wwPDB X-ray Structure Validation Summary Report (i)

Oct 15, 2023 – 04:05 PM EDT

PDB ID	:	2YZ7
Title	:	X-ray analyses of 3-hydroxybutyrate dehydrogenase from Alcaligenes faecalis
Authors	:	Hoque, M.M.; Juan, E.C.M.; Shimizu, S.; Hossain, M.T.; Takenaka, A.
Deposited on	:	2007-05-04
Resolution	:	2.19 Å(reported)

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

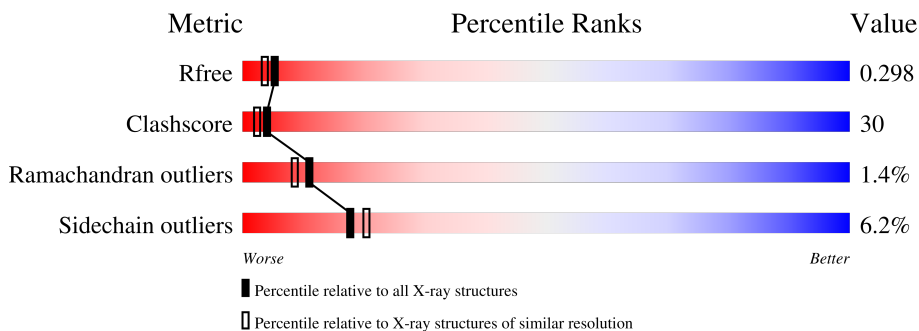|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.19 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	Percentile Ranks	Value
Rfree		0.298
Clashscore		30
Ramachandran outliers		1.4%
Sidechain outliers		6.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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	260	52% 46% •
1	B	260	56% 39% 5%
1	C	260	60% 37% •
1	D	260	55% 42% •
1	E	260	53% 41% 5%
1	F	260	57% 40% •
1	G	260	51% 45% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	260	52% 44%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 15660 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 D-3-hydroxybutyrate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	B	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	C	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	D	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	E	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	F	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	G	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			
1	H	260	Total	C	N	O	S	0	0	0
			1906	1203	331	366	6			

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ca	0	0
			1	1		
2	B	1	Total	Ca	0	0
			1	1		
2	E	1	Total	Ca	0	0
			1	1		
2	F	1	Total	Ca	0	0
			1	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Cl 1	0	0
3	B	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0
3	D	1	Total 1	Cl 1	0	0
3	E	1	Total 1	Cl 1	0	0
3	F	1	Total 1	Cl 1	0	0
3	G	1	Total 1	Cl 1	0	0
3	H	1	Total 1	Cl 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	43	Total 43	O 43	0	0
4	B	40	Total 40	O 40	0	0
4	C	76	Total 76	O 76	0	0
4	D	72	Total 72	O 72	0	0
4	E	42	Total 42	O 42	0	0
4	F	59	Total 59	O 59	0	0
4	G	39	Total 39	O 39	0	0
4	H	29	Total 29	O 29	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. 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: D-3-hydroxybutyrate dehydrogenase

Chain A:

- Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain B:

- Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain C:

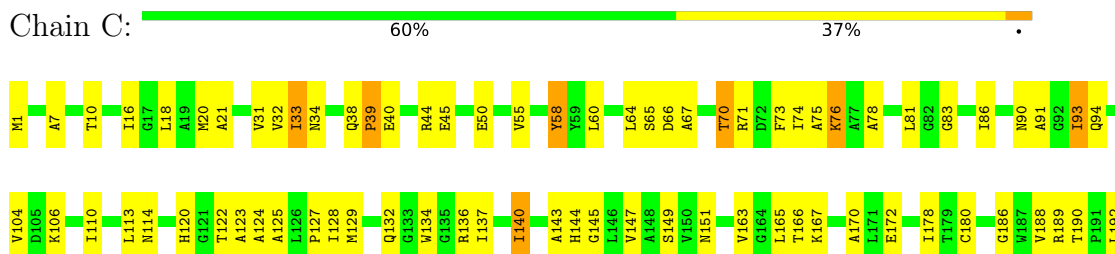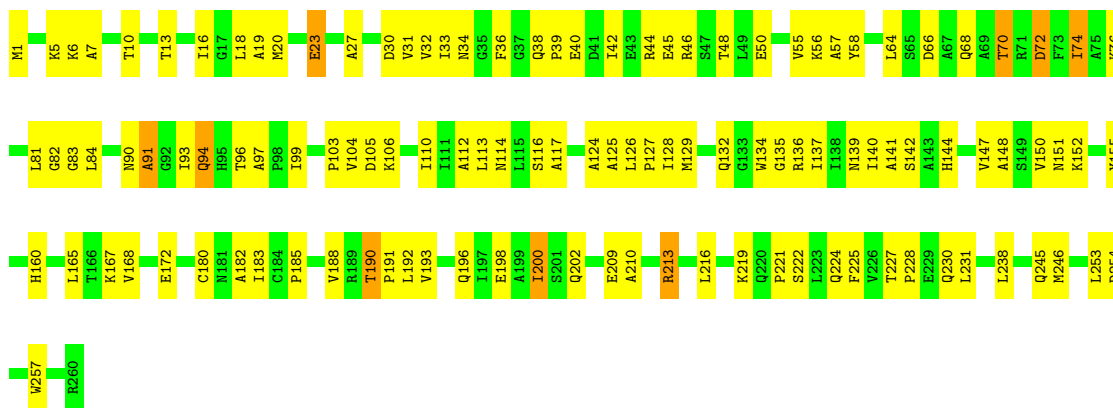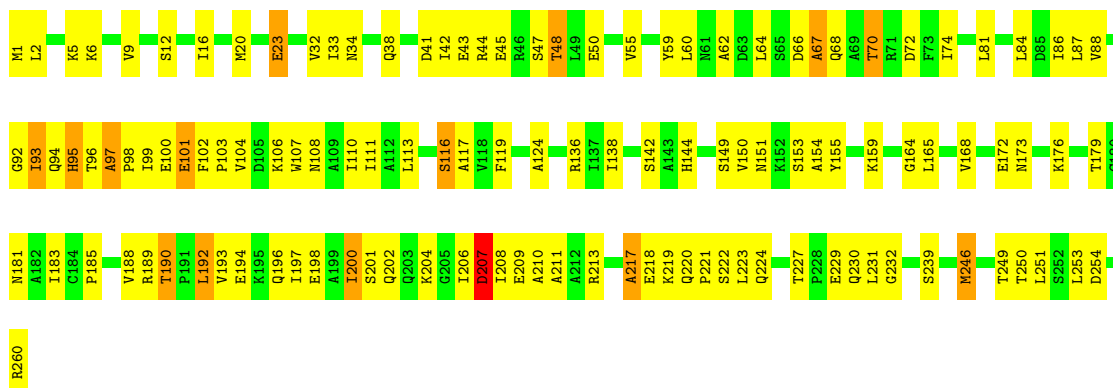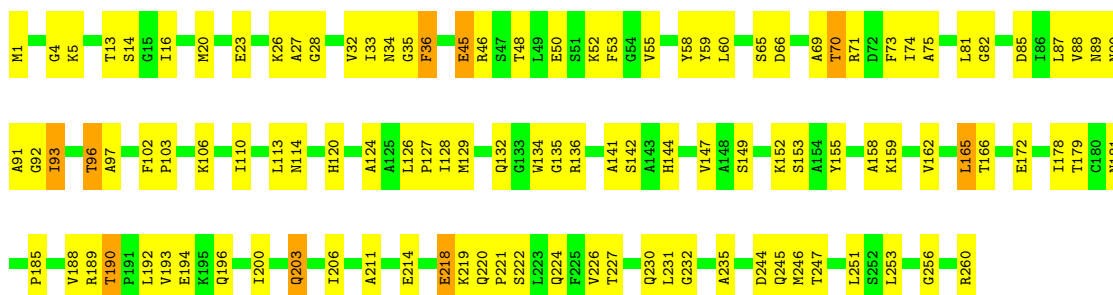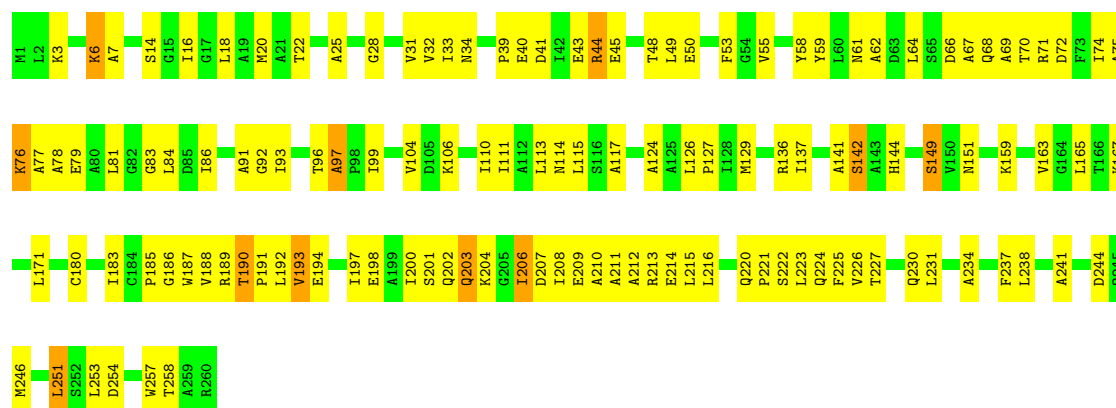V193 Q196 I197 E198 A199 I200 S201 Q202 Q203 K204 G205 I206 D207 I208 A211 A212 R213 E214 L215 L216 P221 S222 L223 Q224 F225 V226 T227 P228 L231 A234 F237 L238 M246 T247 L251 S252 L253 D254 G255 R260

• Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain D:

55%   42%   •

M1 K5 K6 A7 T10 T13 I16 G17 L18 A19 M20 E23 A27 D30 V31 V32 I33 N34 G35 F36 G37 Q38 P39 E40 D41 I42 E43 R44 E45 R46 S47 T48 L49 E50 V55 K56 A57 Y58 L64 S65 D66 A67 Q68 A69 T70 R71 D72 F73 I74 A75 K76

L81 G82 G83 L84 N90 A91 G92 I93 Q94 H95 T96 A97 P98 I99 P103 V104 D105 K106 I110 I111 A112 L113 N114 L115 S116 A117 A124 A125 L126 P127 I128 M129 Q132 G133 W134 G135 R136 I137 I138 N139 I140 A141 S142 A143 H144 V147 A148 S149 V150 N151 K152 Y155

H160 L165 T166 K167 V168 E172 C180 N181 A182 I183 C184 P185 V188 R189 T190 P191 L192 V193 Q196 I197 E198 A199 I200 S201 Q202 E209 A210 R213 L216 K219 Q220 P221 S222 L223 Q224 F225 V226 T227 P228 E229 Q230 L231 L238 Q245 M246 L253 D254

W257 R260

• Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain E:

53%   41%   5%

M1 L2 K5 K6 V9 S12 I16 M20 E23 V32 I33 N34 Q38 D41 I42 E43 R44 E45 R46 S47 T48 L49 E50 V55 Y59 L60 N61 A62 D63 L64 S65 D66 A67 Q68 A69 T70 R71 D72 F73 I74 L81 L84 D85 I86 L87 V88

G92 I93 Q94 H95 T96 A97 P98 I99 E100 E101 F102 P103 V104 D105 K106 W107 N108 A109 I110 I111 A112 L113 S116 A117 V118 F119 A124 R136 I137 I138 S142 A143 H144 S149 V150 N151 K152 S153 A154 Y155 K159 G164 L165 V168 E172 N173 K176 T179 C180

N181 A182 I183 C184 P185 V188 R189 T190 P191 L192 V193 E194 K195 Q196 I197 E198 A199 I200 S201 Q202 Q203 K204 G205 I206 D207 I208 E209 A210 A211 A212 R213 A217 E218 K219 Q220 P221 S222 L223 Q224 T227 P228 E229 Q230 L231 G232 S239 M246 T249 T250 L251 S252 L253 D254

R260

• Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain F:

57%   40%   •

M1 G4 K5 T13 S14 G15 I16 M20 E23 K26 A27 G28 V32 I33 N34 G35 F36 E45 R46 S47 T48 L49 E50 S51 K52 F53 G54 V55 Y58 Y59 L60 S65 D66 A69 T70 R71 D72 F73 I74 A75 L81 G82 D85 I86 L87 V88 N89 N90

A91 G92 I93 T96 A97 F102 P103 K106 I110 L113 N114 H120 A124 A125 L126 P127 I128 M129 Q132 G133 W134 G135 R136 A141 S142 A143 H144 V147 A148 S149 K152 S153 A154 Y155 A158 K159 V162 L165 T166 E172 I178 T179 C180 N181

P185 V188 R189 T190 P191 L192 V193 E194 K195 Q196 I200 Q203 I206 A211 E214 E218 K219 Q220 P221 S222 L223 Q224 F225 V226 T227 Q230 L231 G232 A235 D244 Q245 M246 T247 L251 S252 L253 G256 R260

• Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain G:

51%   45%

M1 L2 K3 K6 A7 S14 G15 I16 G17 L18 A19 M20 A21 T22 A25 G28 V31 V32 I33 N34 P39 E40 D41 I42 E43 R44 E45 T48 L49 E50 F53 G54 V55 Y58 Y59 L60 N61 A62 D63 L64 S65 D66 A67 Q68 A69 T70 R71 D72 F73 I74 A75

K76 A77 A78 E79 A80 L81 G82 G83 L84 D85 I86 A91 G92 I93 T96 A97 P98 I99 V104 D105 K106 I110 I111 A112 L113 N114 L115 S116 A117 A124 A125 L126 P127 I128 M129 R136 I137 A141 S142 A143 H144 S149 V150 N151 K159 V163 G164 L165 T166 K167

L171 C180 I183 C184 P185 G186 W187 V188 R189 T190 P191 L192 V193 E194 I197 E198 A199 I200 S201 Q202 Q203 K204 G205 I206 D207 I208 E209 A210 A211 A212 R213 E214 L215 L216 Q220 P221 S222 L223 Q224 F225 V226 T227 Q230 L231 A234 F237 L238 A241 D244 Q245

M246 L251 S252 L253 D254 W257 T258 A259 R260

• Molecule 1: D-3-hydroxybutyrate dehydrogenase

Chain H:

52%   44%

M1 L2 K3 K6 V9 S12 T13 I16 M20 L24 A27 G28 V32 I33 N34 G35 F36 G37 Q38 P39 E40 D41 I42 E43 R44 E45 T48 L49 E50 V55 Y58 Y59 D66 A69 T70 F73 I74 A75 L81 G82 L87 V88 N89 N90

I93 Q94 H95 T96 I99 E100 E101 F102 D105 K106 W107 N108 A109 I110 I111 A112 L113 N114 A117 V118 F119 H120 G121 T122 A123 A124 A125 L126 P127 I128 M129 Q130 K131 Q132 R136 I137 I140 A141 S142 S149 V150 N151 K152 V156 A157 A158 V162 L165 A174

T179 C180 I183 C184 P185 T190 P191 L192 V193 I197 E198 A199 I200 S201 Q202 Q203 K204 G205 I206 D207 I208 E209 A210 A211 A212 L215 L216 A217 E218 K219 Q220 P221 S222 L223 Q224 F225 V226 T227 Q230 L231 L238 D244 Q245 M246 L251 S252 L253 D254 G255 G256 W257

R260

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	77.30Å 118.83Å 118.73Å 90.00° 93.90° 90.00°	Depositor
Resolution (Å)	50.00 – 2.19 45.50 – 2.19	Depositor EDS
% Data completeness (in resolution range)	54.8 (50.00-2.19) 93.4 (45.50-2.19)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.84 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.197 , 0.270 0.239 , 0.298	Depositor DCC
R_{free} test set	10253 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.377	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 49.2	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	15660	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.81% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.37	0/1935	0.63	0/2623
1	B	0.38	0/1935	0.63	0/2623
1	C	0.40	0/1935	0.66	0/2623
1	D	0.43	0/1935	0.65	0/2623
1	E	0.37	0/1935	0.62	0/2623
1	F	0.41	0/1935	0.66	0/2623
1	G	0.36	0/1935	0.63	0/2623
1	H	0.36	0/1935	0.62	0/2623
All	All	0.39	0/15480	0.64	0/20984

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 ⓘ

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	1906	0	1941	121	0
1	B	1906	0	1941	131	0
1	C	1906	0	1941	90	0
1	D	1906	0	1941	112	0
1	E	1906	0	1941	129	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	1906	0	1941	108	0
1	G	1906	0	1941	137	0
1	H	1906	0	1941	143	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	1	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	1	0
3	F	1	0	0	1	0
3	G	1	0	0	0	0
3	H	1	0	0	1	0
4	A	43	0	0	3	0
4	B	40	0	0	3	0
4	C	76	0	0	4	0
4	D	72	0	0	4	0
4	E	42	0	0	6	0
4	F	59	0	0	2	0
4	G	39	0	0	0	0
4	H	29	0	0	1	0
All	All	15660	0	15528	926	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 30.

The worst 5 of 926 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:92:GLY:HA2	1:F:114:ASN:HD21	1.13	1.07
1:E:190:THR:HG22	1:E:193:VAL:HG23	1.37	1.06
1:G:200:ILE:HG23	1:G:211:ALA:HB1	1.37	1.06
1:E:16:ILE:HD11	1:E:188:VAL:HG11	1.40	1.04
1:B:204:LYS:HB2	1:B:206:ILE:HD12	1.44	0.99

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	258/260 (99%)	232 (90%)	22 (8%)	4 (2%)	9	7
1	B	258/260 (99%)	236 (92%)	20 (8%)	2 (1%)	19	19
1	C	258/260 (99%)	243 (94%)	12 (5%)	3 (1%)	13	10
1	D	258/260 (99%)	229 (89%)	27 (10%)	2 (1%)	19	19
1	E	258/260 (99%)	226 (88%)	24 (9%)	8 (3%)	4	2
1	F	258/260 (99%)	241 (93%)	16 (6%)	1 (0%)	34	37
1	G	258/260 (99%)	229 (89%)	24 (9%)	5 (2%)	8	5
1	H	258/260 (99%)	227 (88%)	28 (11%)	3 (1%)	13	10
All	All	2064/2080 (99%)	1863 (90%)	173 (8%)	28 (1%)	11	8

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	ALA
1	B	208	ILE
1	E	67	ALA
1	E	95	HIS
1	E	97	ALA

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	192/192 (100%)	187 (97%)	5 (3%)	46	58

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	192/192 (100%)	177 (92%)	15 (8%)	12	13
1	C	192/192 (100%)	174 (91%)	18 (9%)	8	8
1	D	192/192 (100%)	181 (94%)	11 (6%)	20	24
1	E	192/192 (100%)	178 (93%)	14 (7%)	14	15
1	F	192/192 (100%)	183 (95%)	9 (5%)	26	33
1	G	192/192 (100%)	181 (94%)	11 (6%)	20	24
1	H	192/192 (100%)	179 (93%)	13 (7%)	16	17
All	All	1536/1536 (100%)	1440 (94%)	96 (6%)	18	20

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	190	THR
1	F	218	GLU
1	E	200	ILE
1	F	70	THR
1	G	76	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	173	ASN
1	F	139	ASN
1	H	203	GLN
1	E	196	GLN
1	F	89	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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 12 are monoatomic - leaving 0 for Mogul analysis.

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 ⓘ

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 ⓘ

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.