Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 02:42 am BST

PDB ID	:	4I04
Title	:	Structure of zymogen of cathepsin B1 from Schistosoma mansoni
Authors	:	Rezacova, P.; Jilkova, A.; Brynda, J.; Horn, M.; Mares, M.
Deposited on	:	2012-11-16
Resolution	:	1.95 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

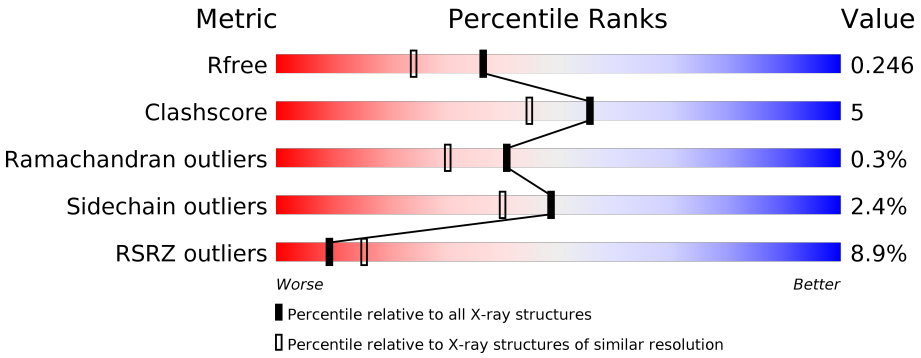
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance

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

The reported resolution of this entry is 1.95 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}	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 <=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	323	16% 80% 12% • 6%
1	B	323	6% 88% 7% •
1	C	323	4% 85% 9% • 5%
1	D	323	8% 84% 11% 5%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	EDO	D	401	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10371 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.

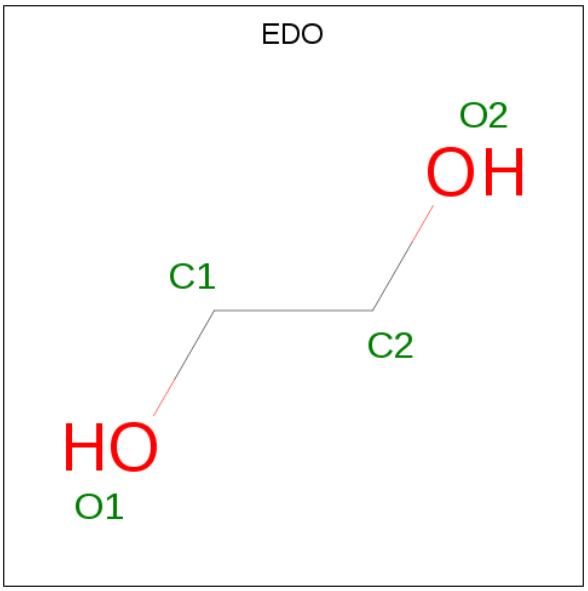
- Molecule 1 is a protein called Cathepsin B-like peptidase (C01 family).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	1	0
			2428	1515	437	461	15			
1	B	310	Total	C	N	O	S	0	2	0
			2482	1547	446	473	16			
1	C	306	Total	C	N	O	S	0	3	0
			2455	1531	441	468	15			
1	D	308	Total	C	N	O	S	0	1	0
			2463	1535	443	470	15			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	100	SER	CYS	ENGINEERED MUTATION	UNP Q8MNY2
A	168	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
A	283	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
B	100	SER	CYS	ENGINEERED MUTATION	UNP Q8MNY2
B	168	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
B	283	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
C	100	SER	CYS	ENGINEERED MUTATION	UNP Q8MNY2
C	168	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
C	283	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
D	100	SER	CYS	ENGINEERED MUTATION	UNP Q8MNY2
D	168	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2
D	283	ALA	THR	ENGINEERED MUTATION	UNP Q8MNY2

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

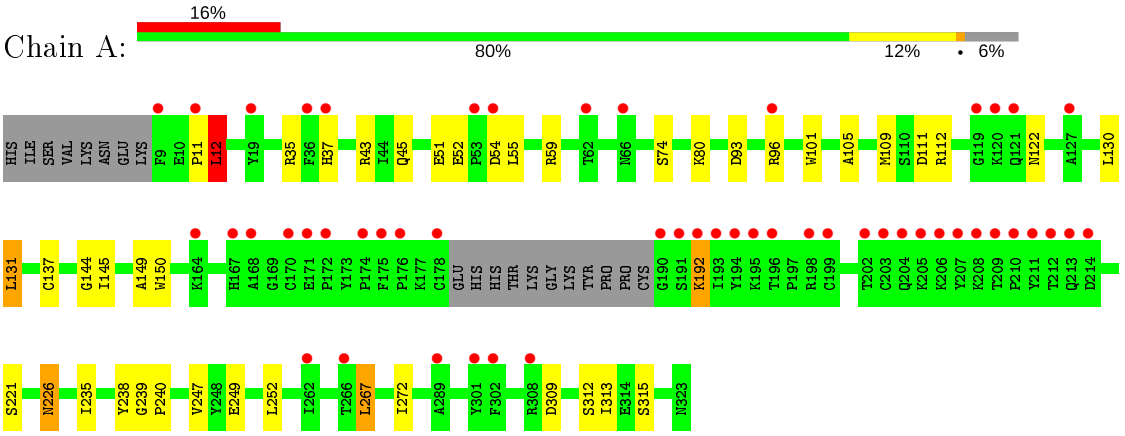
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	84	Total O 84 84	0	0
3	B	137	Total O 137 137	0	0
3	C	172	Total O 172 172	0	0
3	D	126	Total O 126 126	0	0

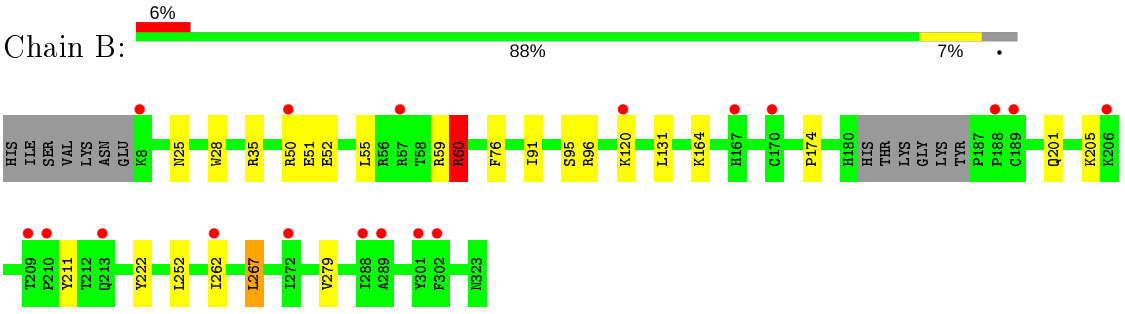
3 Residue-property plots (i)

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

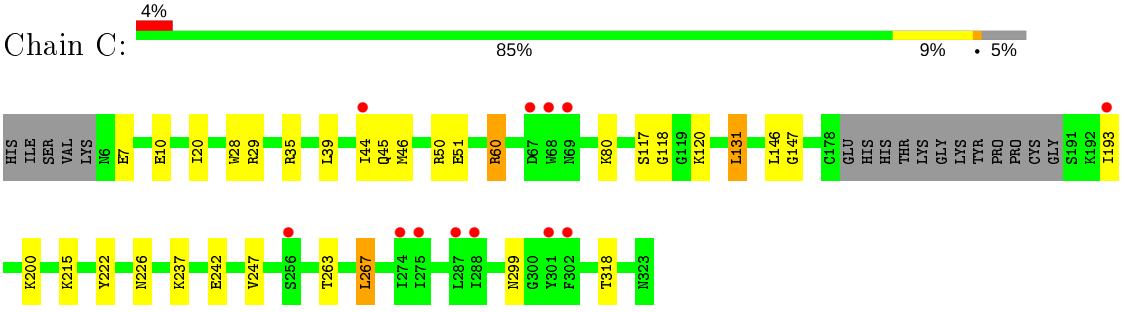
- Molecule 1: Cathepsin B-like peptidase (C01 family)



- Molecule 1: Cathepsin B-like peptidase (C01 family)

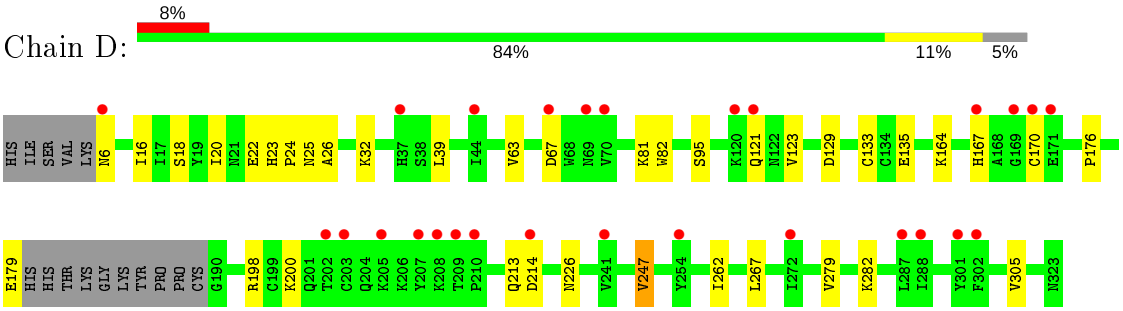


- Molecule 1: Cathepsin B-like peptidase (C01 family)



- Molecule 1: Cathepsin B-like peptidase (C01 family)

Chain D:


8%
84%
11%
5%
HIS ILE SER VAL LYS N6 I16 I17 S18 Y19 I20 N21 E22 H23 P24 N25 A26 K32 H37 S38 L39 I44 V63 D67 W68 N69 V70 K81 W82 S95 K120 Q121 N122 V123 D129 C133 C134 E135 K164 H167 A168 G169 C170 E171 P176
E179 HIS HIS THR LYS GLY LYS TYR PRO PRO CYS G190 R198 C199 K200 Q201 T202 C203 Q204 K205 K206 Y207 K208 T209 P210 Q213 D214 N226 V241 V247 Y254 I262 L267 I272 V279 K282 L287 I288 Y301 F302 V305 N323

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	54.28Å 83.17Å 84.57Å 76.97° 86.60° 71.11°	Depositor
Resolution (Å)	38.40 – 1.95 38.40 – 1.95	Depositor EDS
% Data completeness (in resolution range)	75.3 (38.40-1.95) 75.4 (38.40-1.95)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.29 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.3.0037	Depositor
R, R_{free}	0.199 , 0.252 0.195 , 0.246	Depositor DCC
R_{free} test set	3756 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	23.9	Xtriage
Anisotropy	0.177	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 47.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10371	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.36% 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: EDO

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	0.52	0/2491	0.62	1/3359 (0.0%)
1	B	0.59	0/2551	0.67	2/3440 (0.1%)
1	C	0.63	0/2524	0.69	1/3403 (0.0%)
1	D	0.58	0/2526	0.65	0/3405
All	All	0.58	0/10092	0.66	4/13607 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	60	ARG	NE-CZ-NH2	-7.89	116.36	120.30
1	C	131	LEU	CA-CB-CG	6.22	129.60	115.30
1	B	60	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	A	12	LEU	CA-CB-CG	5.07	126.97	115.30

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	2428	0	2323	35	0
1	B	2482	0	2366	20	0
1	C	2455	0	2353	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2463	0	2354	22	0
2	A	8	0	12	1	0
2	B	12	0	18	5	0
2	D	4	0	6	4	0
3	A	84	0	0	2	0
3	B	137	0	0	4	0
3	C	172	0	0	7	0
3	D	126	0	0	5	0
All	All	10371	0	9432	98	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 5.

All (98) 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:A:51:GLU:HG3	1:A:145[A]:ILE:HD11	1.30	1.09
1:B:91:ILE:H	2:B:402:EDO:H11	1.28	0.96
1:A:51:GLU:CG	1:A:145[A]:ILE:HD11	2.08	0.83
1:D:20:ILE:HD12	1:D:262[B]:ILE:HD11	1.59	0.82
1:C:60:ARG:HG3	1:C:222:TYR:HA	1.61	0.82
1:C:29:ARG:HD2	3:C:521:HOH:O	1.81	0.81
1:C:80:LYS:HE3	3:C:500:HOH:O	1.82	0.80
1:A:80:LYS:HB2	2:A:401:EDO:H22	1.63	0.79
1:C:117:SER:O	1:C:120:LYS:HG2	1.84	0.76
1:D:282:LYS:H	2:D:401:EDO:C2	1.98	0.75
1:A:55:LEU:HB3	1:A:59:ARG:HE	1.52	0.74
1:A:137:CYS:O	1:A:145[A]:ILE:HG22	1.88	0.73
1:D:282:LYS:H	2:D:401:EDO:H22	1.52	0.72
1:C:146:LEU:HD13	1:C:242:GLU:HG2	1.69	0.72
1:B:91:ILE:H	2:B:402:EDO:C1	2.04	0.71
1:C:46:MET:SD	1:C:247[B]:VAL:HG11	2.34	0.68
1:C:215:LYS:HE2	3:C:485:HOH:O	1.94	0.66
1:A:43:ARG:HG2	1:A:267:LEU:CD2	2.27	0.64
1:B:205:LYS:HE3	3:B:621:HOH:O	1.97	0.64
1:D:32:LYS:HG3	3:D:560:HOH:O	1.97	0.63
1:B:28:TRP:HE1	1:B:262:ILE:HD12	1.63	0.63
1:A:43:ARG:HG2	1:A:267:LEU:HD22	1.81	0.62
1:C:50:ARG:NH1	3:C:535:HOH:O	2.33	0.62
1:B:91:ILE:N	2:B:402:EDO:H11	2.09	0.61
1:A:192:LYS:HD2	1:A:192:LYS:H	1.66	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:252:LEU:HD11	1:B:267:LEU:HD21	1.83	0.59
1:B:51:GLU:OE1	1:B:60:ARG:NH2	2.32	0.58
1:D:39:LEU:HD22	1:D:267:LEU:HD21	1.85	0.57
1:D:279:VAL:HG12	2:D:401:EDO:H22	1.86	0.57
1:A:35:ARG:NE	1:A:45:GLN:OE1	2.36	0.57
1:D:18:SER:O	1:D:22:GLU:HG2	2.05	0.57
1:A:111:ASP:OD2	1:A:240:PRO:HG2	2.05	0.56
1:A:11:PRO:O	1:A:12:LEU:HD12	2.07	0.55
1:B:60:ARG:HG3	1:B:222:TYR:HA	1.88	0.54
1:B:76:PHE:CE1	2:B:403:EDO:H21	2.42	0.54
1:D:129:ASP:O	1:D:133:CYS:HB2	2.08	0.54
1:B:76:PHE:CD1	2:B:403:EDO:H21	2.42	0.54
1:B:205:LYS:CE	3:B:621:HOH:O	2.53	0.54
1:A:54:ASP:OD1	1:A:55:LEU:N	2.41	0.54
1:C:46:MET:SD	1:C:247[A]:VAL:HG21	2.49	0.52
1:A:226:ASN:ND2	1:A:315:SER:HA	2.25	0.51
1:A:150:TRP:HB3	1:A:221:SER:HB2	1.91	0.51
1:C:118:GLY:HA2	3:C:458:HOH:O	2.10	0.51
1:C:44:ILE:HD13	1:C:193:ILE:HG12	1.93	0.50
1:D:305:VAL:HG22	3:D:555:HOH:O	2.10	0.50
1:A:35:ARG:NH2	1:A:93:ASP:OD1	2.38	0.50
1:A:101:TRP:CD1	1:A:144:GLY:HA3	2.47	0.50
1:C:39:LEU:HD22	1:C:267:LEU:HD11	1.93	0.50
1:D:170:CYS:HB2	1:D:214:ASP:OD2	2.12	0.49
1:D:23:HIS:ND1	1:D:24:PRO:HD2	2.27	0.49
1:C:299:ASN:HB2	3:C:490:HOH:O	2.13	0.49
1:A:249:GLU:O	1:A:252:LEU:HD13	2.13	0.49
1:D:164:LYS:O	1:D:167:HIS:CE1	2.66	0.48
1:B:28:TRP:NE1	1:B:262:ILE:HD12	2.27	0.48
1:B:279:VAL:HG12	2:D:401:EDO:H12	1.95	0.48
1:D:164:LYS:O	1:D:167:HIS:ND1	2.47	0.48
1:B:164:LYS:HD2	1:B:174:PRO:O	2.14	0.48
1:D:20:ILE:CD1	1:D:262[B]:ILE:HD11	2.37	0.48
1:C:51:GLU:OE1	1:C:60:ARG:NH2	2.41	0.47
1:C:20:ILE:HD13	1:C:28:TRP:CH2	2.48	0.47
1:C:35:ARG:NE	1:C:45:GLN:OE1	2.48	0.47
1:C:147:GLY:HA2	1:C:318:THR:HG21	1.97	0.46
1:D:63:VAL:HA	3:D:533:HOH:O	2.15	0.46
1:D:67:ASP:N	1:D:67:ASP:OD1	2.47	0.46
1:A:43:ARG:CG	1:A:267:LEU:CD2	2.93	0.46
1:A:112:ARG:HG2	1:A:238:TYR:O	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:96:ARG:NH1	3:A:583:HOH:O	2.50	0.45
1:C:46:MET:CE	1:C:247[B]:VAL:CG1	2.95	0.45
1:A:252:LEU:HD11	1:A:267:LEU:HD12	1.97	0.45
1:A:55:LEU:HD13	1:A:59:ARG:HH21	1.82	0.44
1:B:201:GLN:HA	1:B:211:TYR:CZ	2.52	0.44
1:C:237:LYS:HE2	3:C:447:HOH:O	2.17	0.44
1:B:52:GLU:HB3	1:B:55:LEU:HD12	2.00	0.44
1:D:81:LYS:HD3	1:D:82:TRP:CZ2	2.52	0.44
1:D:6:ASN:HB3	3:D:578:HOH:O	2.17	0.44
1:A:272:ILE:HD11	1:A:313:ILE:HD11	1.99	0.44
1:C:44:ILE:O	1:C:44:ILE:HG22	2.18	0.44
1:A:51:GLU:HG3	1:A:145[A]:ILE:CD1	2.22	0.44
1:D:247:VAL:HA	3:D:501:HOH:O	2.17	0.44
1:B:50:ARG:HD2	3:B:576:HOH:O	2.19	0.43
1:A:309:ASP:OD1	1:A:312:SER:HA	2.18	0.43
1:A:130:LEU:HD21	1:A:149:ALA:HB1	2.00	0.43
1:A:101:TRP:CG	1:A:144:GLY:HA3	2.53	0.43
1:A:101:TRP:CZ2	1:A:131:LEU:HA	2.53	0.43
1:A:105:ALA:O	1:A:109:MET:HG3	2.19	0.42
1:A:235:ILE:HA	1:A:239:GLY:O	2.19	0.42
1:C:7:GLU:O	1:C:263:THR:OG1	2.38	0.42
1:A:252:LEU:N	1:A:252:LEU:HD12	2.35	0.41
1:D:16:ILE:O	1:D:20:ILE:HG12	2.19	0.41
1:D:95:SER:HB3	1:D:176:PRO:O	2.20	0.41
1:A:101:TRP:CH2	1:A:131:LEU:HA	2.55	0.41
1:B:55:LEU:O	1:B:59:ARG:HB2	2.20	0.41
1:A:52:GLU:HB3	1:A:55:LEU:HD12	2.02	0.41
1:B:25:ASN:HB2	3:B:572:HOH:O	2.21	0.40
1:A:252:LEU:N	1:A:252:LEU:CD1	2.84	0.40
1:A:37:HIS:CD2	3:A:576:HOH:O	2.73	0.40
1:B:95:SER:OG	1:B:96:ARG:N	2.54	0.40
1:D:26:ALA:HB1	1:D:262[A]:ILE:HD13	2.04	0.40

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	301/323 (93%)	287 (95%)	13 (4%)	1 (0%)	41 30
1	B	308/323 (95%)	289 (94%)	19 (6%)	0	100 100
1	C	305/323 (94%)	294 (96%)	10 (3%)	1 (0%)	41 30
1	D	305/323 (94%)	289 (95%)	14 (5%)	2 (1%)	22 11
All	All	1219/1292 (94%)	1159 (95%)	56 (5%)	4 (0%)	41 30

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	25	ASN
1	A	226	ASN
1	C	226	ASN
1	D	226	ASN

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	260/277 (94%)	253 (97%)	7 (3%)	44 34
1	B	267/277 (96%)	262 (98%)	5 (2%)	57 50
1	C	265/277 (96%)	260 (98%)	5 (2%)	57 50
1	D	264/277 (95%)	256 (97%)	8 (3%)	41 30
All	All	1056/1108 (95%)	1031 (98%)	25 (2%)	49 40

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	LEU
1	A	74	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	122	ASN
1	A	131	LEU
1	A	192	LYS
1	A	247	VAL
1	A	267	LEU
1	B	35	ARG
1	B	60	ARG
1	B	120	LYS
1	B	131	LEU
1	B	267	LEU
1	C	10	GLU
1	C	60	ARG
1	C	131	LEU
1	C	200	LYS
1	C	267	LEU
1	D	121	GLN
1	D	123	VAL
1	D	135	GLU
1	D	179	GLU
1	D	198	ARG
1	D	200	LYS
1	D	213	GLN
1	D	247	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	EDO	D	401	-	3,3,3	0.24	0	2,2,2	0.60	0
2	EDO	B	401	-	3,3,3	0.39	0	2,2,2	0.67	0
2	EDO	B	402	-	3,3,3	0.42	0	2,2,2	0.43	0
2	EDO	A	402	-	3,3,3	0.27	0	2,2,2	0.92	0
2	EDO	B	403	-	3,3,3	0.29	0	2,2,2	0.48	0
2	EDO	A	401	-	3,3,3	0.30	0	2,2,2	0.85	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	EDO	D	401	-	-	1/1/1/1	-
2	EDO	B	401	-	-	0/1/1/1	-
2	EDO	B	402	-	-	0/1/1/1	-
2	EDO	A	402	-	-	1/1/1/1	-
2	EDO	B	403	-	-	0/1/1/1	-
2	EDO	A	401	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	401	EDO	O1-C1-C2-O2
2	A	402	EDO	O1-C1-C2-O2
2	A	401	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	401	EDO	4	0
2	B	402	EDO	3	0
2	B	403	EDO	2	0
2	A	401	EDO	1	0

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

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, 95^{th} 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	304/323 (94%)	0.82	52 (17%) 1 2	12, 33, 60, 80	1 (0%)
1	B	310/323 (95%)	0.22	18 (5%) 23 31	3, 20, 43, 60	0
1	C	306/323 (94%)	0.17	12 (3%) 39 49	4, 17, 36, 54	0
1	D	308/323 (95%)	0.47	27 (8%) 10 16	7, 24, 54, 72	0
All	All	1228/1292 (95%)	0.42	109 (8%) 9 15	3, 23, 51, 80	1 (0%)

All (109) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	167	HIS	6.4
1	A	203	CYS	5.8
1	A	168	ALA	5.4
1	A	262	ILE	5.3
1	A	190	GLY	5.2
1	A	207	TYR	4.9
1	A	19	TYR	4.8
1	D	208	LYS	4.7
1	D	69	ASN	4.5
1	A	209	THR	4.3
1	A	176	PRO	4.0
1	C	69	ASN	4.0
1	A	127	ALA	4.0
1	B	188	PRO	4.0
1	A	208	LYS	3.9
1	D	205	LYS	3.9
1	D	207	TYR	3.9
1	A	204	GLN	3.8
1	A	205	LYS	3.8
1	B	189	CYS	3.7
1	B	167	HIS	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	210	PRO	3.6
1	A	171	GLU	3.6
1	A	199	CYS	3.5
1	A	195	LYS	3.5
1	A	170	CYS	3.5
1	A	175	PHE	3.5
1	A	9	PHE	3.4
1	D	203	CYS	3.4
1	A	198	ARG	3.3
1	C	193	ILE	3.3
1	B	170	CYS	3.3
1	A	289	ALA	3.3
1	A	194	TYR	3.3
1	A	212	THR	3.3
1	A	37	HIS	3.3
1	D	209	THR	3.2
1	A	54	ASP	3.2
1	A	206	LYS	3.1
1	B	262	ILE	3.1
1	C	67	ASP	3.0
1	D	170	CYS	3.0
1	D	120	LYS	3.0
1	A	120	LYS	3.0
1	D	167	HIS	2.9
1	A	178	CYS	2.9
1	A	211	TYR	2.8
1	A	191	SER	2.7
1	A	36	PHE	2.7
1	D	67	ASP	2.7
1	D	272	ILE	2.7
1	D	171	GLU	2.7
1	B	210	PRO	2.7
1	B	209	THR	2.7
1	D	70	VAL	2.6
1	A	62	THR	2.6
1	A	202	THR	2.6
1	A	193	ILE	2.6
1	A	96	ARG	2.6
1	A	213	GLN	2.6
1	A	174	PRO	2.6
1	C	256	SER	2.6
1	C	287	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	119	GLY	2.6
1	A	121	GLN	2.5
1	A	192	LYS	2.5
1	B	8	LYS	2.5
1	D	287	LEU	2.5
1	A	53	PRO	2.4
1	B	302	PHE	2.4
1	C	274	ILE	2.4
1	C	288	ILE	2.4
1	D	37	HIS	2.4
1	A	266	THR	2.3
1	A	210	PRO	2.3
1	C	301	TYR	2.3
1	B	288	ILE	2.3
1	A	302	PHE	2.2
1	A	164	LYS	2.2
1	D	202	THR	2.2
1	B	57	ARG	2.2
1	D	302	PHE	2.2
1	D	301	TYR	2.2
1	C	44	ILE	2.2
1	A	308	ARG	2.2
1	B	289	ALA	2.2
1	D	254	TYR	2.2
1	A	214	ASP	2.2
1	C	302	PHE	2.2
1	C	275	ILE	2.1
1	D	44	ILE	2.1
1	D	288	ILE	2.1
1	D	214	ASP	2.1
1	A	11	PRO	2.1
1	C	68	TRP	2.1
1	A	172	PRO	2.1
1	B	301	TYR	2.1
1	A	66	ASN	2.1
1	D	6	ASN	2.1
1	D	241	VAL	2.1
1	D	121	GLN	2.1
1	D	169	GLY	2.1
1	A	196	THR	2.0
1	B	213	GLN	2.0
1	B	206	LYS	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	272	ILE	2.0
1	A	301	TYR	2.0
1	B	120	LYS	2.0
1	B	50	ARG	2.0

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 carbohydrates in this entry.

6.4 Ligands (i)

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, 95^{th} 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
2	EDO	B	402	4/4	0.86	0.17	28,30,31,33	0
2	EDO	A	402	4/4	0.87	0.23	28,29,31,32	0
2	EDO	B	403	4/4	0.87	0.18	13,19,21,29	0
2	EDO	D	401	4/4	0.89	0.20	13,16,17,26	0
2	EDO	B	401	4/4	0.93	0.27	22,25,26,28	0
2	EDO	A	401	4/4	0.96	0.13	17,21,21,28	0

6.5 Other polymers (i)

There are no such residues in this entry.