WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 5, 2026 – 05:24 PM JST

PDB ID	:	9LF6 / pdb_00009lf6
Title	:	Crystal of Glucosidase form Gelsemium sempervirens
Authors	:	Hao, X.L.; Kai, G.Y.
Deposited on	:	2025-01-08
Resolution	:	2.95 Å(reported)

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

|		
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.47

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 2.95 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.
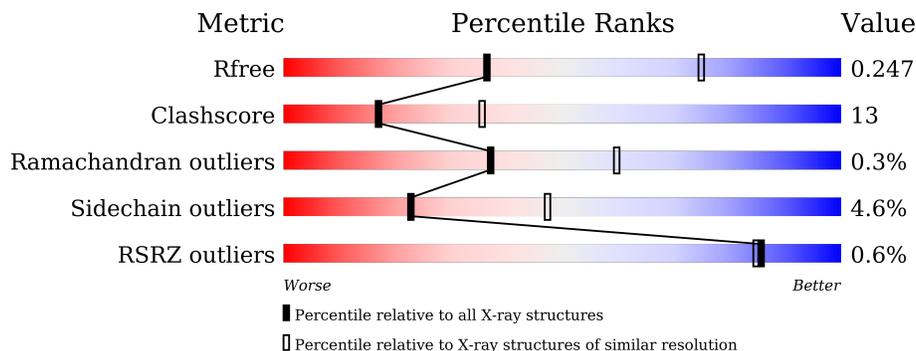
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1044 (2.98-2.94)
Clashscore	180529	1097 (2.98-2.94)
Ramachandran outliers	177936	1049 (2.98-2.94)
Sidechain outliers	177891	1049 (2.98-2.94)
RSRZ outliers	164620	1044 (2.98-2.94)

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	534	68% 20% · 11%
1	B	534	% 62% 25% · 12%
1	C	534	% 61% 26% · 12%
1	D	534	% 49% 33% 5% 12%
1	E	534	65% 22% · 13%
1	F	534	% 64% 22% · 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	534	58% 28% 13%
1	H	534	57% 27% 13%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PEG	B	602	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 29334 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 Strictosidine glucosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	474	Total	C	N	O	S	0	0	0
			3670	2363	609	691	7			
1	B	469	Total	C	N	O	S	0	0	0
			3676	2366	609	693	8			
1	C	471	Total	C	N	O	S	0	0	0
			3687	2371	612	696	8			
1	D	469	Total	C	N	O	S	0	0	0
			3673	2364	608	693	8			
1	E	466	Total	C	N	O	S	0	0	0
			3656	2351	608	689	8			
1	F	466	Total	C	N	O	S	0	0	0
			3629	2333	597	691	8			
1	G	466	Total	C	N	O	S	0	0	0
			3625	2331	600	686	8			
1	H	464	Total	C	N	O	S	0	0	0
			3613	2321	598	686	8			

- Molecule 2 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).

PEG
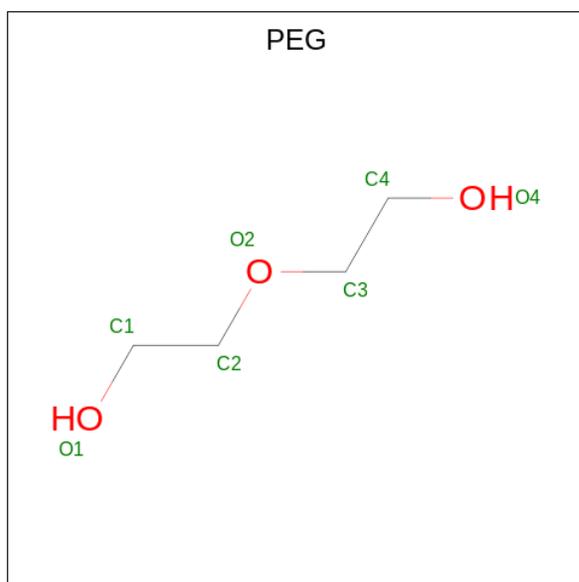
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 7	C 4	O 3	0	0
2	A	1	Total 7	C 4	O 3	0	0
2	B	1	Total 7	C 4	O 3	0	0
2	B	1	Total 7	C 4	O 3	0	0
2	C	1	Total 7	C 4	O 3	0	0
2	C	1	Total 7	C 4	O 3	0	0
2	C	1	Total 7	C 4	O 3	0	0
2	C	1	Total 7	C 4	O 3	0	0
2	E	1	Total 7	C 4	O 3	0	0
2	F	1	Total 7	C 4	O 3	0	0
2	H	1	Total 7	C 4	O 3	0	0

- Molecule 3 is SULFATE ION (CCD ID: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2
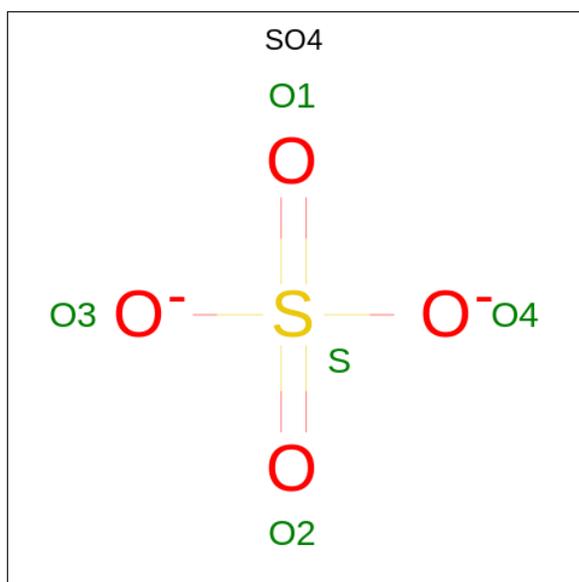
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total 5	O 4	S 1	0	0
3	D	1	Total 5	O 4	S 1	0	0
3	D	1	Total 5	O 4	S 1	0	0
3	E	1	Total 5	O 4	S 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: $C_2H_6O_2$).

EDO

O2

OH

C1

C2

HO

O1
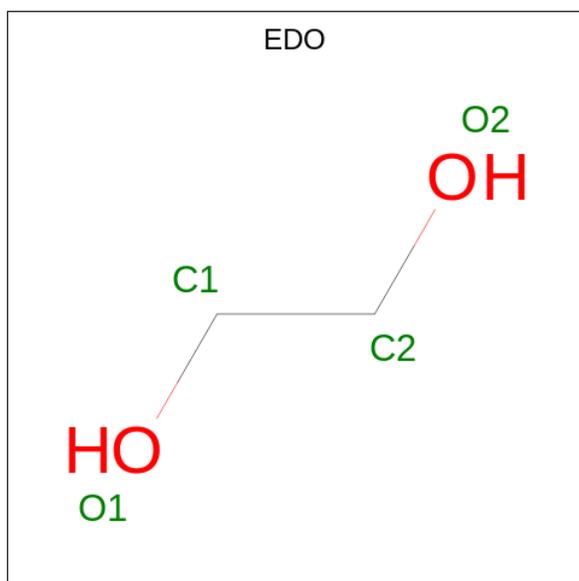
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total C O 4 2 2			0	0
4	H	1	Total C O 4 2 2			0	0

3 Residue-property plots ⓘ

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: Strictosidine glucosidase

Chain A:

- Molecule 1: Strictosidine glucosidase

Chain B:

V459 K460 G461 W465 S466 L467 L468 E472 W473 N474 E475 V479 R480 F481 G482 M483 I484 H485 V486 Q491 Y492 A493 K497 N505 N506 F507 H508 K509 LYS PHE GLY PRO PRO VAL LYS ARG SER VAL GLU GLU ASN GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS

• Molecule 1: Strictosidine glucosidase

Chain C:

%

61%   26%   •   12%

MET ALA THR PRO SER SER THR ILE VAL PRO D11 F20 F26 A29 S30 S31 A32 Y33 Q34 R44 G45 D50 T51 F52 T53 K54 E58 M59 V60 K61 S64 S72 Y76 K77 E78 D79 G87 L88 S94 I95 S96 W97 S98 R99 I100 Y118 T135

H138 D147 P155 V158 F161 Y164 G173 N178 L182 N183 E184 P185 W186 T187 V190 D191 G192 Y193 V194 A195 G196 T197 R202 G203 ALA THR PRO ASP GLN VAL LYS GLY PRO ILE LYS ARG HIS ARG CYS SER GLY TRP GLY PRO GLN CYS S227 N228

S229 D230 G231 Q244 A247 H248 V252 R256 N257 K258 F259 K260 A261 S262 Q263 I267 G268 I269 T270 I271 V272 A273 Q274 W275 M276 E277 P278 S283 V287 R292 M297 F301 M302 E303 P304 M314 Q329 D338 F339 L340 G341 L342 N343 N348 Y349 V350

N355 PRO THR GLY GLY ILE V361 S362 Y363 T367 T370 D374 I380 S386 L389 H390 I391 Y392 P393 R397 K398 Y407 N408 L411 I412 Y413 I414 T415 E416 N417 G418 V419 D420 E421 L422 N423 D424 L427 T428 L429 S430 E431 A432 R433 V434 D435 P436 I437 R438

I439 K440 F441 I442 Q443 D444 H445 V459 K460 G461 W465 S466 L467 L468 D469 N470 T478 V479 R480 F481 G482 Y495 P496 S499 W502 L503 M504 N505 H508 K509 LYS PHE SER GLY PRO PRO VAL LYS SER VAL GLU GLU ASN GLN GLU THR ASP SER ARG LYS ARG SER

ARG LYS

• Molecule 1: Strictosidine glucosidase

Chain D:

%

49%   33%   5%   12%

MET ALA THR PRO SER SER THR I8 V9 P10 D11 A12 I15 N16 R17 R18 D19 S22 A29 S30 S31 I35 E36 G37 G38 A39 S40 E41 G42 G43 R44 S47 D50 T51 K54 R55 E58 M59 V60 K61 N65 G66 N67 I70 D71 S72 V76 K77

V80 L83 L86 G87 L88 D89 A90 I95 S96 W97 S98 R99 I100 G104 G109 I110 N111 K112 E113 G114 I115 Y118 F121 I122 L125 I126 I130 V134 F137 H138 W139 D140 V141 P142 Q143 A144 L145 E146 D147 S154 P155 K156 I157 V158 Y164 A165 E166

L167 C168 G173 N183 E184 P185 S189 V190 Y193 V194 T197 F198 A199 P200 G203 ALA THR PRO THR ASP GLN VAL LYS GLY PRO ILE LYS ARG HIS ARG CYS SER GLY TRP GLY PRO GLN CYS SER ASN SER D230 G231 N232 P233 G234 T235 E236 P237 T241 H242 A247

H248 V252 E253 S254 Y255 R256 N257 K258 F259 K260 A261 S262 I267 G268 I269 T270 I271 V272 A273 Q274 W275 P278 L279 N280 E281 K282 S283 D284 S285 L294 D295 F296 M297 Y298 F301 M302 P304 S307 E312 I313 M314 K315 V318 G319 S320 R321 L322 S326 A327 E328

Q329 S330 Y337 D338 G341 L342 Y345 T346 A347 V350 T351 S352 A353 PRO ASN PRO THR GLY GLY ILE V361 S362 Y363 D366 T367 T370 S373 D374 R375 G381 P382 L383 A384 S386 E387 W388 L389 H390 I391 Y392 P393 E394 G395 I396 R397 K398 L399 Y407 N408 V409

P410 L411 I412 E416 E421 T428 L429 S430 R433 V434 D435 P436 I437 R438 I442 Q443 D444 H445 L446 L447 Q448 L449 R450 L451 A452 I463 D454 D455 N458 V459 W465 S466 L467 L468 D469 N474 E475 V479 G482 M483 I484 H485 V486 N487 Q491 Y495 P496

S499 A500 I501 W502 L503 M504 N505 N506 F507 H508 K509 LYS PHE SER GLY PRO PRO VAL LYS ARG SER VAL GLU GLU GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS
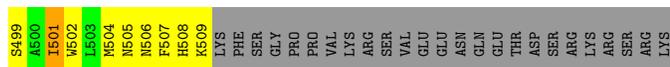
● Molecule 1: Strictosidine glucosidase

Chain E:

65%   22%   • 13%

MET ALA THR PRO SER SER THR ILE VAL PRO D11 N16 R17 F20 S30 S31 A32 I35 E36 R44 T51 R55 K61 V80 K81 I82 D89 A90 Y91 R92 F93 S94 W97 S98 R99 I100 G104 N105 L106 I115 L125 I126 Q131 P132 Y133

V134 T135 H138 L145 E146 D147 F152 L153 S154 I157 Y164 W170 I180 E184 P185 W186 S189 V190 Y193 V194 R202 G203 ALA THR PRO THR ASP GLN VAL LYS GLY PRO ILE LYS ARG HIS ARG CYS SER GLY TRP GLY PRO GLN CYS SER ASN SER ASP

G231 Q234 T235 E236 P237 V240 T241 H248 V252 E253 S254 Y255 R256 K260 A261 S262 T270 I271 V272 A273 Q274 W275 M276 E277 S285 D286 R292 F296 M297 F301 M302 E303 P304 I305 T306 S307 I313 I317 R321 Y337 N343 Y344 N348 Y349

V350 T351 P354 ASN PRO THR GLY GLY ILE V361 S362 T367 Q368 V369 G381 S386 E387 W388 I391 Y392 P393 E394 G395 I396 R397 K398 L399 L400 I412 Y413 I414 T415 E416 N417 G418 V419 D420 E421 T425 E431 I439 Q443 I449 R450 L451 V459 W465

S466 E475 G476 R480 F481 I484 Y492 A493 R494 Y495 P496 W502 N505 K509 LYS PHE SER GLY PRO PRO VAL LYS ARG SER VAL GLU GLU GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS

● Molecule 1: Strictosidine glucosidase

Chain F:

%

64%   22%   • 13%

MET ALA THR PRO SER SER THR ILE VAL PRO ASP A12 I15 N16 R17 R18 D19 A29 S30 S31 A39 S40 R44 G45 T51 R55 E58 M59 V60 S64 I82 L86 G87 L88 D89 A90 Y91 R92 S98 R99 I100 L101 I110 Y118 F121 I122

H138 W139 D147 F152 L153 S154 I157 V158 E166 L167 W170 E184 P185 W186 V190 Y193 V194 R202 G203 ALA THR PRO THR ASP GLN VAL LYS GLY PRO ILE LYS ARG HIS ARG CYS SER GLY TRP GLY PRO GLN CYS SER N228 S229 D230 G231 N232 P233 G234

T235 E236 P237 Y238 L239 V240 T241 H242 A247 H248 V252 E253 S254 Y255 R256 K260 A261 S262 Q266 I269 V272 A273 Q274 W275 M276 E277 S285 D286 A289 R292 A293 L294 D295 F296 M297 F301 M302 E303 P304 M314 R321 L322 P323 S326 Q329 S330

L333 S336 Y337 D338 N343 T346 A347 N348 Y349 V350 T351 S352 A353 PRO ASN PRO THR GLY GLY ILE VAL S362 Y363 D366 D374 R375 N376 L379 I380 L383 A384 L389 H390 E394 K398 E416 N417 G418 V419 D420 L429 R433 I437 L446

L449 V459 Y462 W465 S466 W473 N474 F481 G482 M483 I484 Y495 P496 M504 F507 H508 K509 LYS PHE SER GLY PRO PRO VAL LYS ARG SER VAL GLU GLU GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS

● Molecule 1: Strictosidine glucosidase

Chain G:

%

58%   28%   • 13%

MET ALA THR PRO SER SER THR ILE VAL PRO D11 N16 R17 R18 G27 A28 S31 A32 I35 E36 R44 I48 T51 R55 R56 M59 S64 K84 N85 L86 D89 A90 Y91 R92 F93 S94 I95 S96 R99 I100 L101 Y118 I122 L125

I126 A127 S128 Y133 H138 W139 D140 Q43 E146 D147 G151 F161 E166 L167 W170 G173 D174 E184 P185 W186 T187 V190 D191 G192 Y193 V194 F198 R202 G203 ALA THR PRO THR ASP GLN VAL LYS GLY PRO ILE LYS ARG HIS ARG CYS SER GLY

TRP GLY PRO GLN CYS SER ASN SER D230 G231 N232 P233 G234 T235 E236 P237 V240 T241 H242 H243 Q244 I245 L246 A247 S254 K258 F259 S262 Q263 Q266 V272 A273 Q274 W275 M276 E277 N280 S283 D284 S285 R292 F296 M297 F301 M302 E303 P304 D309

M314 R321 L322 P323 K324 F325 L333 S336 F339 L340 G341 L342 N343 Y349 V350 T351 S352 A353 PRO ASN PRO THR GLY GLY ILE V361 S362 Y363 T367 Q368 V369 S373 D374 R375 N376 G381 P382 L383 S386 E387 W388 L389 H390 I391 E394 G395 I396 R397 K398

L399 L400 V401 Y402 L411 I412 Y413 I414 T415 E416 D420 E421 D424 L427 R433 V434 D435 R438 L446 L447 Q448 L449 V459 K460 G461 Y462 F463 V464 W465 S466 L467 L468 F471 E475 G476 F477 T478 V479 R480 F481 G482 M483 I484 N489 P496 S499 A500

M504 N505 N506 F507 H508 K509 LYS PHE SER GLY PRO PRO VAL LYS ARG SER VAL GLU GLU ASN GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS

• Molecule 1: Strictosidine glucosidase

Chain H:

57%   27%   •   13%

MET ALA THR PRO SER SER THR ILE VAL PRO D11 A12 I15 N16 R17 S30 S31 Q34 I35 E36 R44 T51 K54 R55 R56 P57 E58 M59 N67 I70 D71 S72 Y76 I82 G87 R92 W97 S98 R99 N105 Y118 I122 I126

Q131 L136 F137 H138 A144 D147 E148 Y149 G150 G151 F152 L153 I157 V158 D159 D160 Y164 L167 F172 G173 D174 R175 N178 L182 N183 E184 P185 W186 S189 V190 Y193 V194 A199 P200 G201 R202 G203 ALA THR PRO THR ASP GLN VAL LYS GLY PRO

ILE LYS ARG HIS ARG CYS SER GLY TRP GLY PRO GLN CYS SER ASN SER ASP G231 N232 P233 G234 T235 E236 P237 Y238 L239 V240 T241 H242 A247 H248 V252 Y255 K260 A261 S262 I267 T270 I271 V272 A273 Q274 W275 M276 R292 F296 M297 W300 F301 T306

I313 M314 I317 V318 G319 S320 R321 L322 S326 Q329 L340 N343 N348 Y349 V350 A353 PRO ASN PRO THR GLY GLY ILE VAL S362 T370 I380 G381 I391 Y392 P393 I396 L400 I412 Y413 E416 N417 G418 V419 D420 E421 L422 N423 D424 T425

S426 L427 T428 L429 V434 D435 R438 I439 K440 F441 I442 Q443 D444 H445 L446 L447 Q448 L451 A452 I453 D454 D455 G456 V457 N458 V459 K460 G461 W465 S466 L467 L468 D469 N470 V479 R480 F481 G482 M483 I484 H485 V486 Q491 Y492 A493 R494 Y495 P496 K497 D498 S499 L503 M504

N505 N506 F507 H508 K509 LYS PHE SER GLY PRO PRO VAL LYS ARG SER VAL GLU GLU ASN GLN GLU THR ASP SER ARG LYS ARG SER ARG LYS

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	237.43Å 107.84Å 220.27Å 90.00° 118.46° 90.00°	Depositor
Resolution (Å)	48.99 – 2.95 48.99 – 2.95	Depositor EDS
% Data completeness (in resolution range)	95.8 (48.99-2.95) 95.8 (48.99-2.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.84 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.206 , 0.248 0.206 , 0.247	Depositor DCC
R_{free} test set	4958 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	22.0	Xtriage
Anisotropy	0.484	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 35.4	EDS
L-test for twinning[2]	$< |L| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	29334	wwPDB-VP
Average B, all atoms (Å^2)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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)

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

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	0.46	0/3776	0.62	0/5154
1	B	0.77	4/3780 (0.1%)	0.75	6/5150 (0.1%)
1	C	0.70	2/3791 (0.1%)	0.71	5/5166 (0.1%)
1	D	1.50	5/3777 (0.1%)	1.15	26/5145 (0.5%)
1	E	0.52	1/3760 (0.0%)	0.61	0/5124
1	F	0.53	0/3732	0.64	3/5089 (0.1%)
1	G	0.66	1/3728 (0.0%)	0.69	2/5084 (0.0%)
1	H	0.84	3/3715 (0.1%)	0.73	7/5065 (0.1%)
All	All	0.81	16/30059 (0.1%)	0.76	49/40977 (0.1%)

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	341	GLY	CA-C	-6.96	1.47	1.52
1	D	487	ASN	CA-C	-6.50	1.44	1.52
1	B	473	TRP	CA-C	-6.24	1.47	1.53
1	C	421	GLU	N-CA	-6.05	1.38	1.45
1	B	472	GLU	CA-C	-6.00	1.44	1.52

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	421	GLU	N-CA-C	-9.68	94.53	108.96
1	C	459	VAL	N-CA-C	-9.57	94.68	108.36
1	D	260	LYS	N-CA-C	8.93	121.01	111.28
1	D	134	VAL	N-CA-C	8.74	120.48	107.80
1	B	456	GLY	N-CA-C	7.84	126.29	115.00

There are no chirality outliers.

There are no planarity outliers.

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	3670	0	3395	74	0
1	B	3676	0	3428	92	0
1	C	3687	0	3432	98	0
1	D	3673	0	3422	119	0
1	E	3656	0	3404	75	0
1	F	3629	0	3342	83	0
1	G	3625	0	3340	105	0
1	H	3613	0	3327	101	0
2	A	14	0	20	0	0
2	B	14	0	20	6	0
2	C	28	0	40	4	0
2	E	7	0	10	0	0
2	F	7	0	10	0	0
2	H	7	0	10	0	0
3	B	5	0	0	1	0
3	D	10	0	0	0	0
3	E	5	0	0	0	0
4	E	4	0	6	0	0
4	H	4	0	6	0	0
All	All	29334	0	27212	709	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 13.

The worst 5 of 709 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:143:GLN:HE22	2:B:602:PEG:H11	1.20	1.06
1:A:147:ASP:HB3	1:B:147:ASP:HB3	1.47	0.94
1:G:449:LEU:HD11	1:G:459:VAL:HG11	1.50	0.92
1:G:56:ARG:HE	1:G:59:MET:HE3	1.37	0.88
1:A:51:THR:HG23	1:A:55:ARG:HD2	1.56	0.87

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	470/534 (88%)	446 (95%)	24 (5%)	0	100	100
1	B	463/534 (87%)	434 (94%)	26 (6%)	3 (1%)	22	46
1	C	465/534 (87%)	440 (95%)	25 (5%)	0	100	100
1	D	463/534 (87%)	429 (93%)	28 (6%)	6 (1%)	10	27
1	E	460/534 (86%)	437 (95%)	21 (5%)	2 (0%)	30	54
1	F	460/534 (86%)	435 (95%)	25 (5%)	0	100	100
1	G	460/534 (86%)	421 (92%)	39 (8%)	0	100	100
1	H	458/534 (86%)	426 (93%)	31 (7%)	1 (0%)	44	67
All	All	3699/4272 (87%)	3468 (94%)	219 (6%)	12 (0%)	37	60

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	230	ASP
1	D	36	GLU
1	E	476	GLY
1	H	491	GLN
1	D	408	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	A	369/456 (81%)	360 (98%)	9 (2%)	44	68

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	375/456 (82%)	364 (97%)	11 (3%)	37	61
1	C	376/456 (82%)	359 (96%)	17 (4%)	23	48
1	D	374/456 (82%)	337 (90%)	37 (10%)	6	18
1	E	373/456 (82%)	361 (97%)	12 (3%)	34	59
1	F	367/456 (80%)	356 (97%)	11 (3%)	36	60
1	G	365/456 (80%)	346 (95%)	19 (5%)	19	42
1	H	364/456 (80%)	343 (94%)	21 (6%)	17	38
All	All	2963/3648 (81%)	2826 (95%)	137 (5%)	23	47

5 of 137 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	479	VAL
1	H	136	LEU
1	H	370	THR
1	D	157	ILE
1	D	130	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	143	GLN
1	G	244	GLN
1	H	65	ASN
1	C	372	HIS
1	C	348	ASN

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 (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

17 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	PEG	A	602	-	6,6,6	0.23	0	5,5,5	0.16	0
2	PEG	F	601	-	6,6,6	0.50	0	5,5,5	0.30	0
2	PEG	E	601	-	6,6,6	0.49	0	5,5,5	0.32	0
2	PEG	A	601	-	6,6,6	0.42	0	5,5,5	0.30	0
3	SO4	D	602	-	4,4,4	0.19	0	6,6,6	0.31	0
2	PEG	C	601	-	6,6,6	0.39	0	5,5,5	0.27	0
2	PEG	B	601	-	6,6,6	0.30	0	5,5,5	0.32	0
3	SO4	E	603	-	4,4,4	0.22	0	6,6,6	0.43	0
4	EDO	E	602	-	3,3,3	0.71	0	2,2,2	0.19	0
2	PEG	C	604	-	6,6,6	0.33	0	5,5,5	0.14	0
4	EDO	H	602	-	3,3,3	0.42	0	2,2,2	0.99	0
2	PEG	B	602	-	6,6,6	0.57	0	5,5,5	0.44	0
2	PEG	C	602	-	6,6,6	0.34	0	5,5,5	0.16	0
2	PEG	C	603	-	6,6,6	0.38	0	5,5,5	0.30	0
3	SO4	D	601	-	4,4,4	0.37	0	6,6,6	0.38	0
2	PEG	H	601	-	6,6,6	0.41	0	5,5,5	0.21	0
3	SO4	B	603	-	4,4,4	0.25	0	6,6,6	0.27	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	PEG	A	602	-	-	2/4/4/4	-
2	PEG	F	601	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PEG	E	601	-	-	1/4/4/4	-
2	PEG	A	601	-	-	1/4/4/4	-
2	PEG	C	601	-	-	1/4/4/4	-
2	PEG	B	601	-	-	3/4/4/4	-
4	EDO	E	602	-	-	1/1/1/1	-
4	EDO	H	602	-	-	0/1/1/1	-
2	PEG	C	604	-	-	1/4/4/4	-
2	PEG	B	602	-	-	1/4/4/4	-
2	PEG	C	602	-	-	1/4/4/4	-
2	PEG	C	603	-	-	4/4/4/4	-
2	PEG	H	601	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	601	PEG	O1-C1-C2-O2
2	C	602	PEG	O1-C1-C2-O2
2	C	603	PEG	O2-C3-C4-O4
2	C	604	PEG	O2-C3-C4-O4
2	B	601	PEG	O2-C3-C4-O4

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	601	PEG	1	0
2	C	604	PEG	3	0
2	B	602	PEG	5	0
2	C	602	PEG	1	0
3	B	603	SO4	1	0

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	A	474/534 (88%)	-0.35	2 (0%)	89	89	5, 10, 19, 34	0
1	B	469/534 (87%)	-0.27	3 (0%)	85	85	5, 11, 22, 33	0
1	C	471/534 (88%)	-0.21	6 (1%)	74	72	8, 15, 24, 45	0
1	D	469/534 (87%)	0.16	4 (0%)	81	79	8, 16, 26, 31	0
1	E	466/534 (87%)	-0.38	1 (0%)	92	92	6, 10, 18, 32	0
1	F	466/534 (87%)	-0.20	3 (0%)	85	85	7, 13, 26, 41	0
1	G	466/534 (87%)	-0.15	3 (0%)	85	85	9, 16, 28, 40	0
1	H	464/534 (86%)	-0.07	1 (0%)	92	92	9, 18, 31, 42	0
All	All	3745/4272 (87%)	-0.18	23 (0%)	85	85	5, 14, 26, 45	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	229	SER	4.2
1	E	231	GLY	3.7
1	C	230	ASP	3.4
1	F	376	ASN	3.3
1	A	376	ASN	3.3

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 oligosaccharides in this entry.

6.4 Ligands ⓘ

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, 95th 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^2)	Q<0.9
4	EDO	H	602	4/4	0.83	0.16	22,23,27,34	0
3	SO4	D	601	5/5	0.85	0.25	37,44,49,86	0
2	PEG	C	603	7/7	0.87	0.17	14,15,25,28	0
2	PEG	C	604	7/7	0.87	0.15	12,21,36,39	0
2	PEG	A	601	7/7	0.88	0.15	15,29,31,37	0
2	PEG	B	602	7/7	0.88	0.14	9,15,24,30	0
4	EDO	E	602	4/4	0.89	0.13	11,22,24,25	0
2	PEG	C	602	7/7	0.89	0.13	21,23,29,35	0
3	SO4	E	603	5/5	0.91	0.20	27,32,48,72	0
2	PEG	E	601	7/7	0.91	0.12	7,9,14,23	0
3	SO4	D	602	5/5	0.91	0.21	34,38,50,72	0
2	PEG	B	601	7/7	0.92	0.12	11,16,19,23	0
2	PEG	H	601	7/7	0.93	0.10	17,19,21,32	0
3	SO4	B	603	5/5	0.94	0.15	30,39,51,82	0
2	PEG	F	601	7/7	0.94	0.10	9,20,24,27	0
2	PEG	A	602	7/7	0.94	0.09	9,11,12,20	0
2	PEG	C	601	7/7	0.96	0.08	13,13,15,21	0

6.5 Other polymers ⓘ

There are no such residues in this entry.