wwPDB X-ray Structure Validation Summary Report (i)

Feb 5, 2024 – 05:47 AM EST

PDB ID	:	1RXT
Title	:	Crystal structure of human myristoyl-CoA:protein N-myristoyltransferase.
Authors	:	Yang, J.; Wang, Y.; Frey, G.; Abeles, R.H.; Petsko, G.A.; Ringe, D.
Deposited on	:	2003-12-18
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 (i)) were used in the production of this report:

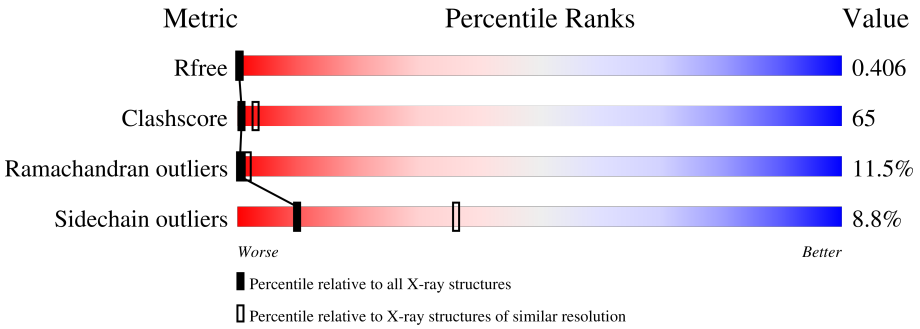
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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)

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%

Mol	Chain	Length	Quality of chain
1	A	496	20% 40% 8% • 31%
1	B	496	21% 39% 9% • 31%
1	C	496	17% 38% 9% • 34%
1	D	496	17% 39% 8% • 34%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 10420 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.

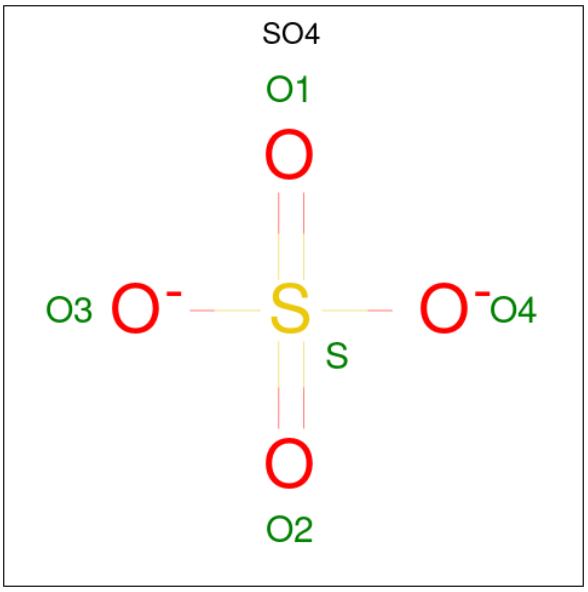
- Molecule 1 is a protein called Glycylpeptide N-tetradecanoyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	342	Total	C	N	O	S	0	0	0
			2646	1724	439	471	12			
1	B	342	Total	C	N	O	S	0	0	0
			2640	1718	438	472	12			
1	C	325	Total	C	N	O	S	0	0	0
			2538	1652	424	450	12			
1	D	326	Total	C	N	O	S	0	0	0
			2523	1642	418	451	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	452	ILE	LEU	conflict	UNP P30419
B	452	ILE	LEU	conflict	UNP P30419
C	452	ILE	LEU	conflict	UNP P30419
D	452	ILE	LEU	conflict	UNP P30419

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0

- Molecule 3 is COBALT (II) ION (three-letter code: CO) (formula: Co).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	C	1	Total Co 1 1	0	0

- Molecule 4 is water.

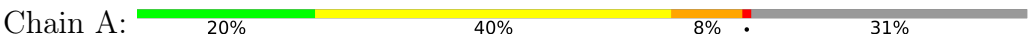
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	13	Total O 13 13	0	0
4	B	13	Total O 13 13	0	0
4	C	10	Total O 10 10	0	0
4	D	16	Total O 16 16	0	0

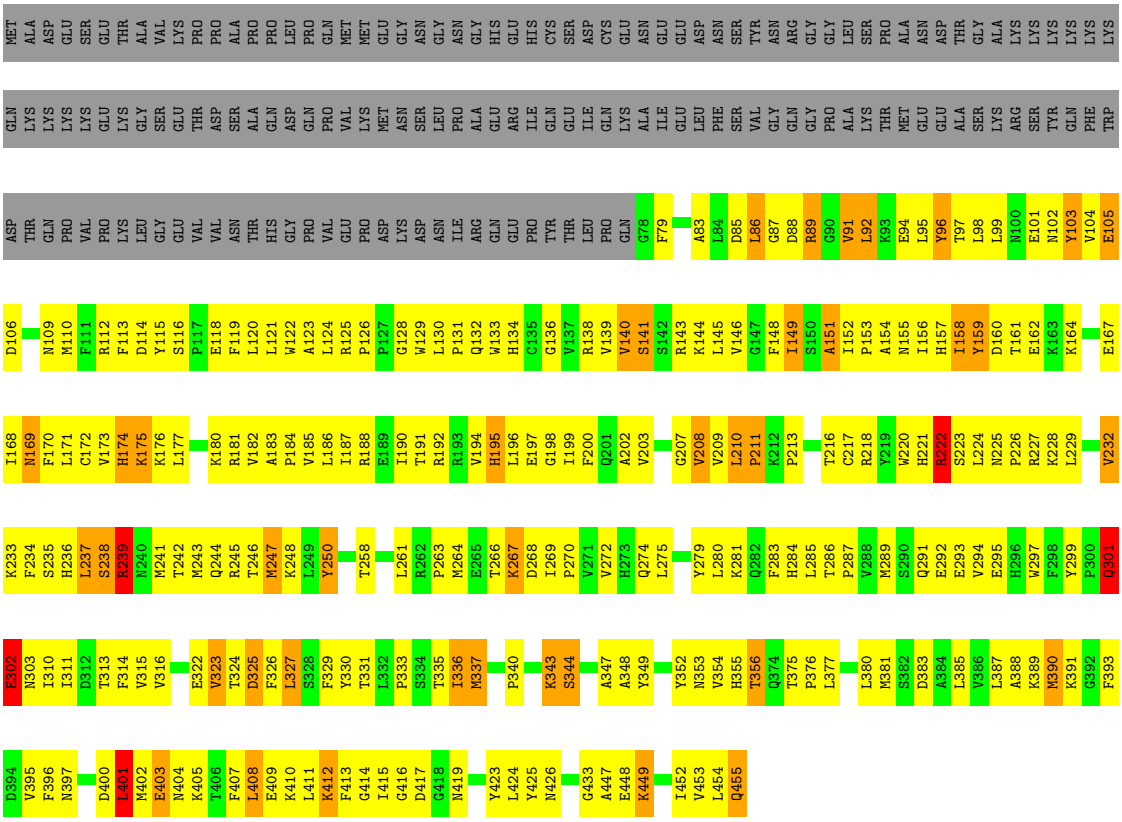
3 Residue-property plots (i)

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. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Glycylpeptide N-tetradecanoyltransferase 1

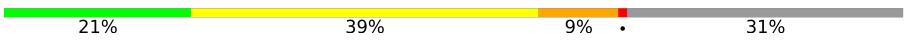
Chain A:

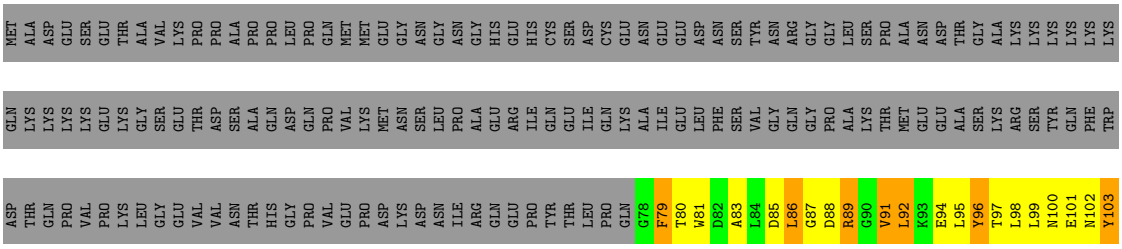


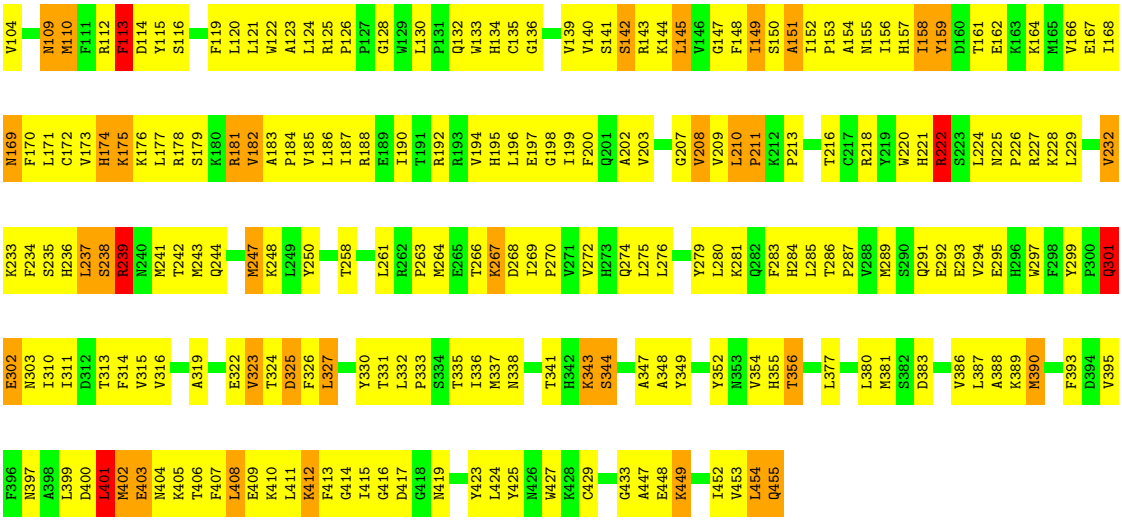


- Molecule 1: Glycylpeptide N-tetradecanoyltransferase 1

Chain B:

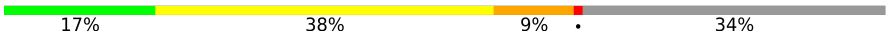


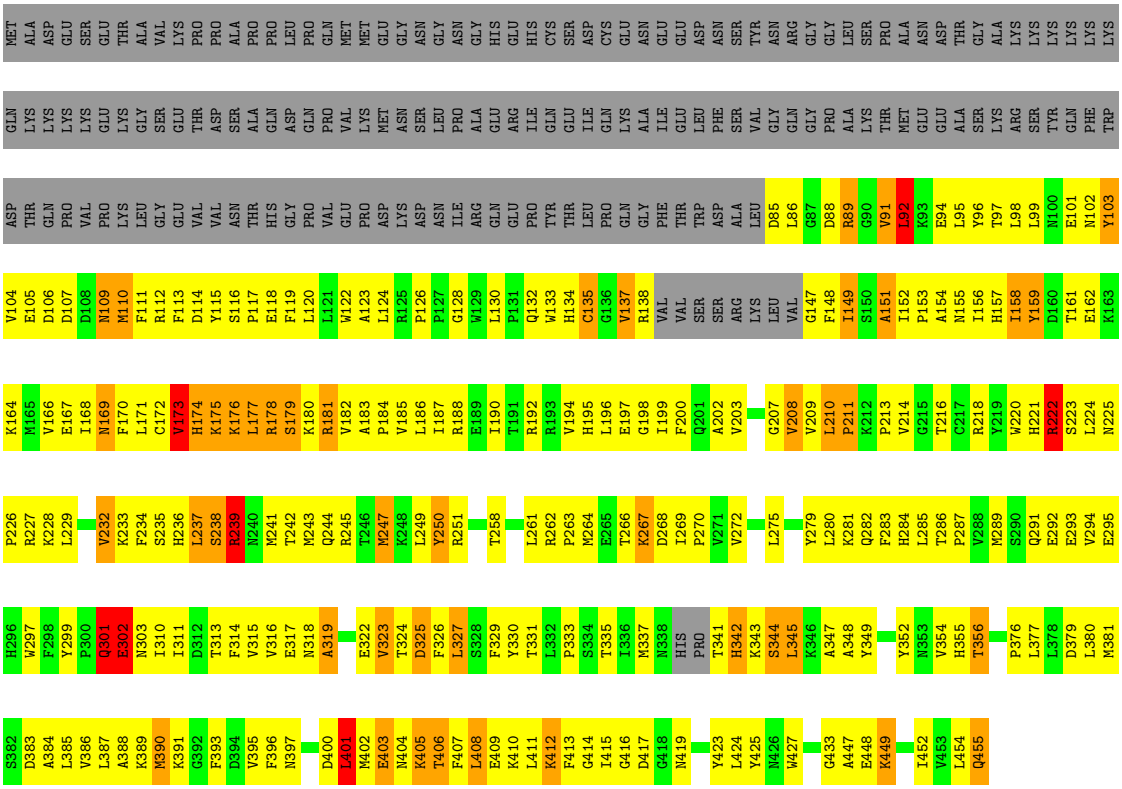




• Molecule 1: Glycylpeptide N-tetradecanoyltransferase 1

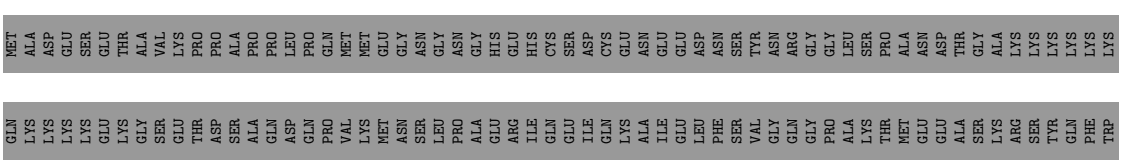
Chain C:

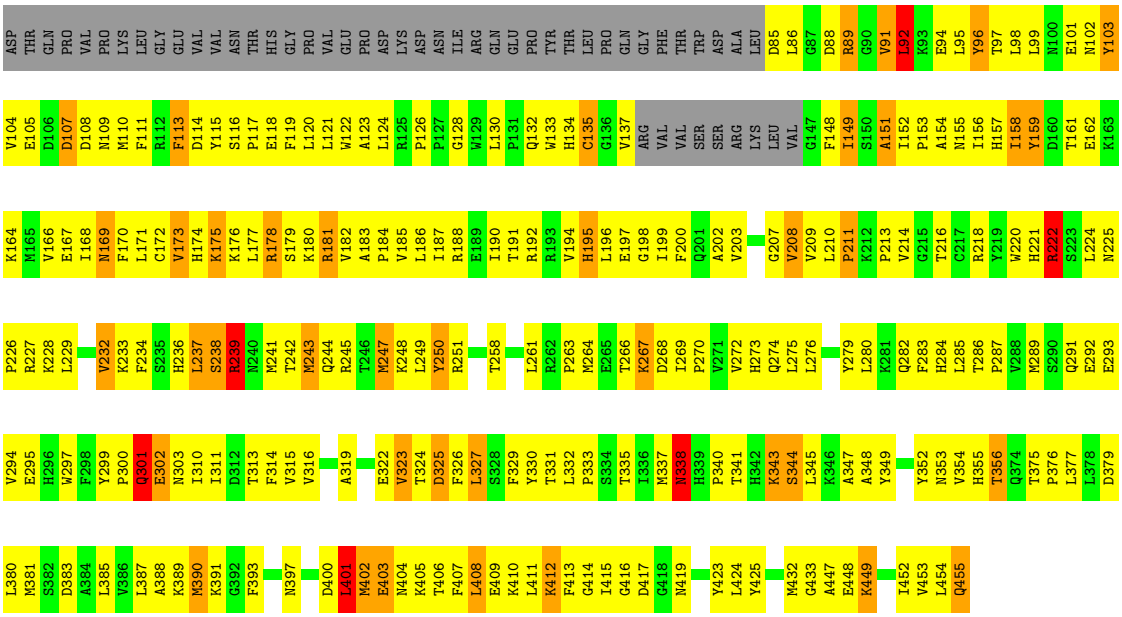




• Molecule 1: Glycylpeptide N-tetradecanoyltransferase 1

Chain D:




ASP THR GLN PRO VAL PRO LYS LEU GLY GLU VAL VAL ASN THR HIS GLY PRO VAL GLU PRO ASP LYS ASP ASN ILE ARG GLN GLU PRO TYR THR LEU PRO GLN GLY PHE THR TRP ASP ALA LEU
D85 L86 G87 D88 R89 G90 V91 L92 K93 E94 L95 Y96 T97 L98 L99 N100 E101 N102 Y103
V104 E105 D106 D107 D108 N109 M110 F111 R112 F113 D114 Y115 S116 P117 E118 F119 L120 L121 W122 A123 L124 R125 P126 P127 G128 W129 L130 P131 Q132 W133 H134 C135 G136 V137 ARG VAL VAL SER SER ARG LYS LEU VAL G147 F148 I149 S150 A151 I152 P153 A154 N155 I156 H157 I158 Y159 D160 T161 E162 K163
K164 M165 V166 E167 I168 N169 F170 L171 C172 V173 H174 K175 K176 L177 R178 S179 K180 R181 V182 A183 P184 V185 L186 I187 R188 E189 I190 T191 R192 R193 V194 H195 L196 E197 G198 I199 F200 Q201 A202 V203 G207 V208 V209 L210 P211 K212 P213 V214 G215 T216 C217 R218 Y219 W220 H221 R222 S223 L224 N225
P226 R227 K228 L229 V232 K233 F234 S235 H236 L237 S238 R239 N240 M241 T242 M243 Q244 R245 T246 M247 K248 L249 Y250 R251 T258 L261 R262 P263 M264 E265 T266 K267 D268 I269 P270 V271 V272 H273 Q274 L275 L276 Y279 L280 K281 Q282 F283 H284 L285 T286 P287 V288 M289 S290 Q291 E292 E293
V294 E295 H296 W297 F298 Y299 P300 Q301 E302 N303 I310 I311 D312 T313 F314 V315 V316 A319 E322 V323 T324 D325 F326 L327 S328 F329 Y330 T331 L332 P333 S334 T335 I336 M337 N338 H339 P340 T341 H342 K343 S344 L345 K346 A347 A348 Y349 Y352 N353 V354 H355 T356 Q374 T375 P376 L377 L378 D379
L380 M381 S382 D383 A384 L385 V386 L387 A388 K389 M390 K391 G392 F393 N397 D400 L401 M402 E403 N404 K405 T406 F407 L408 E409 K410 L411 K412 F413 G414 I415 G416 D417 G418 N419 Y423 L424 Y425 M432 G433 A447 E448 K449 T452 V453 L454 Q455

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	71.63Å 116.42Å 90.36Å 90.00° 90.13° 90.00°	Depositor
Resolution (Å)	20.00 – 3.00 29.44 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.00) 96.5 (29.44-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 3.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.283 , 0.391 0.338 , 0.406	Depositor DCC
R_{free} test set	1412 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	45.2	Xtriage
Anisotropy	0.694	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 121.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.57$, $< L^2 > = 0.43$	Xtriage
Estimated twinning fraction	0.380 for h,-k,-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	10420	wwPDB-VP
Average B, all atoms (Å^2)	47.0	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 59.63 % 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 1.7162e-05. 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CO, SO4

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	0.76	3/2718 (0.1%)	1.01	8/3715 (0.2%)
1	B	1.04	4/2712 (0.1%)	1.15	8/3709 (0.2%)
1	C	0.74	3/2605 (0.1%)	0.96	6/3551 (0.2%)
1	D	1.10	4/2592 (0.2%)	1.12	9/3542 (0.3%)
All	All	0.92	14/10627 (0.1%)	1.06	31/14517 (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	A	0	3
1	B	0	3
1	C	0	4
1	D	0	3
All	All	0	13

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	401	LEU	C-N	-32.53	0.59	1.34
1	D	401	LEU	C-N	-32.37	0.59	1.34
1	B	401	LEU	C-N	-31.27	0.62	1.34
1	B	402	MET	C-N	28.57	1.99	1.34
1	D	302	GLU	C-N	-27.80	0.70	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	301	GLN	O-C-N	-47.53	46.66	122.70
1	B	301	GLN	O-C-N	-44.79	51.04	122.70
1	C	301	GLN	O-C-N	-37.26	63.09	122.70
1	A	301	GLN	O-C-N	-29.43	75.62	122.70
1	B	401	LEU	O-C-N	-24.32	83.78	122.70

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	301	GLN	Mainchain
1	A	401	LEU	Peptide,Mainchain
1	B	301	GLN	Mainchain
1	B	401	LEU	Peptide,Mainchain
1	C	301	GLN	Mainchain

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	2646	0	2523	316	0
1	B	2640	0	2514	336	0
1	C	2538	0	2434	338	1
1	D	2523	0	2395	332	1
2	A	10	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
3	C	1	0	0	0	0
4	A	13	0	0	3	0
4	B	13	0	0	3	0
4	C	10	0	0	4	0
4	D	16	0	0	2	0
All	All	10420	0	9866	1322	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 65.

The worst 5 of 1322 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:B:302:GLU:CA	1:B:303:ASN:N	1.95	1.29
1:D:302:GLU:CA	1:D:303:ASN:N	1.94	1.29
1:D:302:GLU:C	1:D:303:ASN:CA	2.03	1.26
1:C:302:GLU:O	1:C:303:ASN:N	1.61	1.26
1:D:302:GLU:O	1:D:303:ASN:N	1.68	1.25

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:196:LEU:O	1:D:196:LEU:O[1_655]	2.17	0.03

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	338/496 (68%)	224 (66%)	80 (24%)	34 (10%)	0	2
1	B	340/496 (68%)	226 (66%)	73 (22%)	41 (12%)	0	1
1	C	317/496 (64%)	212 (67%)	65 (20%)	40 (13%)	0	1
1	D	322/496 (65%)	213 (66%)	73 (23%)	36 (11%)	0	2
All	All	1317/1984 (66%)	875 (66%)	291 (22%)	151 (12%)	0	2

5 of 151 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	141	SER
1	A	151	ALA
1	A	232	VAL
1	A	238	SER
1	A	239	ARG

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	272/443 (61%)	247 (91%)	25 (9%)	9 34
1	B	271/443 (61%)	247 (91%)	24 (9%)	9 35
1	C	263/443 (59%)	240 (91%)	23 (9%)	10 37
1	D	259/443 (58%)	237 (92%)	22 (8%)	10 38
All	All	1065/1772 (60%)	971 (91%)	94 (9%)	10 36

5 of 94 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	177	LEU
1	C	455	GLN
1	C	210	LEU
1	C	302	GLU
1	D	101	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	244	GLN
1	D	338	ASN
1	D	455	GLN
1	B	273	HIS
1	B	244	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)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

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	SO4	C	904	-	4,4,4	0.27	0	6,6,6	0.05	0
2	SO4	A	902	-	4,4,4	0.27	0	6,6,6	0.04	0
2	SO4	A	903	-	4,4,4	0.26	0	6,6,6	0.06	0
2	SO4	B	901	-	4,4,4	0.29	0	6,6,6	0.10	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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	4

Continued on next page...

Continued from previous page...

Mol	Chain	Number of breaks
1	A	4
1	B	4
1	D	4

The worst 5 of 16 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	402:MET	C	403:GLU	N	2.12
1	A	402:MET	C	403:GLU	N	2.01
1	B	402:MET	C	403:GLU	N	1.99
1	D	402:MET	C	403:GLU	N	1.97
1	A	302:GLU	C	303:ASN	N	1.15

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.