W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report ⓘ

May 29, 2020 – 12:21 am BST

PDB ID	:	2MSN
Title	:	NMR structure of a putative phosphoglycolate phosphatase (NP_346487.1) from Streptococcus pneumoniae TIGR4
Authors	:	Jaudzems, K.; Serrano, P.; Pedrini, B.; Geralt, M.; Wuthrich, K.; Joint Center for Structural Genomics (JCSG)
Deposited on	:	2014-08-04

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

|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
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 (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 74%.

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.
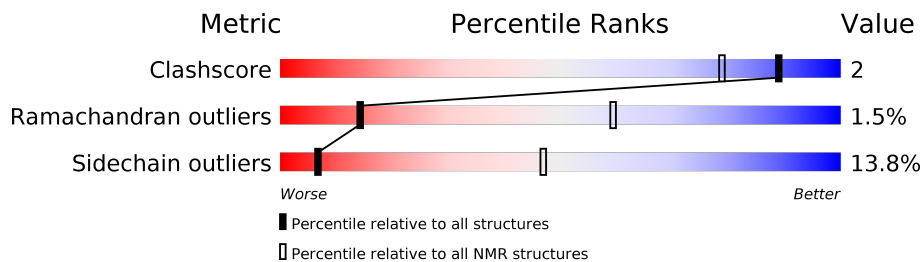
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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	208	

2 Ensemble composition and analysis ⓘ

This entry contains 20 models. Model 1 is the overall representative, medoid model (most similar to other models).

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:1-A:17, A:84-A:204 (138)	0.39	1
2	A:19-A:73 (55)	0.32	18

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 10, 12, 14, 16, 18, 20
2	1, 3, 5, 6, 8, 19
3	4, 9, 13, 17
4	11, 15
Single-model clusters	7

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3302 atoms, of which 1628 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Hydrolase, haloacid dehalogenase-like family.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	208	3302	1066	1628	276	330	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP Q97NG6
A	-1	HIS	-	EXPRESSION TAG	UNP Q97NG6

4 Residue-property plots ⓘ

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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 outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

• Molecule 1: Hydrolase, haloacid dehalogenase-like family
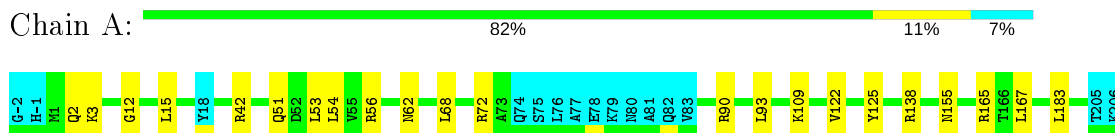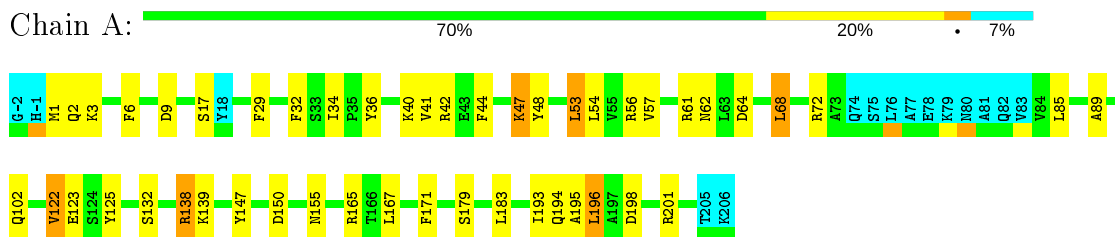

4.2 Scores per residue for each member of the ensemble

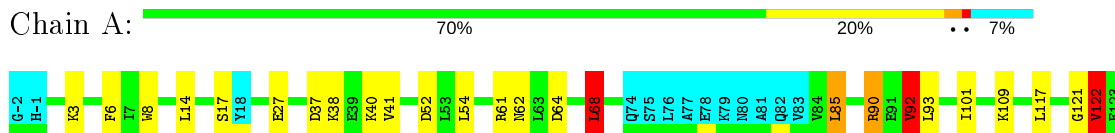Colouring as in section 4.1 above.

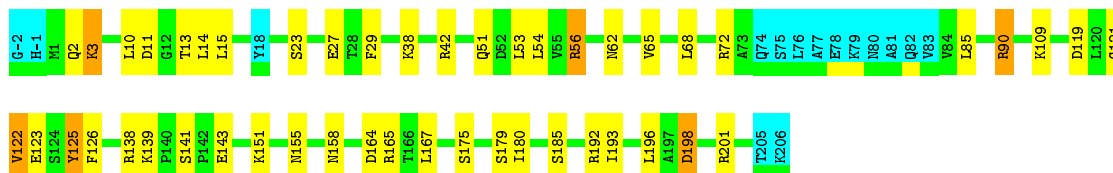### 4.2.1 Score per residue for model 1 (medoid)

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.2 Score per residue for model 2

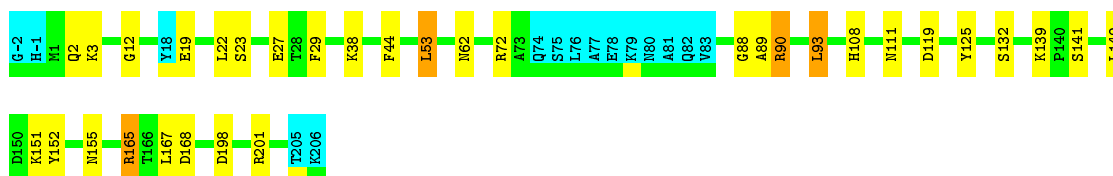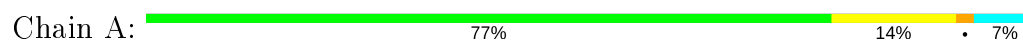• Molecule 1: Hydrolase, haloacid dehalogenase-like family

S124 Y125 F126 T127 E128 T131 R138 S141 D150 K151 Y152 N155 T159 I162 R165 T166 L167 E170 S175 S179 L183 E184 S185 I193 Q194 A195 D198 T205 K206

4.2.3 Score per residue for model 3

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 70%   20%   •   7%

G-2 H-1 M1 Q2 K3 L10 D11 G12 T13 L14 L15 Y18 S23 E27 T28 F29 K38 R42 Q51 D52 L53 L54 V55 R56 N62 V65 L68 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 V84 L85 R90 K109 D119 L120 G121

V122 E123 S124 Y125 F126 R138 K139 P140 S141 P142 E143 K151 N155 N158 D164 R165 T166 L167 S175 S179 I180 S185 R192 I193 L196 A197 D198 R201 T205 K206

4.2.4 Score per residue for model 4

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 77%   14%   •   7%

G-2 H-1 M1 Q2 K3 G12 Y18 E19 L22 S23 E27 T28 F29 K38 F44 L53 N62 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 G88 A89 R90 L93 H108 N111 D119 Y125 S132 K139 P140 S141 L149

D150 K151 Y152 N155 R165 T166 L167 D168 D198 R201 T205 K206

4.2.5 Score per residue for model 5

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 75%   17%   •   7%

G-2 H-1 M1 Q2 K3 W8 D9 G12 T13 L14 Y18 F29 A30 Q31 F32 D37 K38 R42 E43 F44 Y48 S49 V50 Q51 D52 L53 L54 D60 R61 N62 L68 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 V84 L85 Y106 N111 K118

E123 R138 D157 R165 T166 L167 N181 N190 H191 R192 L196 T205 K206
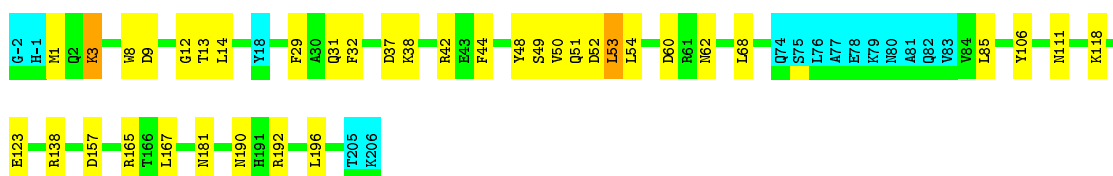
4.2.6 Score per residue for model 6

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 80%   12%   7%

G-2 H-1 K3 L10 D16 S17 Y18 T28 F29 A30 Q31 F32 S33 K38 F44 D52 L53 N62 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 V84 N111 N112 D119 E128 D150 K151 Y152 N155 S156 D157 R165 S175

N181 F182 L183 R192 T205 K206
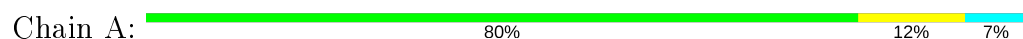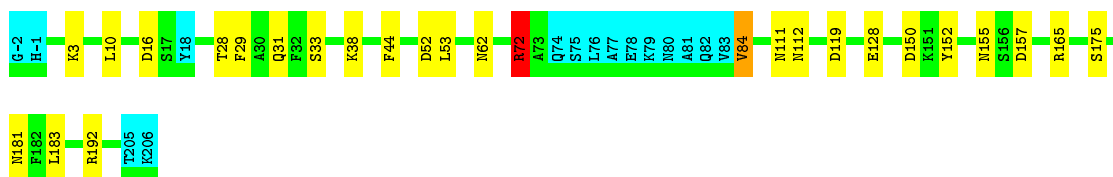
4.2.7 Score per residue for model 7

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 78% 13% 7%

G-2 H-1 M1 Q2 K3 F6 D9 L10 D11 G12 L15 Y18 S33 V50 Q51 L54 R61 N62 L68 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 L93 E128 S132 F136 V137 R138 S141 L149 N155 L167 D168 V169

S185 N190 H191 R192 D198 R201 T205 K206
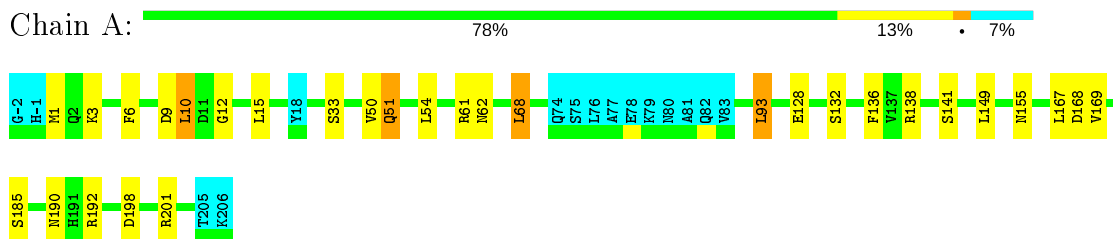
4.2.8 Score per residue for model 8

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 76% 14% 7%

G-2 H-1 M1 Q2 K3 G12 L15 D16 S17 Y18 Y36 D37 R42 I45 D52 R56 D60 R61 N62 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 S99 Q102 K109 G110 N111 L117 Y125 E128 R138 K139 D150 K151

Y152 Q153 L154 D157 Y161 R165 T166 L167 L183 Y187 R192 T205 K206
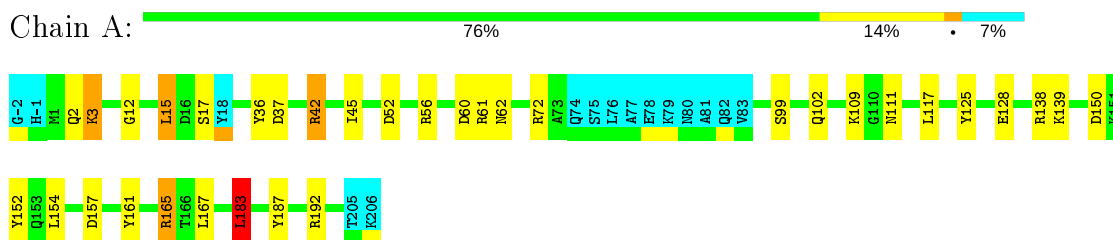
4.2.9 Score per residue for model 9

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 75% 15% 7%

G-2 H-1 M1 Q2 K3 W8 L14 Y18 S23 F29 S33 K38 R42 Q51 D52 L53 L54 V55 R56 R61 N62 L68 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 R90 L93 E98 D119 V122 E123 S124 Y125 F126 R138

L149 D150 K151 Y152 S156 D164 L167 F171 F182 N190 Q194 T205 K206

4.2.10 Score per residue for model 10

- Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 75% 17% 7%

G-2 H-1 K3 F6 I7 W8 D9 L10 D11 G12 Y18 E19 Y36 K40 V41 Q51 D52 L53 R56 D60 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 V84 L85 G88 V92 L93 K109 N112 L149 D150 K151 Y152 N155 I162

R165 T166 L167 S179 L183 Y187 E188 G189 N190 T205 K206

4.2.11 Score per residue for model 11

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 77% 14% • 7%

G-2 H-1 M1 Q2 G12 L15 Y18 E27 D37 F46 S49 L53 R56 V57 N62 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 V84 L85 R90 E91 V92 L93 K109 G110 N111 N112 K118 E123 F126 L149 D150 K151

Y152 S156 Y161 D164 R165 T166 L167 F171 R192 T205 K206

4.2.12 Score per residue for model 12

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 80% 13% 7%

G-2 H-1 K3 F6 G12 T13 L14 L15 Y18 E27 Y36 F46 Q51 R56 N62 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 M86 V92 L93 N111 E123 R138 K139 D150 K151 S156 T166 L167 D168 S179

L196 A197 D198 T205 K206

4.2.13 Score per residue for model 13

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

Chain A: 79% 11% •• 7%

G-2 H-1 M1 Q2 K3 L10 D11 G12 Y18 F29 F32 F44 S49 V50 Q51 D52 L53 N62 L63 D64 L68 R72 A73 Q74 S75 L76 A77 E78 K79 N80 A81 Q82 V83 K109 D119 V122 E123 S124 Y125 E128 S132 K139 P140 L149

D150 K151 L167 S179 L183 N190 T205 K206

4.2.14 Score per residue for model 14

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.15 Score per residue for model 15
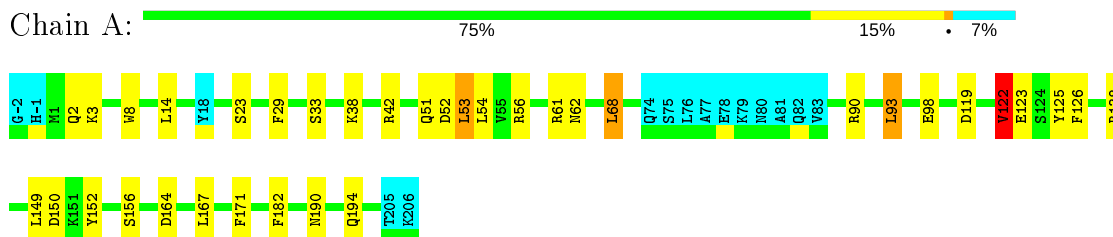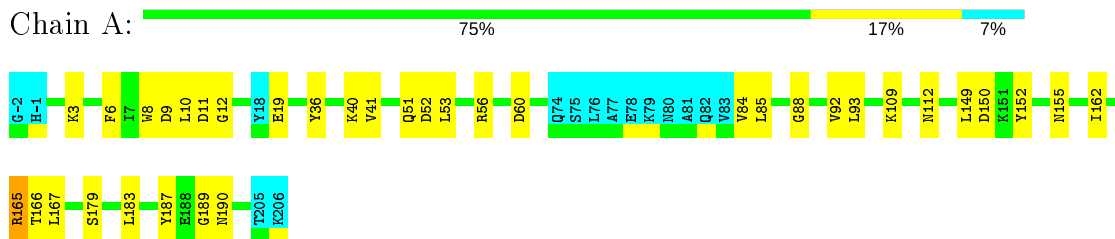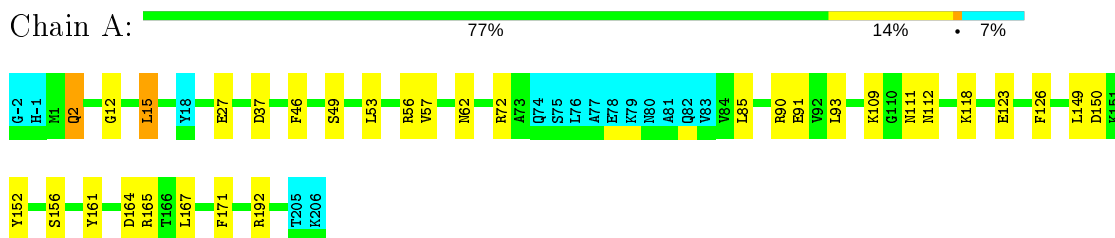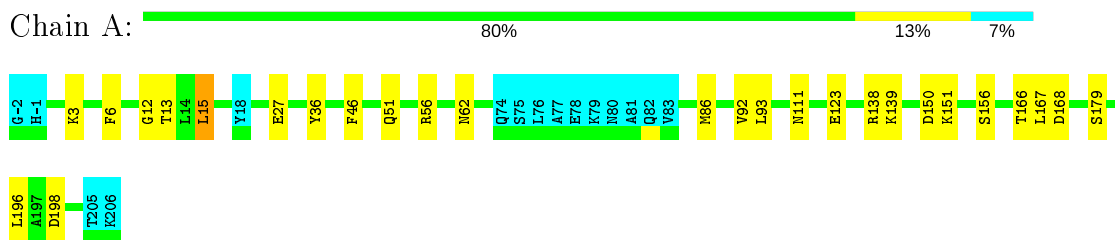
• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.16 Score per residue for model 16

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.17 Score per residue for model 17

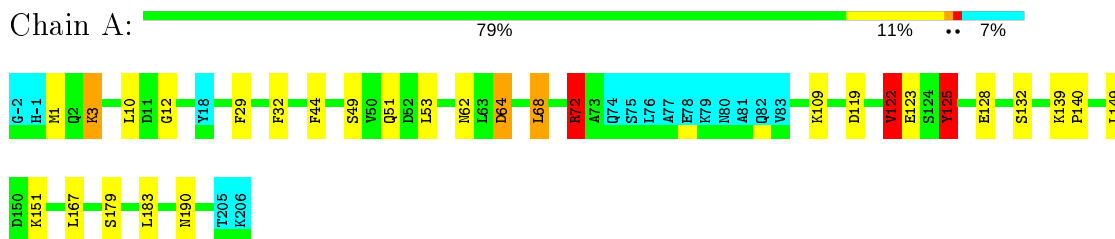• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.18 Score per residue for model 18

• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.19 Score per residue for model 19

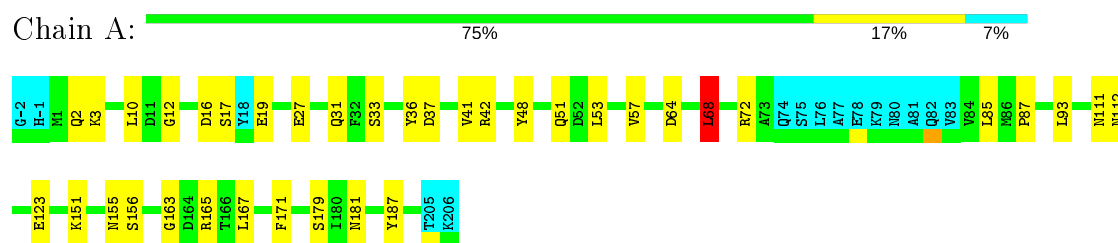• Molecule 1: Hydrolase, haloacid dehalogenase-like family

4.2.20 Score per residue for model 20

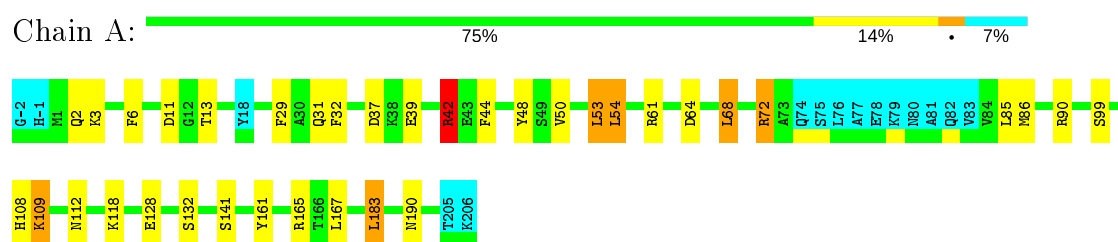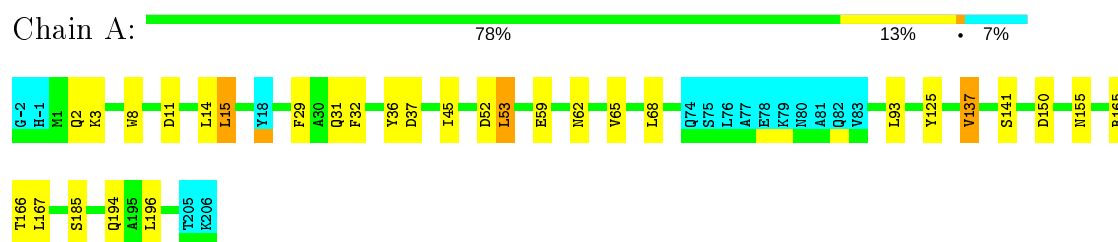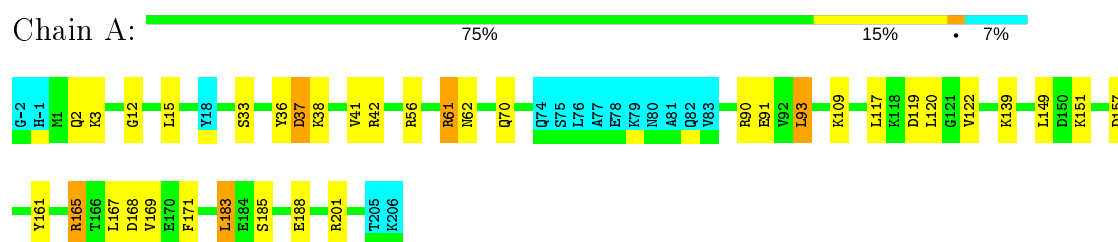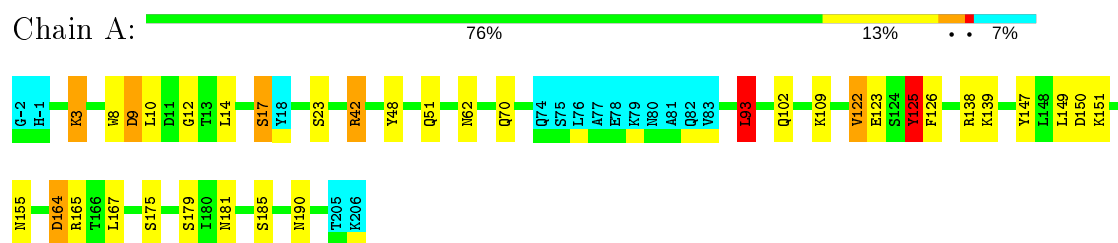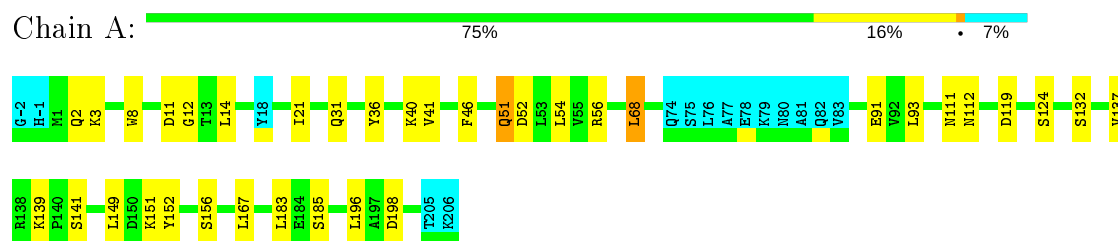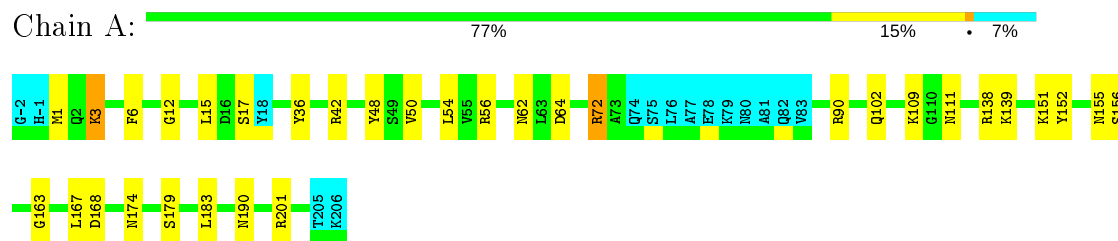• Molecule 1: Hydrolase, haloacid dehalogenase-like family

5   Refinement protocol and experimental data overview (i)

The models were refined using the following method: *torsion angle dynamics, simulated annealing*.

Of the 100 calculated structures, 20 were deposited, based on the following criterion: *target function*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CYANA	refinement	3.0
CYANA	structure solution	3.0
OPAL	refinement	
UNIO-ATNOS/CANDID	structure solution	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 6 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	2213
Number of shifts mapped to atoms	2213
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	74%

No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1   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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	1557	1513	1513	6±3
All	All	31140	30260	30260	110

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 2.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:93:LEU:HD23	1:A:125:TYR:CD2	0.83	2.08	4	1
1:A:90:ARG:HD3	1:A:93:LEU:HD22	0.72	1.59	4	1
1:A:85:LEU:HD21	1:A:122:VAL:HG11	0.69	1.63	2	1
1:A:36:TYR:CE2	1:A:41:VAL:HG21	0.67	2.25	1	1
1:A:36:TYR:CE1	1:A:41:VAL:HG21	0.66	2.25	19	4
1:A:28:THR:HG21	1:A:72:ARG:NH2	0.65	2.06	6	1
1:A:22:LEU:HD13	1:A:38:LYS:HD2	0.64	1.70	4	1
1:A:54:LEU:H	1:A:54:LEU:HD12	0.63	1.54	7	3
1:A:169:VAL:HG11	1:A:185:SER:HB3	0.60	1.72	7	1
1:A:85:LEU:CD2	1:A:122:VAL:HG11	0.60	2.27	2	1
1:A:15:LEU:HD13	1:A:15:LEU:N	0.59	2.12	11	2
1:A:122:VAL:HG23	1:A:125:TYR:CD2	0.59	2.32	3	6
1:A:32:PHE:CZ	1:A:68:LEU:HD12	0.59	2.32	1	3
1:A:89:ALA:HB3	1:A:90:ARG:NH2	0.55	2.16	4	1
1:A:15:LEU:HD13	1:A:15:LEU:H	0.53	1.63	11	2
1:A:29:PHE:CZ	1:A:53:LEU:HD11	0.53	2.39	1	9
1:A:37:ASP:O	1:A:41:VAL:HG23	0.53	2.03	14	3
1:A:44:PHE:HA	1:A:47:LYS:CG	0.52	2.34	1	1
1:A:28:THR:HG21	1:A:72:ARG:HH22	0.51	1.64	6	1
1:A:32:PHE:CE1	1:A:68:LEU:HD23	0.51	2.41	5	2
1:A:85:LEU:HD13	1:A:86:MET:N	0.51	2.21	15	1
1:A:122:VAL:HG22	1:A:126:PHE:CE2	0.51	2.41	2	4
1:A:64:ASP:O	1:A:68:LEU:HD13	0.50	2.06	1	2
1:A:44:PHE:HA	1:A:47:LYS:HG3	0.50	1.83	1	1
1:A:193:ILE:HD12	1:A:195:ALA:O	0.49	2.07	2	2
1:A:8:TRP:CZ3	1:A:162:ILE:HG21	0.49	2.42	2	2
1:A:8:TRP:CE3	1:A:14:LEU:HD13	0.49	2.42	9	4
1:A:15:LEU:N	1:A:15:LEU:CD1	0.48	2.76	16	1
1:A:53:LEU:O	1:A:57:VAL:HG23	0.48	2.08	11	2
1:A:15:LEU:CD1	1:A:15:LEU:N	0.48	2.77	11	2
1:A:64:ASP:O	1:A:68:LEU:HD22	0.48	2.07	2	2
1:A:50:VAL:O	1:A:54:LEU:HD12	0.47	2.09	5	3
1:A:54:LEU:HD22	1:A:68:LEU:HB3	0.47	1.86	2	2
1:A:36:TYR:CZ	1:A:41:VAL:HG21	0.47	2.44	1	1
1:A:15:LEU:HD22	1:A:120:LEU:CD1	0.46	2.41	17	1
1:A:90:ARG:CD	1:A:93:LEU:HD22	0.46	2.36	4	1
1:A:90:ARG:HA	1:A:93:LEU:HD12	0.46	1.86	2	1
1:A:29:PHE:CZ	1:A:57:VAL:HG21	0.44	2.48	1	1
1:A:15:LEU:N	1:A:15:LEU:HD13	0.44	2.27	16	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:164:ASP:HB2	1:A:182:PHE:CD2	0.44	2.48	9	1
1:A:44:PHE:CD1	1:A:53:LEU:HB2	0.43	2.47	13	4
1:A:8:TRP:CZ3	1:A:14:LEU:HD13	0.43	2.48	5	1
1:A:34:ILE:HG21	1:A:57:VAL:HG11	0.43	1.89	1	1
1:A:183:LEU:H	1:A:183:LEU:HD12	0.43	1.72	2	1
1:A:14:LEU:C	1:A:15:LEU:HD12	0.43	2.34	3	1
1:A:16:ASP:CG	1:A:84:VAL:HG23	0.43	2.34	6	1
1:A:89:ALA:HB3	1:A:90:ARG:CZ	0.43	2.44	4	1
1:A:183:LEU:HD13	1:A:183:LEU:N	0.43	2.29	15	1
1:A:51:GLN:HA	1:A:51:GLN:HE21	0.42	1.74	19	1
1:A:12:GLY:H	1:A:15:LEU:HD11	0.42	1.74	12	2
1:A:10:LEU:HA	1:A:15:LEU:HD21	0.42	1.90	7	1
1:A:51:GLN:HA	1:A:54:LEU:HD13	0.42	1.90	7	1
1:A:93:LEU:HB3	1:A:125:TYR:CZ	0.42	2.50	18	1
1:A:137:VAL:HG23	1:A:141:SER:CB	0.42	2.45	16	1
1:A:15:LEU:N	1:A:15:LEU:HD22	0.42	2.30	8	1
1:A:169:VAL:HG11	1:A:185:SER:CB	0.41	2.45	17	1
1:A:180:ILE:HG22	1:A:193:ILE:HD13	0.41	1.91	3	1
1:A:54:LEU:HD12	1:A:54:LEU:H	0.41	1.75	20	1
1:A:50:VAL:HG23	1:A:51:GLN:N	0.41	2.30	5	1
1:A:183:LEU:N	1:A:183:LEU:HD13	0.41	2.31	17	2
1:A:6:PHE:CD1	1:A:101:ILE:HG21	0.41	2.51	2	1
1:A:89:ALA:HB2	1:A:196:LEU:HD23	0.40	1.92	1	1
1:A:8:TRP:CH2	1:A:92:VAL:HG11	0.40	2.51	2	1
1:A:44:PHE:CE1	1:A:48:TYR:HB2	0.40	2.52	5	1
1:A:54:LEU:HB2	1:A:65:VAL:HG23	0.40	1.93	3	1

5.2 Torsion angles ⓘ

5.2.1 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 PDB entries followed by that with respect to all NMR entries. 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	193/208 (93%)	171±4 (88±2%)	20±4 (10±2%)	3±2 (1±1%)	14	59
All	All	3860/4160 (93%)	3413 (88%)	391 (10%)	56 (1%)	14	59

All 19 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	3	LYS	13
1	A	12	GLY	11
1	A	122	VAL	5
1	A	1	MET	4
1	A	139	LYS	4
1	A	138	ARG	2
1	A	121	GLY	2
1	A	88	GLY	2
1	A	163	GLY	2
1	A	2	GLN	2
1	A	17	SER	1
1	A	87	PRO	1
1	A	137	VAL	1
1	A	136	PHE	1
1	A	164	ASP	1
1	A	140	PRO	1
1	A	50	VAL	1
1	A	189	GLY	1
1	A	109	LYS	1

5.2.2 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 PDB entries followed by that with respect to all NMR entries. 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	169/181 (93%)	146±4 (86±2%)	23±4 (14±2%)	7	47
All	All	3380/3620 (93%)	2913 (86%)	467 (14%)	7	47

All 95 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	167	LEU	19
1	A	62	ASN	16
1	A	3	LYS	12
1	A	165	ARG	12
1	A	155	ASN	11

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	109	LYS	10
1	A	151	LYS	10
1	A	2	GLN	10
1	A	150	ASP	10
1	A	183	LEU	10
1	A	123	GLU	9
1	A	152	TYR	9
1	A	179	SER	9
1	A	111	ASN	9
1	A	51	GLN	9
1	A	149	LEU	9
1	A	52	ASP	8
1	A	72	ARG	8
1	A	190	ASN	8
1	A	138	ARG	7
1	A	119	ASP	7
1	A	85	LEU	7
1	A	53	LEU	7
1	A	10	LEU	7
1	A	141	SER	6
1	A	38	LYS	6
1	A	37	ASP	6
1	A	27	GLU	6
1	A	112	ASN	6
1	A	31	GLN	6
1	A	156	SER	6
1	A	198	ASP	6
1	A	42	ARG	6
1	A	132	SER	6
1	A	196	LEU	6
1	A	128	GLU	6
1	A	168	ASP	5
1	A	171	PHE	5
1	A	11	ASP	5
1	A	15	LEU	5
1	A	68	LEU	5
1	A	185	SER	5
1	A	139	LYS	5
1	A	56	ARG	5
1	A	33	SER	5
1	A	9	ASP	4
1	A	102	GLN	4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	23	SER	4
1	A	17	SER	4
1	A	93	LEU	4
1	A	90	ARG	4
1	A	48	TYR	4
1	A	13	THR	4
1	A	40	LYS	4
1	A	181	ASN	4
1	A	175	SER	4
1	A	157	ASP	4
1	A	6	PHE	4
1	A	164	ASP	3
1	A	118	LYS	3
1	A	117	LEU	3
1	A	46	PHE	3
1	A	49	SER	3
1	A	19	GLU	3
1	A	92	VAL	3
1	A	192	ARG	3
1	A	64	ASP	3
1	A	194	GLN	3
1	A	166	THR	3
1	A	124	SER	2
1	A	201	ARG	2
1	A	70	GLN	2
1	A	91	GLU	2
1	A	99	SER	2
1	A	108	HIS	2
1	A	143	GLU	1
1	A	137	VAL	1
1	A	158	ASN	1
1	A	1	MET	1
1	A	131	THR	1
1	A	188	GLU	1
1	A	126	PHE	1
1	A	16	ASP	1
1	A	86	MET	1
1	A	61	ARG	1
1	A	14	LEU	1
1	A	39	GLU	1
1	A	174	ASN	1
1	A	47	LYS	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	170	GLU	1
1	A	84	VAL	1
1	A	159	THR	1
1	A	98	GLU	1
1	A	59	GLU	1
1	A	60	ASP	1

5.2.3 RNA ⓘ

There are no RNA molecules in this entry.

5.3 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.4 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.5 Ligand geometry ⓘ

There are no ligands in this entry.

5.6 Other polymers ⓘ

There are no such molecules in this entry.

5.7 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 74% for the well-defined parts and 74% for the entire structure.

6.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_chem_shift_list_1*

6.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	2213
Number of shifts mapped to atoms	2213
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	7

6.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	205	-0.08 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	196	0.44 ± 0.11	None needed (< 0.5 ppm)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	196	0.24 ± 0.44	None needed (< 0.5 ppm)

6.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 74%, i.e. 1786 atoms were assigned a chemical shift out of a possible 2401. 31 out of 31 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	554/957 (58%)	181/382 (47%)	190/386 (49%)	183/189 (97%)
Sidechain	1071/1226 (87%)	656/711 (92%)	386/460 (84%)	29/55 (53%)

Continued on next page...

Continued from previous page...

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Aromatic	161/218 (74%)	105/116 (91%)	54/98 (55%)	2/4 (50%)
Overall	1786/2401 (74%)	942/1209 (78%)	630/944 (67%)	214/248 (86%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 74%, i.e. 1917 atoms were assigned a chemical shift out of a possible 2583. 33 out of 33 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	**^{1}H**	**^{13}C**	**^{15}N**
Backbone	597/1032 (58%)	196/412 (48%)	205/416 (49%)	196/204 (96%)
Sidechain	1149/1318 (87%)	705/765 (92%)	412/493 (84%)	32/60 (53%)
Aromatic	171/233 (73%)	111/124 (90%)	58/104 (56%)	2/5 (40%)
Overall	1917/2583 (74%)	1012/1301 (78%)	675/1013 (67%)	230/269 (86%)

6.1.4 Statistically unusual chemical shifts ⓘ

The following table lists the statistically unusual chemical shifts. These are statistical measures, and large deviations from the mean do not necessarily imply incorrect assignments. Molecules containing paramagnetic centres or hemes are expected to give rise to anomalous chemical shifts.

Mol	Chain	Res	Type	Atom	Shift, *ppm*	Expected range, *ppm*	Z-score
1	A	162	ILE	HD12	-1.47	2.13 − -0.77	-7.4
1	A	162	ILE	HD11	-1.47	2.13 − -0.77	-7.4
1	A	162	ILE	HD13	-1.47	2.13 − -0.77	-7.4
1	A	200	SER	HB2	2.12	5.18 − 2.58	-6.8
1	A	151	LYS	HG2	-0.37	2.67 − 0.07	-6.7
1	A	173	GLN	HB2	0.59	3.30 − 0.80	-5.9
1	A	93	LEU	HB2	-0.14	3.32 − -0.08	-5.2

6.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A: