Carbohydrate Research 472 (2019) 72–75

Contents lists available at ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

N-acetyltransferases from three different organisms displaying distinct selectivity toward hexosamines and N-terminal amine of peptides

Peiru Zhang[a,1], Pei Liu[a,1], Yangyang Xu[a], Yulu Liang[b], Peng George Wang[a], Jiansong Cheng[a,*]

[a] College of Pharmacy and Tianjin Key Laboratory of Molecular Drug Research, Nankai University, Haihe Education Park, 38 Tongyan Road, Tianjin, 300353, PR China
[b] College of Chemistry, Nankai University, Tianjin, 300071, PR China

ARTICLE INFO

Keywords:
N-acetyltransferase
Acetylation
Glucosamine
Peptide
Specificity

ABSTRACT

N-acetyltransferases are a family of enzymes that catalyze the transfer of the acetyl moiety (−$COCH_3$) from acetyl coenzyme A (Acetyl-CoA) to a primary amine of acceptor substrates from small molecules such as aminoglycoside to macromolecules of various proteins. In this study, the substrate selectivity of three N-acetyltransferases falling into different phylogenetic groups was probed against a series of hexosamines and synthetic peptides. GlmA from *Clostridium acetobutylicum* and RmNag from *Rhizomucor miehei*, which have been defined as glucosamine N-acetyltransferases, were herein demonstrated to be also capable of acetylating the free amino group on the very first glycine residue of peptide in spite of varied catalytic efficiency. The human recombinant N-acetyltransferase of Naa10p, however, prefers primary amine groups in the peptides as opposed to glucosamine. The varied preference of GlmA, RmNag and Naa10p probably arose from the divergent evolution of these N-acetyltransferases. The expanded knowledge of acceptor specificity would as well facilitate the application of these N-acetyltransferases in the acetylation of hexosamines or peptides.

1. Introduction

N-terminal acetylation is one of the most common protein post-translational modifications in eukaryotes, and about 85% of different varieties of eukaryotic proteins are acetylated [1,2]. In general, protein N-terminal acetylation plays a role in various cellular functions including stimulation of DNA binding, protein-protein interactions and protein stability [3,4]. Protein N-terminal acetylation also affects varieties of normal and pathological cell biology [5,6].

N-terminal acetyltransferases (NATs) catalyze the transfer of the acetyl moiety (−$COCH_3$) from acetyl coenzyme A (Acetyl-CoA) to a primary amine of acceptor substrates from small molecules such as aminoglycoside to macromolecules including various proteins [7,8]. Generally, NATs belong to the GCN-5 related N-acetyltransferases family (GNAT), which is a significant protein family and containing more than 100,000 members in all life kingdoms [7,8]. There is very limited sequence conservation among members of the GNAT superfamily, in part, reflecting their capacity to bind and modify a diverse array of substrates. In eukaryotes, a group of NATs (NatA-F) have been proved to be involved in protein N-terminal acetylation and differ from each other in subunit composition, substrate preferences and phenotypes [1,9]. In contrast with the high abundance of N-acetylated proteins in eukaryotes, very few proteins were found to be acetylated in bacteria. Up to present, only four proteins including RimI, RimJ and RimL from *Salmonella enterica* and YhY from *Escherichia coli* with NAT activity have been reported and are believed to be homologues of eukaryotic NATs [10,11].

In the present study, two NATs of GlmA and RmNag derived from *C. acetobutylicum* and *R. miehei*, respectively, were unexpectedly found to be capable of acetylating the free amino group on the very first glycine residue of peptide in spite of varied catalytic efficiency. GlmA and RmNag have been defined as glucosamine NATs and clustered into different phylogenetic groups from human NAT Naa10p. Glucosamine N-acetyltransferases (EC 2.3.1.3) catalyze the transfer of an acetyl group from Acetyl-CoA to a glucosamine (GlcN) acceptor generating a N-acetylglucosamine (GlcNAc). The enzymes participating in amino sugar metabolism also have been found in plants. Albeit distinct preferences showed by GlmA and Naa10p toward GlcN and peptide with N-termini of Gly, RmNag exhibits comparable activity toward them. The distinct specificity of GlmA, RmNag and Naa10p probably resulted from the divergent evolution, and meanwhile the extended knowledge of substrate selectivity of these NATs would also facilitate the application of them in the acetylation of hexosamines or specific peptides.

* Corresponding author.
E-mail address: jiansongcheng@nankai.edu.cn (J. Cheng).
[1] The authors contributed to this study equally.

https://doi.org/10.1016/j.carres.2018.11.011
Received 13 October 2018; Received in revised form 18 November 2018; Accepted 21 November 2018
Available online 22 November 2018
0008-6215/

2. Results and discussion

2.1. GlmA, RmNag and Naa10p are distantly related

GlmA, a bacterial glucosamine N-acetyltransferase catalyzing the transfer of an acetyl moiety from Acetyl-CoA to the primary amino group of GlcN [12,13], was recently employed by our group for production of glycopeptides carrying natural eukaryotic N-glycans. In the approach, GlcN linked peptides made by an engineered N-glycosyltransferase (ApNGT^{Q469A}) were further modified by GlmA to generate GlcNAcylated peptides [14]. Herein, GlmA was unexpectedly found to be capable of acetylating the free amino group on the very first glycine residue of peptide in spite of low efficiency. For better understanding the selectivity of various N-acetyltransferases, we next selected and expressed two distantly related enzymes for further study. One of them is characterized as a glucosamine N-acetyltransferase (RmNag) from fungi *R. miehei*, while another is the recombinant the catalytic subunit (Naa10p) of human N-acetyltransferase NatA complex. RmNag is a bifunctional protein with a β-N-acetylglucosaminidase domain in the N-terminal region followed by a glucosamine N-acetyltransferase domain, and was proposed to be involved in chitin metabolism of fungi [15,16]. Though both GlmA and RmNag were defined as glucosamine N-acetyltransferases, they shared only 47% sequence similarity. All three proteins are distantly related and fall into three different phylogenetic groups [17] (Fig. 1).

2.2. All three N-acetyltransferases are capable of acetylating GlcN

The specificity of GlmA, RmNag and Naa10p toward hexosamines was probed by using three monosaccharides including GlcN, GalN and ManN. The thin-layer chromatography (TLC) analysis indicated that all three NATs are capable of acetylating GlcN despite of low efficiency exhibited by Naa10p (Fig. 2). ManN cannot be accepted by all three NATs as substrate (Fig. S1). GalN, however, was partially acetylated by the fungal RmNag after 2 h incubation at 37 °C (Fig. S1).

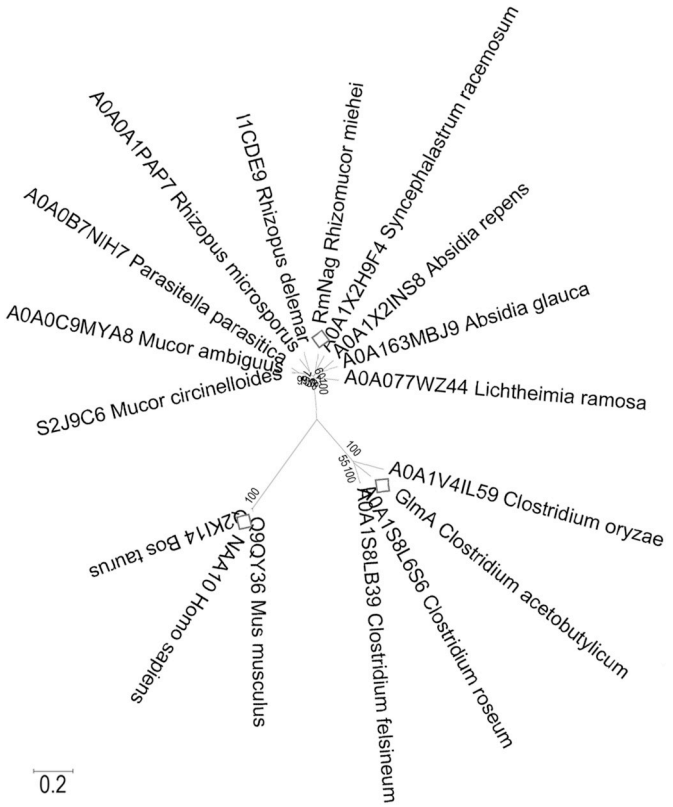

Fig. 1. Phylogenetic tree of three N-acetyltransferases. Neighbor-Joining (NJ) method was used to create this unrooted tree. Bootstrap values (expressed as percentages of 1000 replications) greater than 50% are indicated above the branches. The protein identification numbers are followed by species information.

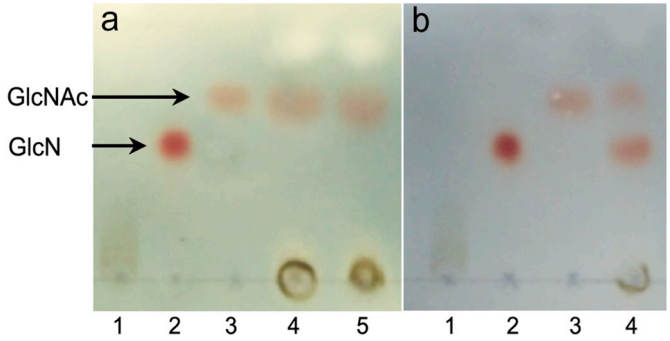

Fig. 2. TLC analyses of acetylation products yielded by GlmA, RmNag (a) and Naa10p (b). Lane 1: Acetyl-CoA; Lane 2: GlcN; Lane 3: GlcNAc; Lane 4a: reaction mixture of GlmA; Lane 5a: reaction mixture of RmNag; Lane 4b: reaction mixture of Naa10p.

2.3. Both GlmA and RmNag are active toward peptide starting with Gly

By scanning a penal of peptides starting with 20 various common amino acids, we found that the peptide containing N-terminal Gly is able to be exclusively recognized by GlmA and RmNag in spite of varied acetylation efficiency (Fig. 3). Further investigation by using two sets of peptides with varied amino acids at next two positions (+2 or +3) following the N-terminal Gly (+1) indicated that negative charged residues are disliked by both enzymes at either +2 or +3 position. In contrast, eukaryotic proteins with Ser, Thr and Ala termini are most frequently acetylated by NatA, and these residues, along with Met and Gly account for over 95% of the N-terminal acetylated residues. Besides, acidic residues, Asp or Glu, at the succeeding position are preferred for NatA [9]. Since blast search revealed the presence of NatA ortholog in *R. miehei* (Fig. S2), it remained unclear whether the NatA ortholog and RmNag might play roles in a collaboration manner.

Generally, the selectivity of both GlmA and RmNag at position +3 is more flexible than that at position +2. In comparison with GlmA, RmNag exhibited more efficient activity and flexible selectivity toward peptides (Fig. 4).

The thin-layer chromatography (TLC) analysis showed that both GlmA and RmNag are able to efficiently modify GlcN (Fig. 2a).

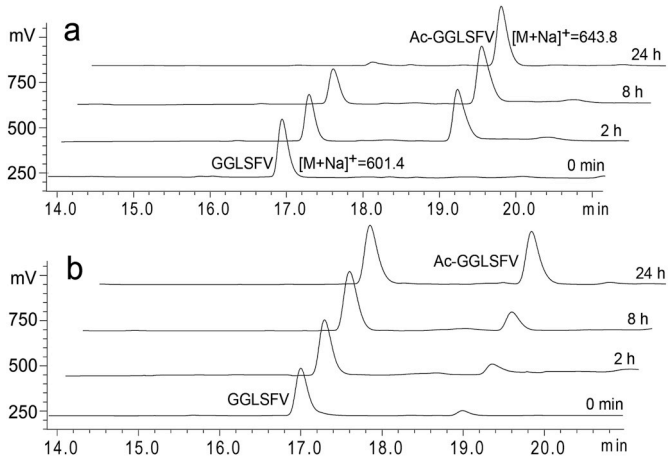

Fig. 3. The peptide GGLSFV starting with glycine can be acetylated by RmNag (a) and GlmA (b), generating Ac-GGLSFV.

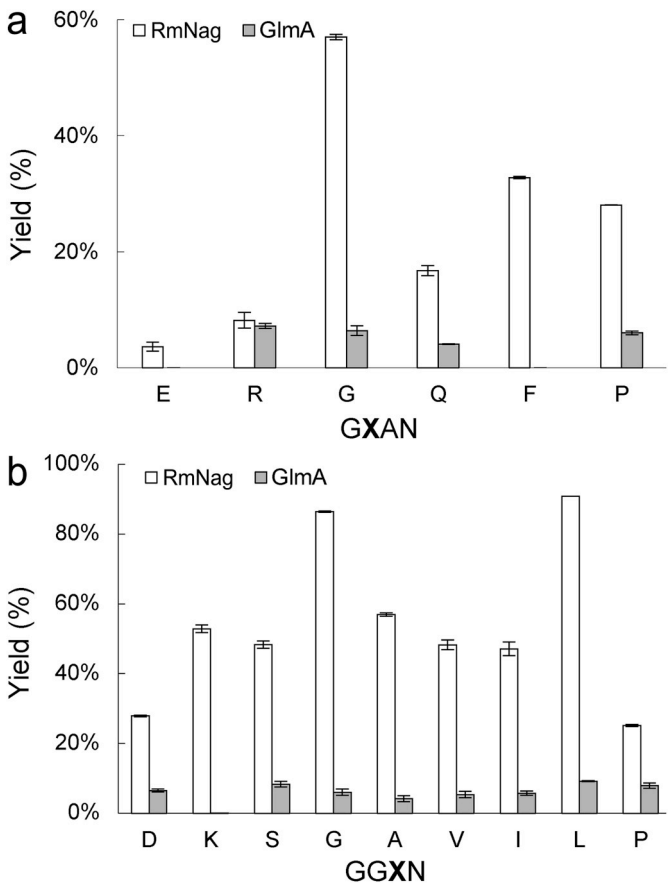

Fig. 4. The acetylation conversion rate of GlmA and RmNag toward a penal of peptides bearing varied residues at +2 position (a) or +3 position (b). For each position, representative amino acids containing basic, acidic, hydroxyl, small hydrophobic, or aromatic side chains were evaluated.

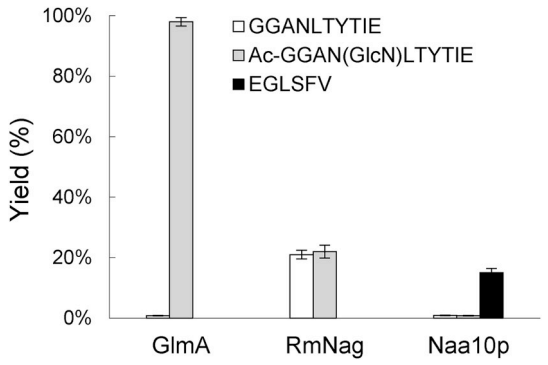

Fig. 5. The acetylation conversion rate of GlmA, RmNag and Naa10p toward GGANLTYTIE and Ac-GGAN(GlcN)LTYTIE, as well as EGLSFV (preferred by Naa10p). All acetylation assays were duplicated, performed at 37 °C for 2 h and analyzed by RP-HPLC. Acetylated products were further identified by MALDI-TOF-MS (Figs. S3–S6).

Therefore, to compare the efficiency of GlmA and RmNag toward GlcN moiety and Gly-starting peptide, a GlcN-linked Ac-peptide (Ac-GGAN(GlcN)LTYTIE) where the GlcN was transferred to Asn residue of the N-glycosylation sequon (Asn-Leu-Thr) by an engineered N-glycosyltransferase ApNGT^{Q469A} [18], and an unmodified peptide with same amino acid sequence (GGANLTYTIE) were used as acceptors and assayed. As shown in Fig. 5, the amine group of GlcN is much more preferred by GlmA in comparison with that of the N-terminal Gly, whereas RmNag showed almost equal activity against both acceptor substrates. The more flexible acceptor selectivity of RmNag might imply more roles in cell than previously proposed (involvement in chitin metabolism of fungi). On the other hand, the preference of GlmA toward GlcN make it a better catalyst for producing GlcNAcylated peptides or proteins. PA3944, a Gcn5-related acetyltransferase derived from *Pseudomonas aeruginosa*, has been proven to accept both glycine and GlcN and exhibits around 3-fold efficiency for the former [19].

2.4. Naa10p is a N-acetyltransferase for specific peptides

Naa10p is the catalytic subunit of the N-terminal acetyltransferase NatA and can independently acetylate specific peptides or proteins [20,21]. It is reported that Naa10p appears to favor peptides starting with acidic residues, such as Glu-starting N-termini [22,23]. Herein, the assays containing different monosaccharides (GlcN, GalN or ManN) indicated that Naa10p can also acetylate GlcN generating N-acetylglucosamine (Fig. 2b), while the acetylation efficiency is relatively low in comparison with that catalyzed by the bacterial GlmA and the fungal RmNag. The preference of Naa10p toward GlcN and peptide was also explored by using a pair of substrates Ac-GGAN(GlcN)LTYTIE and EGLSFV with acidic N-termini. As shown in Fig. 5, peptide with the acidic the N-termini is much more preferred by Naa10p in comparison with GlcN.

3. Conclusions

Taken together, the selectivity of three distantly related NATs was comprehensively explored and compared. GlmA from *C. acetobutylicum* is able to exclusively acetylate GlcN instead of GalN and ManN, and exhibits much more efficiency toward GlcN as opposed to peptide. In contrast, the human Naa10p preferred peptide with acidic N-termini to GlcN. Therefore, GlmA and Naa10p are appropriate to be defined as glucosamine N-acetyltransferase and peptide/protein N-terminal acetyltransferase, respectively, as previously reported. However, RmNag from *R. miehei* showed matched acetylation activity against both GlcN and peptide with N-terminal Gly, indicating that RmNag may play more roles in cell than previously proposed (involvement in chitin metabolism of fungi). NATs diverge from other members of the GNAT superfamily and differ from each other in the type of substrate N-termini preferred. The systematic comparison of their specificity would facilitate selecting suitable catalyst in acetylation of various molecules (GlcN or peptide/protein). For example, the yeast NatB complex has recently been transferred to *E. coli* for exclusively acetylating α-synuclein carrying a Met-Asp-N-termini [24].

4. Experimental

4.1. Materials

Acetyl-CoA was purchased from Sigma-Aldrich (St.Louis, MO, USA). Glucosamine (GlcN), galactosamine (GalN), and mannosamine (ManN) were purchased from Aladdin (Shanghai, China). Peptides were synthesized as previously described [18]. GlcN-peptide was synthesized by an engineered N-glycosyltransferase (ApNGT^{Q469A}) as previously described [14].

4.2. Cloning, expression and purification of recombinant GlmA, RmNag and human Naa10p

Full-length synthetic genes (codon optimized for the expression system of *Escherichia coli*) of *glmA* (GenBank: AE001437) and *rmNag* (GenBank: KC357713, the Asp268 was substituted by Ala) were synthesized by Genewiz (Suzhou, China) and then cloned into expression vector pET22b(+) and pET28a, respectively. The synthetic gene of *naa10p* (GenBank: KJ892483) was cloned into pMAL-c2x as a N-terminal maltose-binding protein (MBP)-fused and His_6-tagged recombinant protein. The expression strain *E. coli* BL21 (DE3) harboring the

corresponding plasmids were cultured in LB medium (50 μg/mL ampicillin for GlmA and Naa10p, or 50 μg/mL kanamycin for RmNag) with vigorous sharking. When the optical density at 600 nm (OD_{600}) reached 0.8, 0.1 mM of isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added and then incubated at 18 °C for 20 h.

Cells were harvested by centrifugation (5000 × *g*, 4 °C, 20 min), resuspended in lysis buffer (100 mM Tris-HCl, 5 mM NaCl, pH 8.0), and then broken by sonication. After centrifugation (18,000 × *g*, 4 °C, 30 min) of the cell lysate, the supernatant was loaded onto a Ni^{2+}-NTA affinity column (Qiagen, Hilden, Germany). Before the lysate was loaded, the column was pre-equilibrated with 10 column volumes of binding buffer (10 mM imidazole, 0.5 M NaCl and 20 mM Tris-HCl, pH 8.0). After washing with 10 column volumes of binding buffer and washing buffer (30 mM imidazole, 0.5 M NaCl and 20 mM Tris-HCl, pH 8.0), the target proteins were eluted with elution buffer (250 mM imidazole, 0.5 M NaCl and 20 mM Tris-HCl, pH 8.0). Elution fractions were analyzed by 12% SDS-PAGE with Coomassie staining. The concentrations of purified enzymes were obtained using a BCA Protein Assay Kit (Pierce Biotechnology, Rockford, IL) with bovine serum albumin as a protein standard.

4.3. Acetyltransferase assays against hexoamines and peptides

The acceptor specificity of GlmA, RmNag and Naa10p toward hexoamines was explored by using GlcN, GalN and ManN. The reactions were carried out in a final volume of 30 μL containing 100 mM Tirs-HCl (pH 7.0 for GlmA; pH 7.5 for RmNag; and pH 8.0 for Naa10p), 5 mM varied hexoamines (GlcN, GalN or ManN), 1.5 mM Acetyl-CoA and 10 μM different enzymes (GlmA, RmNag or Naa10p). The reaction mixtures were incubated at 37 °C for 2 h, and then 1 μL portions of the reaction mixtures were applied to thin-layer chromatography (TLC, silica gel F254, Merck) analysis. The reaction mixtures were developed by *n*-butanol/methanol/ammonia/water = 5:4:2:1 (V/V/V/V) and stained with *p*-anisaldehyde solution (ethanol/sulfuric acid/*p*-anisaldehyde/acetic acid = 500:27:16:5.5). Reaction mixtures without acceptor substrate were used as controls.

The acceptor specificity and preference of GlmA and RmNag toward peptides were explored by using 20 synthetic hexapeptides (**X**-GLSFV where X presents one of the 20 common amino acids) and 15 synthetic peptides (bearing varied amino acid at +2 or +3 position). The reactions were carried out in a final volume of 30 μL containing Tris-HCl (100 mM), peptide (1 mM), Acetyl-CoA (1.5 mM) and enzyme (10 μM). The reaction mixtures were incubated at 37 °C with different incubation times. 10 μL portions of the reaction mixtures were then analyzed by reversed-phase high-performance liquid chromatography (Rp-HPLC, Shimadzu SPD-20A instrument equipped with a Shim-pack VP-ODS C18 column, 5 μm, 4.6 mm × 250 mm). All assays were carried out in duplicate. The acetylated polypeptides were further identified by MALDI-TOF spectrum analysis (Autoflex Speed mass spectrometer, Bruker).

Acknowledgments

Financial support for this work was provided by the National Science Foundation grants: 21372130 and 81302682.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.carres.2018.11.011.

References

[1] B. Polevoda, F. Sherman, Nalpha -terminal acetylation of eukaryotic proteins, J. Biol. Chem. 275 (2000) 36479–36482.
[2] T. Arnesen, Towards a functional understanding of protein N-terminal acetylation, PLoS Biol. 9 (2011) e1001074.
[3] T. Kouzarides, Acetylation: a regulatory modification to rival phosphorylation? EMBO J. 19 (2000) 1176–1179.
[4] K.K. Starheim, K. Gevaert, T. Arnesen, Protein N-terminal acetyltransferases: when the start matters, Trends Biochem. Sci. 37 (2012) 152–161.
[5] T. Freier, H.S. Koh, K. Kazazian, M.S. Shoichet, Controlling cell adhesion and degradation of chitosan films by N-acetylation, Biomaterials 26 (2005) 5872–5878.
[6] H. Aksnes, K. Hole, T. Arnesen, Molecular, cellular, and physiological significance of N-terminal acetylation, Int. Rev. Cell Mol. Biol. 316 (2015) 267–305.
[7] F. Dyda, D.C. Klein, A.B. Hickman, GCN5-related N-acetyltransferases: a structural overview, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 81–103.
[8] L. Favrot, J.S. Blanchard, O. Vergnolle, Bacterial GCN5-related N-acetyltransferases: from resistance to regulation, Biochemistry 55 (2016) 989–1002.
[9] B. Polevoda, F. Sherman, N-terminal acetyltransferases and sequence requirements for N-terminal acetylation of eukaryotic proteins, J. Mol. Biol. 325 (2003) 595–622.
[10] J. Soppa, Protein acetylation in archaea, bacteria, and eukaryotes, Archaea 2010 (2010).
[11] S. Tanaka, Y. Matsushita, A. Yoshikawa, K. Isono, Cloning and molecular characterization of the gene *rimL* which encodes an enzyme acetylating ribosomal protein L12 of *Escherichia coli* K12, Mol. Gen. Genet. 217 (1989) 289–293.
[12] J. Reith, C. Mayer, Characterization of a glucosamine/glucosaminide N-acetyltransferase of Clostridium acetobutylicum, J. Bacteriol. 193 (2011) 5393–5399.
[13] B.J. Dopkins, P.A. Tipton, J.B. Thoden, H.M. Holden, Structural studies on a glucosamine/glucosaminide N-acetyltransferase, Biochemistry 55 (2016) 4495–4508.
[14] Y. Xu, Z. Wu, P. Zhang, H. Zhu, H. Zhu, Q. Song, L. Wang, F. Wang, P.G. Wang, J. Cheng, A novel enzymatic method for synthesis of glycopeptides carrying natural eukaryotic N-glycans, Chem. Commun. (Camb). 53 (2017) 9075–9077.
[15] S. Yang, S. Song, Q. Yan, X. Fu, Z. Jiang, X. Yang, Biochemical characterization of the first fungal glycoside hydrolase family 3 beta-N-acetylglucosaminidase from *Rhizomucor miehei*, J. Agric. Food Chem. 62 (2014) 5181–5190.
[16] Z. Qin, Y. Xiao, X. Yang, J.R. Mesters, S. Yang, Z. Jiang, A unique GCN5-related glucosamine N-acetyltransferase region exist in the fungal multi-domain glycoside hydrolase family 3 beta-N-acetylglucosaminidase, Sci. Rep. 5 (2015) 18292.
[17] K.D. Yamada, S. Omori, H. Nishi, M. Miyagi, Identification of the sequence determinants of protein N-terminal acetylation through a decision tree approach, BMC Bioinf. 18 (2017) 289.
[18] Q. Song, Z. Wu, Y. Fan, W. Song, P. Zhang, L. Wang, F. Wang, Y. Xu, P.G. Wang, J. Cheng, Production of homogeneous glycoprotein with multisite modifications by an engineered N-glycosyltransferase mutant, J. Biol. Chem. 292 (2017) 8856–8863.
[19] M.L. Kuhn, K.A. Majorek, W. Minor, W.F. Anderson, Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity, Protein Sci. 22 (2013) 222–230.
[20] R.S. Magin, Z.M. March, R. Marmorstein, The N-terminal acetyltransferase Naa10/ARD1 does not acetylate lysine residues, J. Biol. Chem. 291 (2016) 5270–5277.
[21] B. Popp, S.I. Stove, S. Endele, L.M. Myklebust, J. Hoyer, H. Sticht, S. Azzarello-Burri, A. Rauch, T. Arnesen, A. Reis, De novo missense mutations in the NAA10 gene cause severe non-syndromic developmental delay in males and females, Eur. J. Hum. Genet. 23 (2015) 602–609.
[22] P. Van Damme, R. Evjenth, H. Foyn, K. Demeyer, P.J. De Bock, J.R. Lillehaug, J. Vandekerckhove, T. Arnesen, K. Gevaert, Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of N(alpha)-acetyltransferases and point to hNaa10p as the post-translational actin N(alpha)-acetyltransferase, Mol. Cell. Proteomics 10 (2011) M110 004580.
[23] G. Liszczak, J.M. Goldberg, H. Foyn, E.J. Petersson, T. Arnesen, R. Marmorstein, Molecular basis for N-terminal acetylation by the heterodimeric NatA complex, Nat. Struct. Mol. Biol. 20 (2013) 1098–1105.
[24] M. Rovere, A.E. Powers, D.S. Patel, T. Bartels, pTSara-NatB, an improved N-terminal acetylation system for recombinant protein expression in *E. coli*, PLoS One 13 (2018) e0198715.