Biomol NMR Assign
DOI 10.1007/s12104-017-9782-3

ARTICLE

^{1}H, ^{13}C and ^{15}N backbone resonance assignments of the β-lactamase BlaP from *Bacillus licheniformis 749/C* and two mutational variants

David Thorn[1,4] · Jennifer Kay[1] · Noureddine Rhazi[1,5] · Mireille Dumoulin[1] · Alessandra Corazza[2] · Christian Damblon[3]

Received: 10 July 2017 / Accepted: 18 September 2017

Abstract Class A β-lactamases have been widely used as versatile scaffolds to create hybrid (or chimeric) proteins for a series of applications ranging from basic research to medicine. We have, in particular, used the β-lactamase BlaP from *Bacillus licheniformis 749/C* (BlaP) as a protein scaffold to create model polyglutamine (polyQ) proteins in order to better understand the mechanism(s) by which an expanded polyQ sequence triggers the formation of amyloid fibrils. The model chimeras were designed by inserting a polyQ sequence of various lengths at two different locations within BlaP (i.e. position 197 or position 216) allowing a detailed comparison of the effects of subtle differences in the environment of the polyQ sequence on its ability to trigger protein aggregation. In order to investigate the effects of the polyQ insertion at both positions on the structure, stability and dynamics of BlaP, a series of NMR experiments including H/D exchange are foreseen. Accordingly, as necessitated by these studies, here we report the NMR assignment of the wild-type BlaP (BlaP-WT) and of the two reference proteins, BlaP197Q0 and BlaP216Q0, wherein a Pro-Gly dipeptide has been introduced at position 197 and 216, respectively; this dipeptide originates from the addition of the Sma1 restriction site at the genetic level to allow further polyQ sequence insertion.

Keywords BlaP hybrid proteins · Polyglutamine model proteins · Protein aggregation · Polyglutamine diseases · Resonance assignment

Biological context

Polyglutamine (polyQ) diseases, including the well-known Huntington's disease and several spinocerebellar ataxias, are progressive and inherited neurodegenerative disorders associated with neural proteins that contain an 'expanded' polyQ sequence encoded by CAG trinucleotides repeats (Hands and Wyttenbach 2010). CAG repeats are present in the genes

David Thorn and Jennifer Kay have equally contributed to this work.

✉ Alessandra Corazza
alessandra.corazza@uniud.it

✉ Christian Damblon
c.damblon@ulg.ac.be

David Thorn
david.thorn@anu.edu.au

Jennifer Kay
j.kay@ulg.ac.be

Mireille Dumoulin
mdumoulin@ulg.ac.be

[1] Laboratory of Enzymology and Protein Folding, Center for Protein Engineering, InBios, University of Liège, Allée du Six Août 13, Sart-Tilman, 4000 Liège, Belgium

[2] Department of Medicine, University of Udine, Piazzale Kolbe, 4, 33100 Udine, Italy

[3] Laboratory of Biological Structural Chemistry, Department of Chemistry, University of Liège, Allée du Six Août 13, Sart-Tilman, 4000 Liège, Belgium

[4] Present Address: Research School of Chemistry, The Australian National University, Acton, ACT 2601, Australia

[5] Present Address: Molecular Biomimetic and Protein Engineering Laboratory, GIGA-Research, University of Liège, Quartier Hôpital, avenue de l'Hopital 1, Sart-Tilman, 4000 Liège, Belgium

Published online: 13 October 2017

of even healthy people; however, the former is inherently unstable and tends to expand in successive generations by mutations leading to an increasingly long polyQ sequence. When the length of this polyQ sequence exceeds a threshold characteristic to each disease (generally comprised between 35 and 45 glutamines), it becomes pathogenic by facilitating the aggregation and conversion of the polyQ protein into amyloid fibrils which are deposited in neurons as nuclear inclusions (Orr and Zoghbi 2007; Almeida et al. 2013). Proteins involved in polyQ diseases have very little in common at the level of sequence, structure or function, except for the presence of the expanded polyQ sequence which is the critical determinant for their amyloid fibril-forming propensity (Almeida et al. 2013). It has been shown that non-polyQ regions (i.e. folded domains and/or sequences flanking the polyQ tract) play an important modulating role, either facilitative or preventive, in the aggregation process (Robertson and Bottomley 2010). The underlying principles of how non-polyQ regions influence polyQ sequences and their ability to induce protein aggregation are not yet fully understood. Disease-associated polyQ proteins cannot be easily handled in vitro due to their large size and/or poor solubility; therefore an alternative strategy to better understand these principles involves the development and characterization of model polyQ proteins whereby a polyQ sequence is inserted into a well-characterized, soluble and easy to handle host protein.

We engineered polyQ model proteins using the β-lactamase from *Bacillus licheniformis 749/C* referred to as BlaP. BlaP has been chosen as a protein scaffold because: (i) it is a relatively small and stable protein (~30 kDa), readily overexpressed in *E.coli* (Huynen et al. 2013); (ii) its three-dimensional structure is well-characterized: the protein is composed of two structural domains (an α domain and an α/β domain) at the interface of which the catalytic site is located; and most importantly, (iii) it tolerates peptide insertions at several positions (Vandevenne et al. 2007; Scarafone et al. 2012; Huynen et al. 2013).

BlaP polyQ-chimeras were created by inserting a polyQ sequence of different lengths (from 23Q to 79Q) at two locations of BlaP, i.e. at position 197 (168) and 216 (187), referred to as 197 and 216 chimeras, respectively. Note that these position numbers correspond to those of standard numbering scheme defined for class A β-lactamases, ABL (Ambler et al. 1991); the underlined number appearing here in parentheses corresponds to the number in BlaP sequence. The correspondence between the two numbering systems is given in Table 1. The ABL numbering has been previously used to describe the characterization of BlaP-based hybrid proteins (Vandevenne et al. 2007; Scarafone et al. 2012; Huynen et al. 2013, 2015) and is therefore used here in the introduction and methods sections; in the remainder of the paper dealing with NMR, only the sequence number (underlined number) will be used. Note also that the BlaP protein used to create the polyQ chimeras is the secreted form following complete N-terminal proteolytic processing and therefore starting at position 27 (Vandevenne et al. 2007) (Table 1).

These two insertion sites were selected based on the following: (i) both sites correspond to permissive insertion sites, e.g. insertion of peptide of various lengths at these positions results in chimeric BlaP proteins that are soluble and functional (Vandevenne et al. 2007); (ii) both sites are located in solvent-exposed loops of similar size (4 residues for the 197 loop versus 5 residues for the 216 loop) connecting two α-helices, therefore they have a comparable overall protein context; (iii) while position 197 is situated in a loop within the α-domain, the loop containing position 216 links the α and α/β domains. This distinction allows us to compare the effects of an intradomain (i.e. at position 197) versus those of an interdomain (i.e. at position 216) insertion of a polyQ sequence on the properties of the host protein.

To insert poly(CAG) sequences within the β-lactamase gene, recognition sites for the Sma1 restriction enzyme have been introduced at the location of the insertion (i.e. between sequences coding for the amino acids 197–198 or 216–217) (Scarafone et al. 2012; Thorn et al. *in preparation*) (Table 1). The insertion of this restriction site at the gene level leads to translation of the dipeptide proline-glycine at position 197 and 216 respectively; these proteins are referred to as BlaP197Q0 and BlaP216Q0. By comparison with these two BlaP variants with a Pro-Gly dipeptide inserted at these two positions, we aim to distinguish between the impact of the polyQ repeat and insertions in general.

The aggregation properties of the BlaP polyQ-chimeras are characteristic of proteins associated with polyQ diseases (Scarafone et al. 2012; Thorn et al. *in preparation*); namely: (i) there is a Q-threshold length above which the polyQ sequence triggers the aggregation of proteins into amyloid fibrils, and (ii) above this threshold, the longer the polyQ sequence, the faster the aggregation. Moreover, the aggregation propensity of BlaP-polyQ chimeras depends on the location of the polyQ sequence within BlaP. Interestingly, the 216 chimeras display increased aggregation propensity: indeed (i) the minimum polyQ length leading to aggregation is lower compared to 197 chimeras, and (ii) above threshold length, the aggregation rate is significantly higher than that observed by 197 chimeras with equivalent polyQ lengths. This observation suggests a significant modulating role of the non-polyQ regions (i.e. the BlaP moiety) on the mechanism of aggregation induced by polyQ sequences. Altogether, these results indicate that BlaP-polyQ chimeras are relevant protein models to further investigate (i) how expanded polyQ sequences trigger the aggregation of proteins into amyloid fibrils and (ii) how non-polyQ regions modulate the propensity of polyQ sequences to induce aggregation. The latter may depend on the (local) stability and dynamics of the BlaP moiety within each polyQ chimera; experiments

such as H/D exchange monitored by NMR are therefore of particular interest. Such work requires the assignment of the [^{1}H-^{15}N] 2D HSQC spectrum of the BlaP chimeras. Here, we report the NMR assignment of the wild-type BlaP (BlaP-WT) and of the two reference proteins, BlaP197Q0 and BlaP216Q0.

Methods and experiments

Molecular biology

The gene coding for BlaP-WT, BlaP197Q0 and BlaP216Q0 were amplified from the pNYBlaP-WT, pNYBlaP197Q0 and pNYBlaP216Q0 vectors, respectively (Scarafone et al. 2012; Huynen et al. 2015), with the following primers: 5′-ACA ATTCCCCTCTAGAAATAATTTTGTTTAACTTTAAGA AGGAGATATACCATGACGGAGATGAAAGATGAT TTTGCAAAAC-3′ and 5′-GACGGCCAGTGAATTCTC AGTGGTGGTGGTGGTGGGGCCCTTTCCCGTTCA TGTTTAAGGC-3′. The resulting DNA was then cloned into the pET28b vector containing a kanamycin resistance gene and for which protein expression is inducible by isopropyl β-D-1-thiogalactopyranoside (IPTG) using T7 promoter. In this vector, the BlaP gene is followed by a C-terminal poly(His)$_5$ tag to facilitate the purification of the protein. A SmaI restriction site has also been introduced into the BlaP variant gene between the trinucleotides coding for residues 197–198 (168–169) or 216–217 (187–188); this restriction site is translated into a Pro-Gly dipeptide at position 197–198 (168–169) or 216–217 (187–188) of the protein respectively. The sequence of the three genes cloned into the pET28b vector was checked by the Sanger method at the GIGA GenoTranscriptomics technology platform (Liège, Belgium). These vectors were then used to transform the BL21(DE3) expression strain of *E. coli* (Novagen, WI, USA).

Expression and purification of ^{15}N and ^{13}C double-labeled BlaP and its variants

BlaP-WT and its PG-inserted variants were expressed in *E. coli* BL21 (DE3) cells and uniformly labeled with ^{15}N and ^{13}C. Firstly, cells were grown in 2 L of LB medium containing 50 µg/mL kanamycin at 37 °C under orbital agitation at 250 rpm until the absorbency at 600 nm reached ≈0.7. Cultures were then centrifuged at 2800 g for 5 min (Avanti J-E Centrifuge, Beckman Coulter). Supernatants were discarded and the pellets were resuspended in 500 mL of M9 minimal medium containing 50 µg/mL kanamycin, 2 g ^{13}C-D-glucose and 0.5 g ^{15}N-NH_4Cl as carbon and nitrogen sources, respectively (Cambridge Isotope Laboratory, Andover, USA). After 1 h of incubation at 37 °C under orbital agitation at 250 rpm, protein expression was induced with 1.2 mM IPTG (final concentration). Cell growth was then allowed for a further 3 h of incubation. Cells were harvested by centrifugation (1400 g for 15 min, Avanti J-E Centrifuge, Beckman Coulter) and then resuspended in 100 mL phosphate buffered saline (PBS, 50 mM sodium phosphate, 150 mM NaCl, pH 7.5) containing 2 µL benzonase (Novagen), 2 mM $MgCl_2$ and 2 tablets of EDTA-free, protease-inhibitor cocktail (cOmplete™, Roche, Penzberg, Germany). Cells were homogenized using an Ultra-turax homogenizer (IKA T18 basic, VWR, Radnor, Pennsylvania) and lysed by three passages through a French press (Emulsiflex-C3, Avestin, Ottawa) at 1000–1500 bar. The cytoplasmic extract obtained was then centrifuged at 2600 g for 30 min (Avanti J-E Centrifuge, Beckman Coulter) and the supernatant successively filtered through a 0.7 and 0.45 µm cut-off pore membrane. BlaP-WT and its variants were purified by nickel affinity chromatography as previously described (Scarafone et al. 2012). Fractions of elution containing more than 95% of proteins of interest, as assessed by SDS–PAGE and densitometry analysis, were pooled and dialyzed twice against 15 L of 20 mM sodium phosphate, 50 mM NaCl, pH 6.5. The protein concentration was determined by measuring the absorbance at 280 nm using a molar extinction coefficient value of 33,000 M^{-1} cm^{-1} (Scarafone et al. 2012) and proteins were stored at −20 °C.

NMR spectroscopy

NMR experiments were acquired at 308 K using ^{15}N-^{13}C double-labeled protein concentrated to 500–800 µM in 20 mM sodium phosphate, 50 mM NaCl, pH 6.5, with added 5% D_2O and 0.2% of DSS. ^{1}H, ^{13}C and ^{15}N backbone resonance assignments for BlaP197Q0 and BlaP216Q0 were determined from [^{1}H-^{15}N] 2D BEST-TROSY HSQC spectra and standard 3D TROSY triple resonance NMR experiments (including HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, HNCO, and HN(CA)CO) acquired on a 900 MHz Bruker spectrometer equipped with a 5-mm cryogenic triple-resonance probe (Biomolecular Magnetic Resonance Center, Goethe University of Frankfurt). Backbone resonance assignments for BlaP-WT were determined from [^{1}H-^{15}N] 2D HSQC spectrum and 3D NMR experiments (i.e. HNCA and HN(CO)CA) acquired on a 500 MHz Bruker spectrometer equipped with a 5-mm cryogenic triple-resonance probe (Centre de REsonance MAgnétique Nucléaire, University of Liège). Data of BlaP-WT and BlaP197Q0 were processed using Topspin (version 2.1, Bruker BioSpin GmbH, Germany) and analyzed using CCPNMR (Vranken et al. 2005). Data for BlaP216Q0 were processed using NMRPipe (Delaglio et al. 1995) and analyzed using Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). All NMR spectra were referenced by assigning to 0 ppm the hydrogen frequency of DSS and using the indirect referencing method described by Wishart et al. (1995).

Table 1 Correspondence between the standard numbering scheme defined for class A β-lactamases (Ambler et al. 1991) and the number in BlaP sequence as well as that of BlaP variants. First column, list of residues; ABL, number according to the standard numbering scheme defined for class A β-lactamase; WT, Blap-WT; 197, Blap197Q0; 216, Blap216Q0. The position of the two insertion sites is framed and residue numbers are underlined when they differ from that of Blap-WT

	ABL	WT	197	216		ABL	WT	197	216		ABL	WT	197	216
T	**27**	1	1	1	Q	**83**	56	56	56	K	**140**	111	111	111
E	**28**	2	2	2	K	**86**	57	57	57	Q	**141**	112	112	112
M	**29**	3	3	3	S	**87**	58	58	58	I	**142**	113	113	113
K	**30**	4	4	4	I	**88**	59	59	59	G	**143**	114	114	114
D	**31**	5	5	5	E	**89**	60	60	60	G	**144**	115	115	115
D	**32**	6	6	6	D	**90**	61	61	61	P	**145**	116	116	116
F	**33**	7	7	7	L	**91**	62	62	62	E	**146**	117	117	117
A	**34**	8	8	8	N	**92**	63	63	63	S	**147**	118	118	118
K	**35**	9	9	9	Q	**93**	64	64	64	L	**148**	119	119	119
L	**36**	10	10	10	R	**94**	65	65	65	K	**149**	120	120	120
E	**37**	11	11	11	I	**95**	66	66	66	K	**150**	121	121	121
E	**38**	12	12	12	T	**96**	67	67	67	E	**151**	122	122	122
Q	**39**	13	13	13	Y	**97**	68	68	68	L	**152**	123	123	123
F	**40**	14	14	14	T	**98**	69	69	69	R	**153**	124	124	124
D	**41**	15	15	15	R	**99**	70	70	70	K	**154**	125	125	125
A	**42**	16	16	16	D	**100**	71	71	71	I	**155**	126	126	126
K	**43**	17	17	17	D	**101**	72	72	72	G	**156**	127	127	127
L	**44**	18	18	18	L	**102**	73	73	73	D	**157**	128	128	128
G	**45**	19	19	19	V	**103**	74	74	74	E	**158**	129	129	129
I	**46**	20	20	20	N	**104**	75	75	75	V	**159**	130	130	130
F	**47**	21	21	21	Y	**105**	76	76	76	T	**160**	131	131	131
A	**48**	22	22	22	N	**106**	77	77	77	N	**161**	132	132	132
L	**49**	23	23	23	P	**107**	78	78	78	P	**162**	133	133	133
D	**50**	24	24	24	I	**108**	79	79	79	E	**163**	134	134	134
T	**51**	25	25	25	T	**109**	80	80	80	R	**164**	135	135	135
G	**52**	26	26	26	E	**110**	81	81	81	F	**165**	136	136	136
T	**53**	27	27	27	K	**111**	82	82	82	E	**166**	137	137	137
N	**54**	28	28	28	H	**112**	83	83	83	P	**167**	138	138	138
R	**55**	29	29	29	V	**113**	84	84	84	E	**168**	139	139	139
T	**56**	30	30	30	D	**114**	85	85	85	L	**169**	140	140	140
V	**57**	31	31	31	T	**115**	86	86	86	N	**170**	141	141	141
A	**59**	32	32	32	G	**116**	87	87	87	E	**171**	142	142	142
Y	**60**	33	33	33	M	**117**	88	88	88	V	**172**	143	143	143
R	**61**	34	34	34	T	**118**	89	89	89	N	**173**	144	144	144
P	**62**	35	35	35	L	**119**	90	90	90	P	**174**	145	145	145
D	**63**	36	36	36	K	**120**	91	91	91	G	**175**	146	146	146
E	**64**	37	37	37	E	**121**	92	92	92	E	**176**	147	147	147
R	**65**	38	38	38	L	**122**	93	93	93	T	**177**	148	148	148
F	**66**	39	39	39	A	**123**	94	94	94	Q	**178**	149	149	149
A	**67**	40	40	40	D	**124**	95	95	95	D	**179**	150	150	150
F	**68**	41	41	41	A	**125**	96	96	96	T	**180**	151	151	151
A	**69**	42	42	42	S	**126**	97	97	97	S	**181**	152	152	152
S	**70**	43	43	43	L	**127**	98	98	98	T	**182**	153	153	153
T	**71**	44	44	44	R	**128**	99	99	99	A	**183**	154	154	154
I	**72**	45	45	45	Y	**129**	100	100	100	R	**184**	155	155	155
K	**73**	46	46	46	S	**130**	101	101	101	A	**185**	156	156	156
A	**74**	47	47	47	D	**131**	102	102	102	L	**186**	157	157	157
L	**75**	48	48	48	N	**132**	103	103	103	V	**187**	158	158	158
T	**76**	49	49	49	A	**133**	104	104	104	T	**188**	159	159	159
V	**77**	50	50	50	A	**134**	105	105	105	S	**189**	160	160	160
G	**78**	51	51	51	Q	**135**	106	106	106	L	**190**	161	161	161
V	**79**	52	52	52	N	**136**	107	107	107	R	**191**	162	162	162
L	**80**	53	53	53	L	**137**	108	108	108	A	**192**	163	163	163
L	**81**	54	54	54	I	**138**	109	109	109	F	**193**	164	164	164
Q	**82**	55	55	55	L	**139**	110	110	110	A	**194**	165	165	165

Table 1 (continued)

	ABL	WT	197	216
L	195	166	166	166
E	196	167	167	167
D	197	168	168	168
P			169	
G			170	
K	198	169	171	169
L	199	170	172	170
P	200	171	173	171
S	201	172	174	172
E	202	173	175	173
K	203	174	176	174
R	204	175	177	175
E	205	176	178	176
L	206	177	179	177
L	207	178	180	178
I	208	179	181	179
D	209	180	182	180
W	210	181	183	181
M	211	182	184	182
K	212	183	185	183
R	213	184	186	184
N	214	185	187	185
T	215	186	188	186
T	216	187	189	187
P				188
G				189
G	217	188	190	190
D	218	189	191	191
A	219	190	192	192
L	220	191	193	193
I	221	192	194	194
R	222	193	195	195
A	223	194	196	196
G	224	195	197	197
V	225	196	198	198
P	226	197	199	199
D	227	198	200	200
G	228	199	201	201
W	229	200	202	202
E	230	201	203	203
V	231	202	204	204
A	232	203	205	205
D	233	204	206	206
K	234	205	207	207
T	235	206	208	208
G	236	207	209	209
A	237	208	210	210
A	238	209	211	211
S	240	210	212	212
Y	241	211	213	213
G	242	212	214	214
T	243	213	215	215
R	244	214	216	216
N	245	215	217	217
D	246	216	218	218
I	247	217	219	219
A	248	218	220	220
I	249	219	221	221
I	250	220	222	222
W	251	221	223	223
P	252	222	224	224
P	254	223	225	225
K	255	224	226	226
G	256	225	227	227
D	257	226	228	228
P	258	227	229	229
V	259	228	230	230
V	260	229	231	231
L	261	230	232	232
A	262	231	233	233
V	263	232	234	234
L	264	233	235	235
S	265	234	236	236
S	266	235	237	237
R	267	236	238	238
D	268	237	239	239
K	269	238	240	240
K	270	239	241	241
D	271	240	242	242
A	272	241	243	243
K	273	242	244	244
Y	274	243	245	245
D	275	244	246	246
D	276	245	247	247
K	277	246	248	248
L	278	247	249	249
I	279	248	250	250
A	280	249	251	251
E	281	250	252	252
A	282	251	253	253
T	283	252	254	254
K	284	253	255	255
V	285	254	256	256
V	286	255	257	257
M	287	256	258	258
K	288	257	259	259
A	289	258	260	260
L	290	259	261	261
N	291	260	262	262
M	292	261	263	263
N	293	262	264	264
G	294	263	265	265
K	295	264	266	266
G	296	265	267	267

Extent of assignment and data deposition

BlaP-WT comprises 265 residues (excluding the poly(His)$_5$ tag and the preceding proline residue introduced with the tag), while the BlaP197Q0 and BlaP216Q0 variants are composed of 267 residues due to the addition of the Pro-Gly dipeptide (Table 1). Despite the size of the proteins (30 kDa), the HSQC spectra of the three proteins are well-dispersed owing to their folded structure and α/β content.

BlaP-WT

Figure 1 shows the [^{1}H-^{15}N] 2D HSQC spectrum of BlaP-WT. A total of 98.8% (excluding the 11 prolines) of the expected backbone ^{1}H-^{15}N correlations were assigned; corresponding to 251 out of 254 residues. Moreover, 94.4% of the C^{α} were assigned. Missing assignments (i.e. residues for which no corresponding resonance was observed) include the two first residues belonging to the N-terminus (T1 & E2) and A208. The chemical shift assignments for BlaP-WT have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under the accession number 25821.

BlaP197Q0

Figure 2 shows the [^{1}H-^{15}N] 2D HSQC spectrum of BlaP197Q0. A total of 95.3% of the 'assignable' backbone ^{1}H-^{15}N correlations (i.e. excluding the 12 prolines) were assigned; corresponding to 243 out of 255 residues. Moreover 95.1% C^{α}, 95.2% C^{β} and 95.1% C' were also assigned. Missing assignments include the two residues belonging to the N-terminus (T1 & E2), A163, A165, all residues juxtaposing the insertion site D168 from E167 to L172 inclusive, A210 and D242. The chemical shift assignments for BlaP197Q0 have been deposited in the BioMagResBank

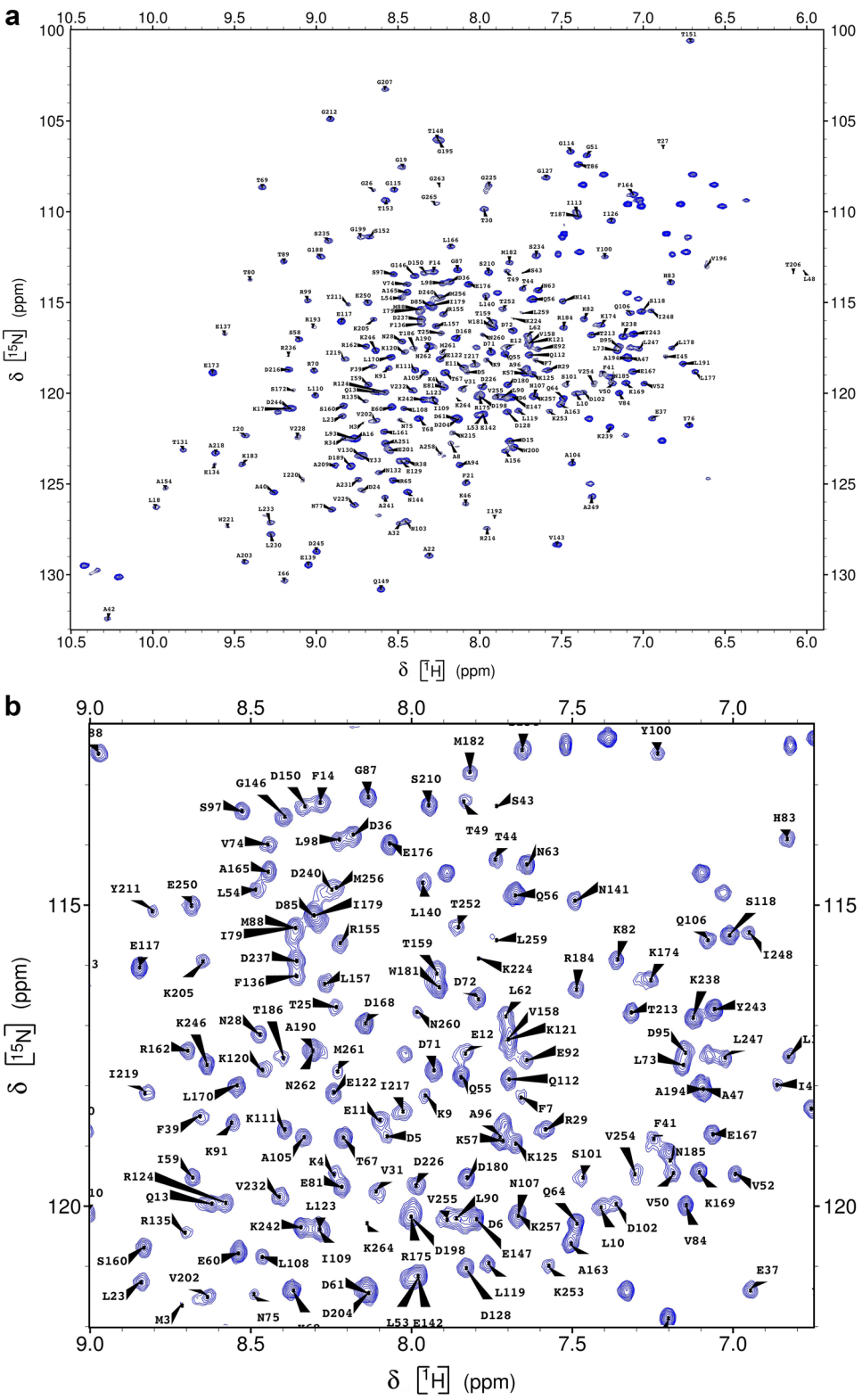

Fig. 1 **a** 500 MHz [^{1}H-^{15}N] 2D HSQC spectrum of ^{15}N-^{13}C double-labeled BlaP-WT in 20 mM phosphate buffer, 50 mM NaCl, 5% D20 and 0.2% DSS at pH 6.5 and 308K. **b** Magnified view of central, clustered region of spectrum

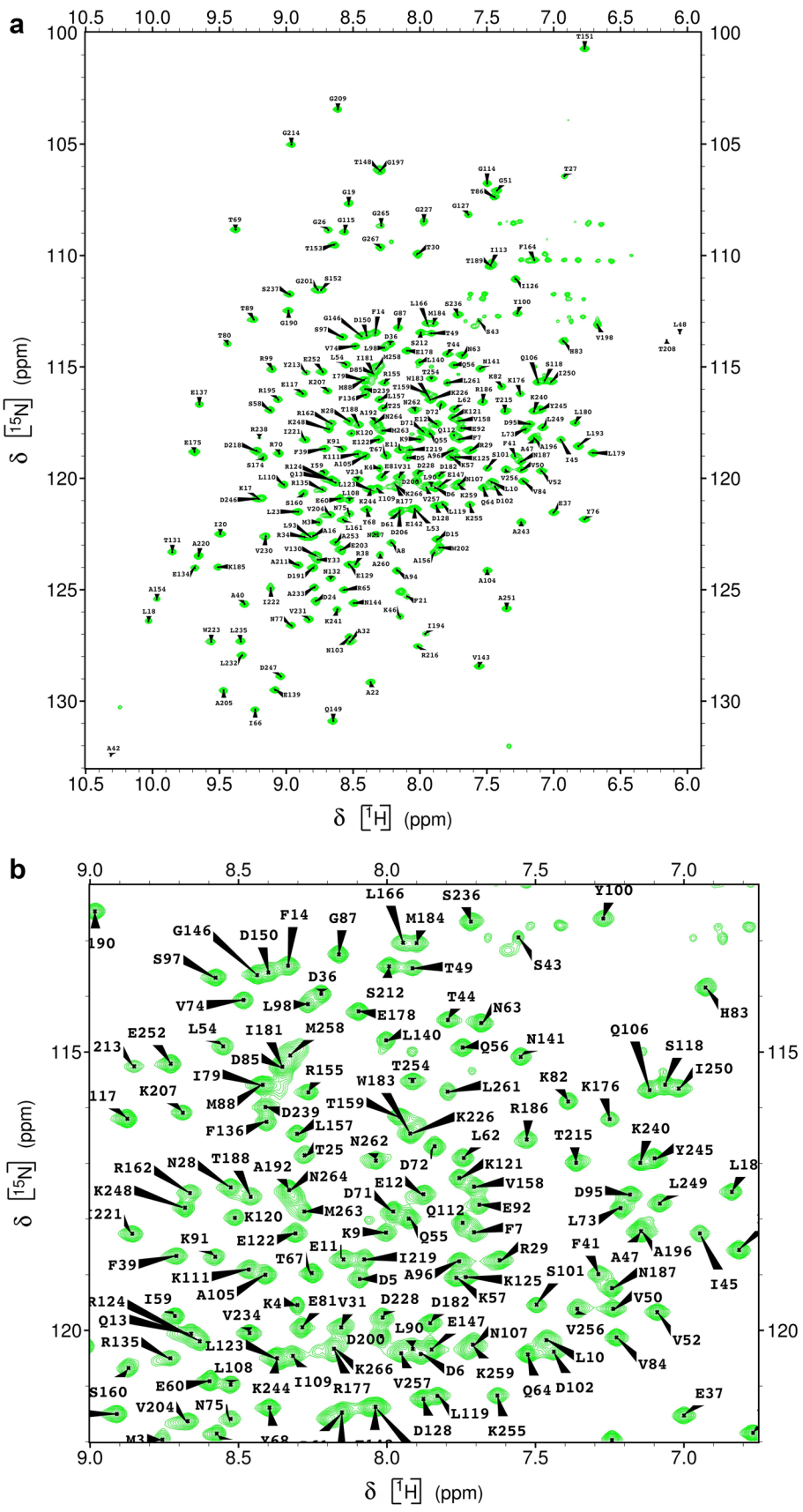

Fig. 2 **a** 900 MHz [^{1}H-^{15}N] 2D BEST-TROSY HSQC spectrum of ^{15}N-^{13}C double-labeled BlaP197Q0 in 20 mM phosphate buffer, 50 mM NaCl, 5% D20 and 0.2% DSS at pH 6.5 and 308K. **b** Magnified view of central, clustered region of spectrum

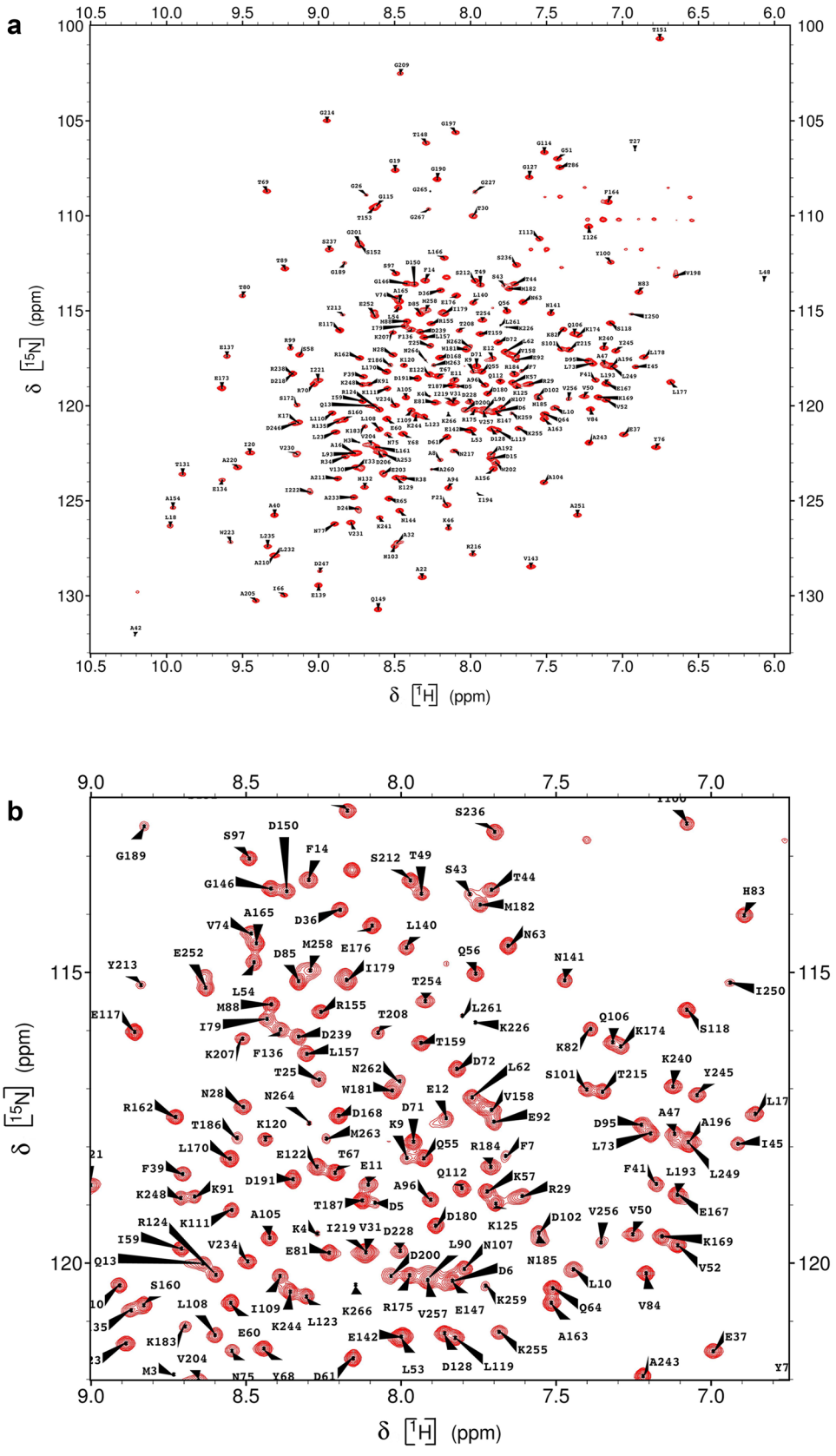

Fig. 3 **a** 900 MHz [^{1}H-^{15}N] 2D BEST-TROSY HSQC spectrum of ^{15}N-^{13}C double-labeled BlaP216Q0 in 20 mM phosphate buffer, 50 mM NaCl, 5% D20 and 0.2% DSS at pH 6.5 and 308K. **b** Magnified view of central, clustered region of spectrum

(http://www.bmrb.wisc.edu) under the accession number 19976.

BlaP216Q0

Figure 3 shows the [^{1}H-^{15}N] 2D HSQC spectrum of BlaP216Q0. A total of 98.4% of the 'assignable' backbone ^{1}H-^{15}N correlations were assigned; corresponding to 251 out of 255 residues. Moreover 98.5% C^{α}, 98.4% C^{β} and 92.9% $C^{'}$ were also assigned. Missing assignments include the first residues belonging to the N-terminus (T1 & E2), R195 and D242. In contrast with the BlaP197Q0 assignment, it was possible to assign residues juxtaposing the insertion site T187. The chemical shift assignments for BlaP216Q0 have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under the accession number 25003.

For all three proteins, the backbone amides belonging to the two N-terminal residues (i.e. T1 & E2) were not observed due to rapid hydrogen exchange with the solvent at neutral pH (Bundi and Wüthrich 1977). Likewise, complete backbone assignment in the flexible and exposed intradomain loop encompassing the 168 insertion site (i.e. between α-helices 8 and 9) was challenging, particularly following dipeptide (Pro-Gly) insertion at this site, as suggested by the high number of missing assignments in BlaP197Q0 within this region, i.e. residues E167 to L172. Interestingly, it was possible to assign residues on the interdomain loop encompassing the 187 insertion site even after dipeptide insertion. Signals for residues E167-L172 are likely too broadened to be observed due to an unfavourable intermediate exchange regime. On the contrary, residues juxtaposing the 187 insertion site in BlaP216Q0, e.g. residues T187 to G190, are readily observed in the HSQC spectrum of the protein, thereby excluding an intermediate conformational exchange regime.

Acknowledgements The authors thank Stéphane Preumont for the production and purification of the labeled proteins, and Fabrice Bouillenne and Anne-Marie Matton for their assistance with the purification of proteins. They would also like to thank Frank Löhr for conducting the NMR experiments at the Center for Biomolecular Magnetic Resonance at the Goethe University of Frankfurt, funded by the European Union (Bio-NMR, Project No. 261863). This work was supported by Grants from Fonds de la Recherche Fondamentale et Collective (CJ and 2.4581.12F to MD), Fonds de la Recherche Scientifique (FRS-FNRS, 1.C039.09 and MIS-F.4505.11 to MD), Fonds Spéciaux from the University of Liège (11/108 to MD) and the Belgian program of Interuniversity Attraction Poles administered by the Federal Office for Scientific Technical and Cultural Affairs (P7/4444 and 7/05). JK is a Recipient of FRIA fellowship of FRS-FNRS, DT was a Research Fellow from FRS-FNRS and MD is a Research Associate of FRS-FNRS. The funding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Almeida B, Fernandes S, Abreu IA, Macedo-Ribeiro S (2013) Trinucleotide repeats: a structural perspective. Front Neurol. doi:10.3389/fneur.2013.00076

Ambler RP, Coulson AF, Frère JM et al (1991) A standard numbering scheme for the class A beta-lactamases. Biochem J 276(Pt 1):269–270. doi:10.1042/bj2760269

Bundi A, Wüthrich K (1977) 1H NMR titration shifts of amide proton in polypeptide chains resonances. FEBS Lett 77(1):11–14. doi:10.1016/0014-5793(77)80182-8

Delaglio F, Grzesiek S, Vuister GW et al (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293. doi:10.1007/BF00197809

Hands SL, Wyttenbach A (2010) Neurotoxic protein oligomerisation associated with polyglutamine diseases. Acta Neuropathol 120:419–437. doi:10.1007/s00401-010-0703-0

Huynen C, Filée P, Matagne A et al (2013) Class A β-Lactamases as versatile scaffolds to create hybrid enzymes: applications from basic research to medicine. Biomed Res Int. doi:10.1155/2013/827621

Huynen C, Willet N, Buell AK et al (2015) Influence of the protein context on the polyglutamine length-dependent elongation of amyloid fibrils. Biochim Biophys Acta - Proteins Proteom 1854(3):239–248. doi:10.1016/j.bbapap.2014.12.002

Robertson AL, Bottomley SP (2010) Towards the treatment of polyglutamine diseases: the modulatory role of protein context. Curr Med Chem 17:3058–3068. doi:10.2174/092986710791959800

Scarafone N, Pain C, Fratamico A et al (2012) Amyloid-like fibril formation by polyq proteins: a critical balance between the polyq length and the constraints imposed by the host protein. PLoS ONE. doi:10.1371/journal.pone.0031253

Vandevenne M, Filee P, Scarafone N et al (2007) The Bacillus licheniformis BlaP beta-lactamase as a model protein scaffold to study the insertion of protein fragments. Protein Sci 16:2260–2271. doi:10.1110/ps.072912407

Vranken WF, Boucher W, Stevens TJ et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins Struct Funct Genet 59:687–696. doi:10.1002/prot.20449

Wishart DS, Bigam CG, Yao J et al (1995) 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR 6:135–140. doi:10.1007/BF00211777

Orr HT, Zoghbi HY (2007) Trinucleotide repeat disorders. Annu Rev Neurosci 30:575–621. doi:10.1146/annurev.neuro.29.051605.113042