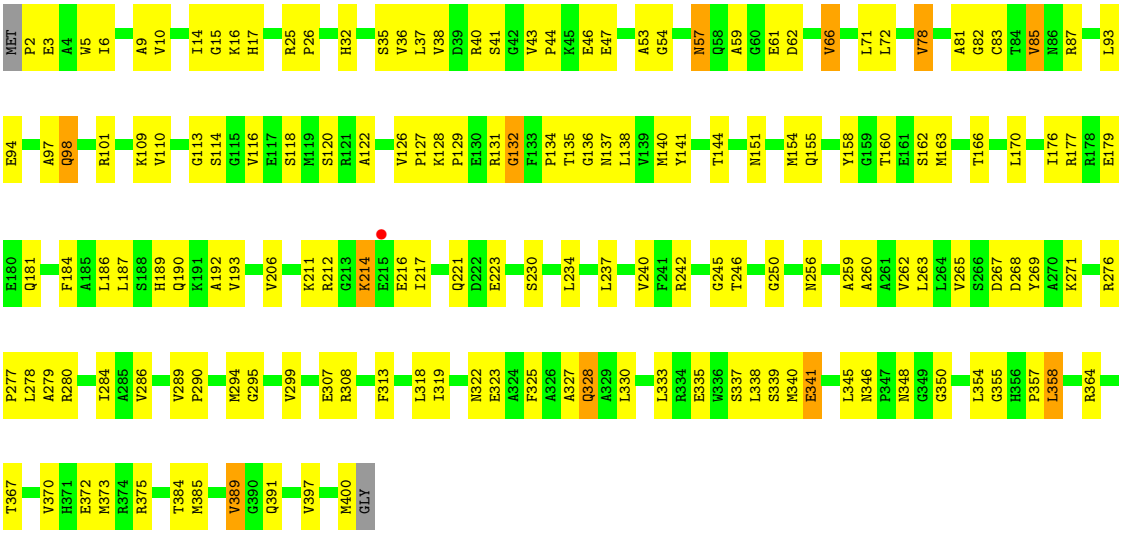
- Molecule 1: putative acetyl-CoA acetyltransferase

Chain G:



- Molecule 1: putative acetyl-CoA acetyltransferase

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	135.17Å 88.87Å 164.87Å 90.00° 100.45° 90.00°	Depositor
Resolution (Å)	48.50 – 3.00 48.52 – 3.00	Depositor EDS
% Data completeness (in resolution range)	93.7 (48.50-3.00) 93.7 (48.52-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.44 (at 3.01Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.218 , 0.275 0.219 , 0.213	Depositor DCC
R_{free} test set	3685 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	25.1	Xtriage
Anisotropy	0.933	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 24.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	23907	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2880e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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)

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	0.40	0/3036	0.66	0/4114
1	B	0.41	0/3036	0.69	0/4114
1	C	0.40	0/3044	0.66	0/4125
1	D	0.40	0/3044	0.66	0/4125
1	E	0.39	0/3044	0.66	0/4125
1	F	0.41	0/3036	0.67	0/4114
1	G	0.38	0/3044	0.66	0/4125
1	H	0.40	0/3044	0.67	0/4125
All	All	0.40	0/24328	0.67	0/32967

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (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	2984	0	3007	151	0
1	B	2984	0	3007	157	0
1	C	2991	0	3015	151	0
1	D	2991	0	3015	151	0
1	E	2991	0	3015	128	0
1	F	2984	0	3007	149	0
1	G	2991	0	3015	159	0
1	H	2991	0	3015	143	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	23907	0	24096	1068	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 22.

The worst 5 of 1068 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:129:PRO:HG2	1:H:129:PRO:HG2	1.18	1.15
1:H:36:VAL:HG11	1:H:206:VAL:HG11	1.25	1.10
1:G:2:PRO:HB2	1:G:5:TRP:HE1	1.17	1.09
1:A:36:VAL:HG21	1:A:206:VAL:HG21	1.37	1.06
1:B:129:PRO:HG2	1:D:129:PRO:HG2	1.09	1.06

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	396/401 (99%)	377 (95%)	15 (4%)	4 (1%)	15	53
1	B	396/401 (99%)	373 (94%)	20 (5%)	3 (1%)	19	57
1	C	397/401 (99%)	371 (94%)	22 (6%)	4 (1%)	15	53
1	D	397/401 (99%)	377 (95%)	17 (4%)	3 (1%)	19	57
1	E	397/401 (99%)	370 (93%)	24 (6%)	3 (1%)	19	57
1	F	396/401 (99%)	372 (94%)	21 (5%)	3 (1%)	19	57
1	G	397/401 (99%)	371 (94%)	23 (6%)	3 (1%)	19	57
1	H	397/401 (99%)	375 (94%)	19 (5%)	3 (1%)	19	57
All	All	3173/3208 (99%)	2986 (94%)	161 (5%)	26 (1%)	19	57

5 of 26 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	358	LEU
1	B	358	LEU
1	C	358	LEU
1	D	358	LEU
1	E	358	LEU

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	302/304 (99%)	284 (94%)	18 (6%)	19	53
1	B	302/304 (99%)	286 (95%)	16 (5%)	22	58
1	C	303/304 (100%)	284 (94%)	19 (6%)	18	51
1	D	303/304 (100%)	279 (92%)	24 (8%)	12	41
1	E	303/304 (100%)	286 (94%)	17 (6%)	21	56
1	F	302/304 (99%)	291 (96%)	11 (4%)	35	70
1	G	303/304 (100%)	287 (95%)	16 (5%)	22	58
1	H	303/304 (100%)	284 (94%)	19 (6%)	18	51
All	All	2421/2432 (100%)	2281 (94%)	140 (6%)	20	55

5 of 140 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	286	VAL
1	G	341	GLU
1	H	214	LYS
1	C	294	MET
1	C	292	ARG

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

Mol	Chain	Res	Type
1	D	322	ASN
1	H	155	GLN
1	E	391	GLN
1	H	98	GLN
1	G	328	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	A	398/401 (99%)	-0.59	0	100	100	1, 10, 27, 58	0
1	B	398/401 (99%)	-0.60	0	100	100	1, 9, 23, 57	0
1	C	399/401 (99%)	-0.64	2 (0%)	91	75	1, 10, 26, 60	0
1	D	399/401 (99%)	-0.46	0	100	100	1, 14, 32, 57	0
1	E	399/401 (99%)	-0.59	2 (0%)	91	75	1, 13, 31, 64	0
1	F	398/401 (99%)	-0.50	2 (0%)	91	75	1, 13, 32, 61	0
1	G	399/401 (99%)	-0.56	2 (0%)	91	75	3, 14, 33, 65	0
1	H	399/401 (99%)	-0.54	1 (0%)	94	84	1, 11, 28, 61	0
All	All	3189/3208 (99%)	-0.56	9 (0%)	94	84	1, 12, 30, 65	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	215	GLU	3.2
1	G	214	LYS	2.6
1	E	215	GLU	2.5
1	F	215	GLU	2.4
1	E	216	GLU	2.4

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 monosaccharides 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.