Identification and characterization of a novel GNAT superfamily N\(^{\alpha}\)-acetyltransferase from Salinicoccus halodurans H3B36

Xiaochen Ma,\(^1\) Kai Jiang,\(^2\) Cheng Zhou,\(^1\) Yanfen Xue\(^1\) and Yanhe Ma\(^1\)

\(^1\)Institute of Microbiology, CAS, Beijing, 100101, China.
\(^2\)College of Life Science and Technology, Inner Mongolia Normal University, Hohhot, Inner Mongolia 010022, China.

Summary

N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine was found as a new type of compatible solutes that acted as an organic cytoprotectant in the strain of Salinicoccus halodurans H3B36. A novel lysine N\(^{\alpha}\)-acetyltransferase gene (shkat), encoding an enzyme that catalysed the acetylation of lysine exclusively at \(\alpha\) position, was identified from this moderate halophilic strain and expressed in Escherichia coli. Sequence analysis indicated ShKAT contained a highly conserved pyrophosphate-binding loop (Arg-Gly-Asn-Gly-Asn-Gly), which was a signature of the GNAT superfamily. ShKAT exclusively recognized free amino acids as substrate, including lysine and other basic amino acids. The enzyme showed a wide range of optimal pH value and was tolerant to high-alkali and high-salinity conditions. As a new member of the GNAT superfamily, the ShKAT was the first enzyme recognized free lysine as substrate. We believe this work gives an expanded perspective of the GNAT superfamily, and reveals great potential of the shkat gene to be applied in genetic engineering for resisting extreme conditions.

Introduction

Microorganisms have to cope with a variety of extreme conditions to survive, grow and proliferate. One common strategy is to accumulate organic protectants termed as compatible solutes (Lentzen and Schwarz, 2006; Köcher et al., 2011). Compatible solutes are water-soluble, low molecular mass organic compounds, which help microbial cells to adapt to osmotic, heat or cold stresses and other changing environments (Brown and Simpson, 1972; Costa et al., 1998). These compounds could accumulate into extremely high concentration through de novo synthesis or external uptake, without interfering with cell metabolism. In the biotechnological field, compatible solutes are used as stabilizers of proteins, nucleic acids and other cellular macromolecules and protectants for cosmetics and medicine (Lentzen and Schwarz, 2006; Czech et al., 2018). According to their chemical structures, compatible solutes can be classified into several groups: carbohydrates, polyols, phosphodiesters, amino acids and their derivatives (Kets et al., 1996; Eibein, 2003; Santos et al., 2007).

In our previous studies, a new type of compatible solutes, N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine, was found as an organic osmolyte and thermolyte in moderate halophile Salinicoccus halodurans H3B36 (Jiang et al., 2015a,b). This strain was isolated from a sediment sample at 3.2 m vertical depth of Qaidam Basin (China), with an optimal growing temperature of 30°C. Under heat stress (42°C), the accumulation of N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine exhibited approximately a threefold increase. The de novo synthesis of N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine has been reported by our group (Jiang et al., 2015a,b), which started from aspartate and went through the acetyl-dependent diaminopimelic acid pathway (DAP) to form lysine. An 8-kb cluster (orf_1582-orf_1589) containing 8 genes was predicted to be involved in this process. Subsequently, an unknown acetyltransferase was responsible for synthesizing the final product. Alternatively, N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine could be generated from exogenous lysine directly. Fig. 1A shows the scheme of the N\(^{\alpha}\)-acetylation of lysine.

Two analogues of N\(^{\alpha}\)-acetyl-\(\alpha\)-lysine, N\(^{\alpha}\)-acetyl-\(\beta\)-lysine and N\(^{\alpha}\)-acetyl-ornithine, were also reported as compatible solutes (Fig. 1B). The osmotic stress caused by NaCl had a significant effect on the accumulation of these two N-acetylated amino acids (Sowers et al., 1990; Wohlfarth et al., 1993; Xavier et al., 2011). At physiological pH, N-acetylation can transform basic amino acids such as lysine and ornithine into uncharged and highly water-soluble zwitterionic molecules, thus
satisfying the requirement of compatible solutes (Empadinhas and Costa, 2008). Nα-acetyl-β-lysine has been widely reported in several microorganisms, such as green sulphur bacteria, Bacillus cereus CECT 148 and methanogenic archaea (Pfluger et al., 2003; Xavier et al., 2011; Majorek et al., 2013). The synthesis of Nα-acetyl-β-lysine involves two enzymes: a lysine-2,3-aminomutase (AblA) which converts α-lysine to the intermediate β-lysine and an acetyltransferase (AblB) which is responsible for its acetylation (Roberts, 2005). Martin et al. reported that the enzymatic activity of AblA in Methanococcus thermolithotrophicus at normal conditions was at least eightfold lower than cells grown in a higher-salinity environment (Martin et al., 2001). The yodP–kamA genes encoding AblA- and AblB-like enzymes were also found in the Bacillaceae family, and Nα-acetyl-β-lysine could be synthesized by overexpressed its own yodP–kamA genes using a heterologous promoter in the B. subtilis strain. However, Nα-acetyl-β-lysine was not involved in cellular defence against osmotic stress (Muller et al., 2011). Another homologous compound, Nα-acetyl-α-lysine, was found only in two strains: Planococcus sp. VITP21 and Salinicoccus hispanicus (Delmoral et al., 1994; Joghee and Jayaraman, 2014). In VITP21, Nα-acetyl-α-lysine was confirmed to be synthesized by direct acetylation of lysine. However, no further information regarding the acetyltransferase has been reported.

In this present study, the Nα-acetyltransferase from Salinicoccus halodurans H3B36 was identified. The corresponding gene shkat was cloned and successfully expressed in E. coli. Enzymatic characteristics and catalytic activity of the recombinant protein were evaluated for the first time.

Results and discussion

Identification of the novel Nα-acetyltransferase from Salinicoccus halodurans H3B36

Previously, we found the Nα-acetyl-α-lysine accumulation increased significantly under heat shock. Therefore, the whole-RNA sequencing of S. halodurans H3B36 was performed under optimal growth (30°C) and heat stress conditions (42°C) to identify the genes involved in Nα-acetyl-α-lysine accumulation. An expression level of > 1.5-fold changes and a P-value of ≤ 0.05 was defined as a threshold. To investigate the differentially expressed genes, we used the KEGG and GO databases for functional classification and enrichment analysis. Four genes annotated as acetyltransferases in the transcriptomic data, orf_1585, orf_793, orf_2506 and orf_442, were found to be up-regulated 7.1-, 1.8-, 1.6- and 1.5-fold, respectively, with heat shock. These four genes were constructed into expression plasmids and transformed into E. coli BL21 (DE3). As E. coli does not produce endogenous Nα-acetyl-α-lysine, it is an ideal system for testing the functionality of the corresponding enzymes. After PCR and sequencing confirmation, the positive clones were cultured and induced with IPTG for gene expression. The compatible solute extract was tested using high-performance liquid chromatography (HPLC).
coupled with a diode array detection (DAD) (Fig. 2A). The red, green, pink and purple lines in the HPLC profiles represented the clones transformed with orf_1585, orf_793, orf_2506 and orf_442 respectively. The N'-acetyl-α-lysine was only detected in the clone transformed with pET28a-orf_793. Mass spectrometry (MS) was also applied to further confirm its molecular weight (Fig. 2A, insert). The result provided strong evidence that orf_793 (shkat) was the gene encoding N'-acetyltransferase. In addition, as this gene can be successfully expressed in E. coli, it has the potential in genetic engineering for producing exogenous compatible solutes.

Sequence and structure analysis of ShKAT

The nucleotide sequence of ShKAT is 432 bp, which encodes a 143 amino acid protein. The sequence was searched in the BLAST database to identify related sequences. From the result, the ShKAT enzyme belongs to the GNAT (general control non-repressible 5 (GCN5)-related N-acetyltransferases) superfamily and shares relatively high sequence similarity with N-acetyltransferases from Jeotgalicoccus coquinae (66%, NCBI Reference Sequence: WP_184281624.1), Jeotgalicoccus sp. ATCC 8456 (65%, WP_198687769.1) and Jeotgalicoccus meleagris (64%, WP_185125999.1).

Until the present, more than 710000 members of the GNAT superfamily (known as GNATs) have been identified in prokaryotes, eukaryotes and archaea (NCBI database https://www.ncbi.nlm.nih.gov/), participating in numerous cellular processes such as transcription regulation, stress resistance, antibiotic resistance and cell protection (Xie et al., 2014). The GNAT superfamily members are capable of transferring an acetyl group from acetyl-CoA to the amino group of a broad range of substrates, including protein, peptides and small molecules. Of those, histone N-acetyltransferase from Saccharomyces cerevisiae acetylates the conserved lysine residues on the histone H4 protein, a process involved in chromatin assembly and DNA repair (Dutnall et al., 1998). Non-Histone N-acetyltransferase of Sulfolobus solfataricus catalyses the acetylation of Lys16 of DNA-binding protein ALBA, which in turn leads to a decrease in DNA-binding affinity (Brent et al., 2009). The aminoglycoside N-acetyltransferases from Enterococcus faecium could acetylate one of the four amino groups present on aminoglycoside antibiotics. Of those, the AAC(6’)-I targets a broad range of substrates including histones and small basic proteins (Wright and Ladak, 1997; Angus-Hill et al., 1999). The N-acetyltransferases from Pseudomonas aeruginosa and methanogenic archaea specifically acetylate the C-terminal N' lysine of proteins and β-lysine respectively (Pfluger et al., 2003; Majorek et al., 2013). Notably, ShKAT is the first enzyme in the GNAT superfamily acting on the free lysine substrate.

Based on sequence or functional similarities, we aligned ShKAT, 3 top candidates from the BLAST search and several typical GNATs (Fig. 3). Consistent with their diverse functions, the sequences were only partially aligned. Subsequently, a phylogenetic tree was generated using GNATs sharing more than 45% amino acid identity with ShKAT (Fig. 4). ShKAT was classified as a unique branch and clustered relatively close to the acetyltransferases from Jeotgalicoccus sp. Most selected proteins are from Macroccocus sp. and Clostridium sp., which are evolutionarily distant from ShKAT.

Although the GNAT enzymes generally exhibit an overall low percentage of sequence identity, their structures often share conserved folds comprising of six or seven β-strands and four α-helices connected in an order of β0–β1–α1–β2–β3–β4–x3–β5–x4–β6 (Vetting et al., 2005a,b). Secondary structure elements jβ0, jβ6 and α2, might be absent in some GNAT enzymes (Salah Ud-Din et al., 2016). In general, there are two highly conserved features of GNAT superfamily proteins. One is the β-bulge at β4 forming a V shape, which accommodates acetyl-CoA. Another highly conserved sequence is the pyrophosphate-binding site (P-loop, Gln/Arg-X-Gly-Ala, where X is for any amino acid) connecting β4 to α3 (Favrot et al., 2016). The P-loop contributes to acetyl-CoA binding by forming hydrogen bonds between amides backbone nitrogen and phosphate oxygen atoms of acetyl-CoA. In ShKAT, the P-loop was composed of Arg-Gly-Asn-Gly-Asn-Gly (residues 80-86) as shown in Fig. 3.

To obtain more clues about the substrate recognition, the high-order structure of ShKAT was predicted by RosettaFold (http://rosetta.bakerlab.org) (Baek et al., 2021). The predicted structure reveals a mixed α/β-fold with a conserved acetyl-CoA-binding core region (Liszczak et al., 2011) composed of one α-helix (α3) and three β-strands (β4–β6) as shown in Fig. 5A. The predicted structure of ShKAT superimposes well with the four typical N-acetyltransferases: aminoglycoside N-acetyltransferase, histone N-acetyltransferase, non-histone protein N-acetyltransferase and aryalkylamine N-acetyltransferase (Fig. 5B), with a root mean square deviation (RMSD) of Cα atoms of 2.331 Å, 4.667 Å, 4.103 Å and 2.916 Å respectively. In addition to a high degree of superposition within the core domain, the entire β-sheet that cuts through the centre of the structure (β1–β7) and the α3 that flanks one side of the protein substrate-binding site superimpose well. A small displacement was found in the conserved substrate-binding region of ShKAT. This minor conformational
Fig. 2. (A) Detection of N³-acetyl-α-lysine in different colonies using HPLC. (B) HPLC chromatograms of Tris-HCl reaction buffer, N³-acetyl-α-lysine standard, N⁶-acetyl-α-lysine standard, lysine, Ac-CoA and the catalytic products of S. halodurans H3B36 N-acetyltransferase.
change in structure might influence the activation loops and cause functional differentiation.

Substrate specificity tests of ShKAT

For comprehensive evaluation of the enzyme, expression plasmid pET28a-shkat was constructed and transformed into E. coli BL21 (DE3). The recombinant ShKAT was purified by Ni-NTA affinity chromatography and HiTrap desalting column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). SDS/PAGE analysis (Fig. S1) revealed a single band at approximately 17 kDa. The specific activity of purified ShKAT was 28.9 U mg⁻¹ with lysine and acetyl-CoA as substrates.

Acetylation of lysine could occur at the α or ε amino group. To confirm the catalytic position of ShKAT in lysine solution, the purified recombinant enzyme was incubated with lysine and acetyl-CoA in Tris-HCl buffer (pH 10.0) at 40°C for 30 min. The product, along with N⁺-acetyl-ε-lysine and N⁺-acetyl-α-lysine standards were run in parallel on a HPLC system equipped with a DAD detector (Fig. 2B). The retention time for N⁺-acetyl-ε-lysine and N⁺-acetyl-α-lysine standard was 12 and 17 min respectively. In the HPLC profile of the catalytic product, the peak of N⁺-acetyl-α-lysine, but not N⁺-acetyl-ε-lysine, was present. These results indicated that this novel acetyltransferase can only catalyse N-acetylation at the α position.

Interestingly, most proteins of the GNAT superfamily specifically catalyse N-acetylation (acetylating the side chain of amino acid residues) in bacteria. Only several α-N-acetyltransferases were reported. RimI, RimJ and RimL target amino groups at α position of ribosomal protein S18, S5 and L12 involved in post-translational modification in Salmonella typhimurium and E. coli. (Yoshikawa et al., 1987; Vetting et al., 2005a,b, 2008). ESAT (secreted antigenic target)-6 protein was responsible for N-terminal Thr acetylation in 4 mycobacterial species (Okkels et al., 2004). Protein YiaC targets sirtuin deacylase CobB long isoform (CobBL) in human pathogen Salmonella enterica was found to modulate the deacylase activity (Parks and Escalante-Semerena, 2020).

To determine whether ShKAT could acetylate other amino acids, 20 more candidates were examined. The enzymatic activity was determined using a colorimetric assay that measures the release of CoA at 412 nm
Discovery of a novel N^α-acetyltransferase

WP 175287616.1 Salinibacillus roseus
WP 040104903.1 Salinibacillus roseus
WP 094906459.1 Salinibacillus roseus
WP 124010419.1 Salinibacillus roseus
WP 021547939.1 Salinibacillus luteus
WP 072709973.1 Salinibacillus alkaliphilus
WP 052256846.1 Salinibacillus sp. YB14-2
WP 092988249.1 Salinibacillus qingdaoensis
WP 198687769.1 Jeotgalicoccus sp. ATCC 8456
WP 184281624.1 Jeotgalicoccus coquinae
WP 046789619.1 Salinibacillus halodurans
WP 185125999.1 Jeotgalicoccus meleagrisis
WP 166626532.1 Jeotgalicoccus sp. S0W5
WP 092595314.1 Jeotgalicoccus aerolatus
WP 225742960.1 Marinilactibacillus sp. Marseille-P9653
WP 208559435.1 Marinilactibacillus sp. M4U5P12
WP 067061036.1 Streptococcus pantotheolipis
RFM18338.1 Clostridium botulinum
NFD28731.1 Clostridium botulinum
WP 012704737.1 Clostridium
WP 058008946.1 Clostridium sporogenes
WP 012720629.1 Clostridium botulinum
WP 021135425.1 Clostridium botulinum
WP 041349502.1 Clostridium botulinum
WP 191590323.1 Clostridium botulinum
WP 003358267.1 Clostridium botulinum
WP 161587090.1 Clostridium botulinum
WP 224187834.1 unclassified Macroccocus
WP 224185755.1 Macroccocus sp. 17Msa1131
WP 224184105.1 Macroccocus sp. JER37
WP 219493897.1 Macroccocus sp.
WP 133445619.1 Macroccocus caseolyticus
WP 101037305.1 Macroccocus caseolyticus
WP 101143712.1 Macroccocus caseolyticus
WP 219522043.1 Macroccocus caseolyticus
WP 012656300.1 Macroccocus caseolyticus
WP 103214328.1 Macroccocus caseolyticus
WP 099482682.1 Macroccocus caseolyticus
WP 120786706.1 Macroccocus caseolyticus
WP 086041926.1 Macroccocus canis
WP 133422148.1 Macroccocus canis
WP 133419411.1 Macroccocus canis
WP 164963133.1 Macroccocus canis
WP 211554142.1 Macroccocus canis
WP 133433830.1 Macroccocus canis
WP 210152648.1 Macroccocus canis
WP 138070557.1 Macroccocus canis
As shown in Table 1, ShKAT only acetylated l-lysine, arginine, histidine and ornithine, indicating this enzyme was specific towards basic amino acids. Ornithine contains one less methylene in the R group, arginine has a basic guanidine group, and histidine carries an imidazole ring. As ShKAT targets variable basic amino acids, it could be concluded substrate recognition is dependent on the positively charge, rather than specific structures. Besides, ShKAT was stereoisomer specific, as d-lysine could not be acetylated.

To demonstrate whether ShKAT works on lysine containing peptides, 3 tetrapeptides containing lysine were synthesized and tested. The tetrapeptide KGGG was used to test if ShKAT works on the a position of lysine in a peptide chain. GKGG, a peptide that contains only an -NH2 group, was used to confirm the acetylation position. And GGGG was used as a control. Compared with its activity on free lysine (100%), the relative activity of the enzyme on KGGG, GKGG and GGGG was only 1.7/C6 0.3, 2.9/C6 0.9 and 0.7/C6 0.5 percentage respectively. These results suggested that ShKAT particularly acts on free amino acids.

Characterization of the purified recombinant ShKAT

The effects of temperature and pH on recombinant ShKAT were studied using lysine and acetyl-CoA as substrates. As shown in Fig. 6, ShKAT exhibited the maximal activity at 40°C and maintained more than 90% of full activity at the temperature range of 30–45°C. The activity declined rapidly to 35% at 60°C. To examine the thermal stability of ShKAT, the enzyme was incubated without substrates at 30, 40 and 50°C respectively. ShKAT was completely inactive after 15 min incubation at 50°C and retained 58% and 28% of initial activity, respectively, after incubation at 30 and 40°C for 45 min.

As shown in Fig. 7, the enzymatic activity of recombinant ShKAT increased as the pH value raised from 6.5 to 8.0.
to an optimal pH of 10.0. Relatively high enzymatic activity (more than 75% of the full activity) was maintained at the pH range of 8.5 to 11.5. However, no enzymatic activity was observed at pH 12. ShKAT was tested in different buffers from pH 6.5 to 12.0 at 4°C for 1 h for pH stability testing. It showed strong pH stability from pH 6.5 to 11.0, with more than 90% of the original enzymatic activity preserved. Furthermore, the enzymatic activity was improved by 2%-10% after being stored in the buffer of pH 9-11, indicating a preference of basic environment for both storage and catalysis.

The effects of metal ions and chemical reagents are given in Table 2, and the enzymatic activity was partly inhibited by Fe³⁺, strongly inhibited by Ni²⁺ and Mn²⁺, and completely inhibited by Cu²⁺, Co²⁺ and Hg²⁺. Two metals, Ca²⁺ and Mg²⁺ facilitated enzymatic activity. As for chemical reagents, 1 mM SDS almost completely abolished the enzymatic activity. However, EDTA had no obvious effect, indicating ShKAT is not a metalloprotein. Methanol and ethanol showed partly inhibitory effects on ShKAT. The effects of NaCl on the enzymatic activity are shown in Fig. 8. Low concentration (0 M–2.0 M) of NaCl only slightly suppressed ShKAT, while higher NaCl concentration leads to further inhibition. When the NaCl concentration reached 3.0 M, 31.7% of enzymatic activity was retained.

Fig. 6. Effect of temperature on the (A) activity and (B) stability of recombinant ShKAT. Values represent means ± standard deviations of results from three technical replications. The 100% specific activity is 3 U mg⁻¹.

© 2022 The Authors. Microbial Biotechnology published by Society for Applied Microbiology and John Wiley & Sons Ltd. Microbial Biotechnology, 15, 1652–1665
The kinetic parameters of the ShKAT were determined using colorimetric enzyme assay with one substrate concentration varied and the other constant. Kinetic data were fitted to the Michaelis–Menten equation using GRAPHPAD PRISM 8 software, and the nonlinear regression plots are shown in Fig. 9. The $K_m$ values for Ac-CoA and lysine were 0.81 mM and 1.89 mM respectively. The $V_{max}$ values for Ac-CoA and lysine were 229.1 mM s$^{-1}$ mg$^{-1}$ and 714.0 mM s$^{-1}$ mg$^{-1}$ respectively. To calculate the $k_{cat}$ value, 18172 Da was used as the molecular weight of the His-tagged recombinant enzyme. The $k_{cat}$ value was 4.2 and 13.0 s$^{-1}$ for Ac-CoA and lysine respectively. Table 3 presents the kinetic parameters of ShKAT and other GNATs. The ShKAT exhibited relatively high values of kinetic parameters compared with other reported enzymes.

Fig. 7. Effect of pH on the (A) activity and (B) stability of recombinant ShKAT. Values represent means ± standard deviations of results from three technical replications. The 100% specific activity is 2.7 U mg$^{-1}$.

| Metal ions/chemical agents | Concentration | Relative activity (%) |
|---------------------------|---------------|-----------------------|
| No addition               | 0 mM          | 100                   |
| Ca$^{2+}$(CaCl$_2$)       | 5 mM          | 123 ± 3               |
| Mg$^{2+}$(MgCl$_2$)       | 5 mM          | 111 ± 3               |
| Cu$^{2+}$(CuSO$_4$)       | 5 mM          | 0.3 ± 0.2             |
| Fe$^{3+}$(FeCl$_3$)       | 5 mM          | 66 ± 8                |
| Mn$^{2+}$(MnCl$_2$)       | 5 mM          | 7 ± 1                 |
| Co$^{2+}$(CoCl$_3$)       | 5 mM          | 1.9 ± 0.7             |
| Hg$^{2+}$(HgCl$_2$)       | 5 mM          | 0.6 ± 0.1             |
| Ni$^{2+}$(Ni SO$_4$)      | 5 mM          | 22 ± 3                |
| SDS                       | 1 mM          | 2.7 ± 0.7             |
| EDTA                      | 1 mM          | 96 ± 4                |
| Methanol                  | 5%            | 52 ± 7                |
| Ethanol                   | 5%            | 66 ± 1                |

Values represent means ± standard deviations of results from three technical replications.
Conclusions

The compatible solutes could be applied in many areas. However, the application of lysine derivatives has not been widely explored regardless of their potential benefits due to their limited accessibility. For the methanogens host strains, anaerobic growth conditions are relatively restricted the growth yields. For the Bacillus family, the heterologous expression of the abl genes in E. coli and B. subtilis host strains is challenging (Muller et al., 2011). Our group previously reported Nα-acetyl-a-lysine as a new compatible solute capable of protecting bacteria from heat stress. In this work, the shkat gene was identified through in parallel RNA sequencing analysis from samples under normal and heat shock conditions. The corresponding protein ShKAT, a novel GNAT superfamily enzyme that acetylated lysine exclusively at the α position, was extensively characterized. The ShKAT is the first reported GNATs member that exclusively recognizes free lysine substrate. The crystal structure is worth further investigation to disclose the substrate selection mechanism. Overexpression of the shkat gene leads to accumulation of Nα-acetyl-α-lysine in E. coli suggesting genetic engineering applications in the construction of stress-tolerant strains and crops.

Experimental procedures

The heat shock experiments and total RNA extraction

The strain Salinicoccus halodurans H3B36 was cultured in GMH medium (5 g l⁻¹ casamino acids, 5 g l⁻¹ yeast extract, 4 g l⁻¹ MgSO₄.7H₂O, 2 g l⁻¹ KCl, 0.036 g l⁻¹, ...
Table 3. Comparison of enzyme kinetic parameters of ShKAT and other GNATs.

| Family                        | Source                     | Substrate             | Name                  | $K_{m}$             | $K_{cat}$         | Ref           |
|-------------------------------|----------------------------|-----------------------|-----------------------|--------------------|------------------|---------------|
| Aminoglycoside N-acetyltransferases | Mycobacterium smegmatis | Aminoglycosides       | AAC(2')-Id           | (6.10 ± 1.20) × 10 µM | 0.34 ± 0.03 s⁻¹ | Jeong et al. (2020) |
|                               |                            |                       | (AcCoA)               | (1.86 ± 0.10) × 10² µM | 3.34 ± 0.10 s⁻¹ |               |
|                               |                            |                       | (Tobramycin)          |                    | (Tobramycin)      |               |
| Histone N-acetyltransferases  | Homo sapiens               | H4 peptide            | HAT1                  | 6.68 ± 0.3 µM (AcCoA) | 4.14 ± 0.35 s⁻¹ | Wu et al. (2012) |
|                               |                            |                       |                       | 20.77 ± 0.47 µM (H4 peptide) | 4.28 ± 0.01 s⁻¹ |               |
| Spermidine/spermineN¹-acetyltransferases | Bacillus subtilis      | Spermidine/ spermine  | PaiA                  | 31 ± 3 µM (AcCoA)   | 0.64 ± 0.04 s⁻¹ | Forouhar et al. (2005) |
|                               |                            |                       |                       | 323 ± 80 µM (Spermidine) |                |               |
| Ribosomal protein N-acetyltransferases | Mycobacterium tuberculosis | Ribosomal protein S18 | MiRiml               | 3.63 ± 0.43 mM      | NA               | Hou et al. (2019) |
| Succinyltransferase           | Mycobacterium tuberculosis | Nucleoid-associated protein HU | Rv0802c       | 1.19 ± 0.23 µM      | 1.33 s⁻¹        | Anand et al. (2021) |
| Arylalkylamine N-acetyltransferases | Tribolium castaneum | Amine substrates       | AANAT1b              | 65 ± 9.8 µM (AcCoA) | 52 ± 2.5 s⁻¹    | O’Flynn et al. (2020) |
|                               |                            |                       |                       | 1400 ± 190 µM (Histamine) |                |               |

FeSO₄·7H₂O, 0.36 mg l⁻¹ MnCl₂·7H₂O and 60 g l⁻¹ NaCl) at 30°C to mid-exponential phase (OD₆₀₀ of 0.6) and continuously cultured at 42°C for 6 h. Then the cells were harvested by centrifugation at 10 000 g for 1 min and frozen in liquid nitrogen. Total RNA was isolated using grinding in liquid nitrogen combined with TRIzol (Invitrogen, Carlsbad, CA, USA) extraction according to the manufacturer’s instructions. The quality and quantity of the total RNA were determined with BIOANALYSER 2100 (Agilent, Palo Alto, CA, USA), NanoDrop (Thermo Fisher Scientific, Cleveland, OH, USA) and agarose gel electrophoresis.

Transcriptomic data analysis

The NGS QC Toolkit (v2.3) was used for sequence analysis. The pretreatment of raw reads was processed by removing low-quality bases, adaptors and the reads shorter than 75 bp after trimming. The software BOWTIE 2 (v2.1.0) was used to map the clean reads to genomic sequence and reference genes of strain S. halodurans H3B36 separately with default parameters. Gene expression values were computed by FPKM (fragments per kilo bases per million reads) normalization. The $P$-value was corrected using Benjamini false discovery rate (FDR). Genes with fold changes > 2 and adjusted $P$-values ≤ 0.05 in FPKM between the heat shock condition and normal condition were defined as differentially expressed.

Compatible solutes extraction and analysis

The cultured cells were obtained by centrifugation at 8000 g for 10 min. The pellets were washed twice, then freeze-dried and weighted. The extraction procedure was adapted from the protocol described by Bligh and Dyer (1959). The extraction was processed overnight with constantly stirring after adding the extraction solution of methanol/chloroform/water (10:5:3.4, v/v). An equal volume ratio of chloroform and water was added to the mixture. The centrifugation at 5000 g for 30 min was performed to promote phase separation after 1 h vigorous shaking. The compatible solutes were recovered in the aqueous top layer for further study.

The Agilent 1200 HPLC system was used to confirm the presence of N²-acetyl-α-lysine in the compatible solutes and 6520 Q-TOF MS was used for mass confirmation. Samples were passed through a 0.22 µm filter and applied to a ZORBAX NH2 column (883952-708, 150 mm × 4.6 mm, 5 µm particle size; Agilent). The mobile phase was isobaric acetonitrile/water (70:30, v/v), the column temperature was set at 30°C, and the flow rate was 1 ml min⁻¹. The effluent was introduced into the mass spectrometer directly and analysed in electrospray ionization mode (ES⁺). The standards were dissolved in a 70% acetonitrile aqueous solution. A LC-DAD (diode array detector, G1315D; Agilent) approach was applied to distinguish N²-acetyl-α-lysine and its isoforms with chromatographic condition described above. The detection wavelength was set at 210 nm.

Gene cloning, sequencing and construction of expression vectors

The putative genes were amplified by PCR using the primer pairs listed in Table S1. The PCR cycling was set as follows: denaturation at 94°C for 1 min; 30 cycles of 94°C for 30 s, 55°C for 1 min, 72°C for 90 s and a final
extension at 72°C for 10 min. The purified PCR product was digested with *BamH*II and *Hind*III to ligate into the corresponding sites of the pET28a vector. The resultant recombinant plasmid was transformed into competent *E. coli* BL21(DE3) cells. Transformed cells were grown on the LB agar plate at 37°C overnight. Selected clones were sequenced to confirm the transformation.

Expression and purification of ShKAT in *E. coli*

The *E. coli* BL21(DE3) clone harbouring pET28a- shkat plasmid was grown in LB (tryptone 10 g l⁻¹, yeast extract 5 g l⁻¹, NaCl 5 g l⁻¹) medium with 50 µg ml⁻¹ of kanamycin at 37°C to the OD₆₀₀ of 0.6–0.8. The cultures were then induced with 0.1 mM IPTG at 30°C (220 rpm) for 8 h. Cells were harvested by centrifugation at 6000 g for 10 min. Collected cells were resuspended in 20 mM binding buffer (20 mM Tris-HCl pH 8.0, 5 mM imidazole and 500 mM NaCl) and disrupted by ultrasonication. The lysate was obtained by centrifugation at 15 000 g for 10 min at 4°C and filtered through a 0.45 µm filter.

The Ni-NTA column was equilibrated with binding buffer and the supernatant was loaded following washing with binding buffer and washing buffer (20 mM Tris-HCl pH 8.0, 60 mM imidazole and 500 mM NaCl). The bound protein was then eluted with elution buffer ((20 mM Tris-HCl pH 8.0, 1.0 M imidazole and 500 mM NaCl), desalted using ÄKTA Puriﬁer and then analysed by SDS-PAGE(12%) and measuring the UV absorbance at 280 nm.

Enzyme assays

Enzyme assays were performed using the method described in Bode et al. (1993). This method measures the absorption at 412 nm caused by 5-thio-2-nitrobenzoate formation resulted from free CoA and 5,5'-dithiobis-(2-nitrobenzoate) (DTNB). Assays were carried out in 50 µl of 50 mM Tris–HCl buffer (pH 10.0) containing 10 mM lysine, 1 mM acetyl coenzyme A and 0.076 µg purified ShKAT at 40°C for 5 min. The reaction was stopped by addition of 150 µl of ethanol; then, 200 µl of 0.5 mM DTNB in 0.1 M Tris-HCl buffer was added for absorption measure. One unit of enzymatic activity was defined as the amount of enzyme that liberates 1 µmol of coenzyme A per min. All the measurements were repeated three times.

The substrate specificity

The substrate specificity of ShKAT was investigated by measuring the enzymatic activity towards various amino acids. All reactions were conducted in 50 mM Tris–HCl buffer (pH 10.0) at 40°C for 5 min. The relative activity with L-lysine as a substrate was set as 100%.

Three tetrapeptides were purchased as synthetic original compounds from Sangon Biotech, China (Shanghai, China). The purity of these samples was determined to be higher than 90% by liquid chromatography. The tetrapeptide KGKG was used to test if ShKAT works on the lysine in a peptide chain, GKKG was used to confirm the acetylation position, and GGGG was used as a control.

Sequences and structure analysis

The amino acid sequence was aligned with JALVIEW software. BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify the related sequences at the default parameters, the non-redundant protein sequences database and the blastp (protein–protein BLAST) algorithm were chosen, and all protein sequences were obtained from NCBI. A phylogenetic tree was constructed using the MEGA X software, a neighbour-joining (NJ) method was used and a bootstrap test was generated with 1000 replicates. The structure of ShKAT from sequence information is predicted by RoseTTAFold (https://robetta.bakerlab.org).

Enzyme characterization

The optimal temperature of the purified recombinant ShKAT was determined in 50 mM Tris–HCl buffer (pH 8.0) at 15°C to 60°C for 5 min. The optimal pH value was investigated at optimal temperature (40°C) in 20 mM MOPS buffer (pH 6.5–8.0), TAPS buffer (pH 8.0–9.0), CHES buffer (pH 9.0–10.0), CAPS buffer (pH 10.0–11.0) and Na₂PHO₄–NaOH (pH 11.0–12.0) for 5 min. The relative activity was defined as the percentage of activity against the highest activity. The thermal stability was determined by measuring the residual activity in standard conditions after incubating the enzyme at 30, 40 and 50°C, respectively, and sampled every 15 min. The stability of ShKAT at different pH values was tested by measuring the residual activity in standard conditions after incubating the enzyme in different buffers as described above at 4°C for 1 h.

To analyse the effects of various metal ions on the enzymatic activity, 5 mM of Ca²⁺, Mg²⁺, Cu²⁺, Fe³⁺, Mn²⁺, Co²⁺, Hg²⁺ and Ni²⁺ were added individually during the activity measurement. The effects of 1mM SDS and EDTA, 5% (v/v) organic reagents were also determined. The relative activity without any additive was set as 100%.

To test the effect of NaCl on enzymatic activity, various concentrations of NaCl (0–4.0 M) were added to the reaction mixture. The relative activity without NaCl addition was set as 100%.

© 2022 The Authors. Microbial Biotechnology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Microbial Biotechnology, 15, 1652–1665
Enzyme kinetic analysis

The kinetic parameters of the ShKAT were determined using colorimetric enzyme assay where one substrate concentration was varied and another was constant. The varied substrate concentrations were 0.5, 1.0, 1.5, 2.0, 3.0, 4.0, and 5.0 mM, while the other substrate concentration kept 5 mM. The absorbance measurements were taken at 412 nm in 50 mM Tris–HCl buffer (pH 10.0) at 40°C containing 0.76 μg purified ShKAT, 0.65 mM DTNB. The first 5 min of every reaction was used to determine the initial velocity. The values of the kinetic constant were calculated by fitting initial velocity data to the Michaelis–Menten equation using nonlinear regression on GRAPHPAD PRISM v 8.0 (GraphPad Software, San Diego, CA, USA).

Acknowledgements

This study was supported by grants from the National Key R&D Program of China (2018YFA0901400) and National Natural Science Foundation of China (31800672).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Anand, C., Santoshi, M., Singh, P.R., and Nagaraja, V. (2021) Rv0802c is an acyltransferase that succinylates and acetylates Mycobacterium tuberculosis nucleoid-associated protein HU. Microbiology 167: 1058.

Angus-Hill, M.L., Dutnall, R.N., Tafrov, S.T., Sternglanz, R., and Ramakrishnan, V. (1999) Crystal structure of the histone acetyltransferase Hpa2: a tetrameric member of the GCN5-related N-acetyltransferase superfamily. J Mol Biol 294: 1311–1325.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., et al. (2021) Accurate prediction of protein structures and interactions using a three-track neural network. Science 373: 871–876.

Bligh, E.G., and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. Can J Biochem Physiol 37: 911–917.

Bode, R.D., Thurau, A.M., and Schmidt, H. (1993) Characterization of acetyl-CoA: L-lysine N6-acetyltransferase, which catalyses the first step of carbon catabolism from lysine in Saccharomyces cerevisiae. Arch Microbiol 160: 397–400.

Brent, M.M., Iwata, A., Carten, J., Zhao, K., and Marmostein, R. (2009) Structure and biochemical characterization of protein acetyltransferase from Sulfolobus solfataricus. J Biol Chem 284: 19412–19419.

Brown, A.D., and Simpson, J.R. (1972) Water relations of sugar-tolerant yeasts: the role of intracellular polyols. J Gen Microbiol 72: 589–591.

Costa, M.S.D., Santos, H., and Galinski, E.A. (1998) An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 61: 117–153.

Czech, L., Hermann, L., Stöveken, N., Richter, A., Höppner, A., Smits, S., et al. (2018) Role of the extremeolys echinoe and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. Genes (Basel) 9: 177.

Delmoral, A., Sevier, J., Ramoscormenzana, A., Truper, H.G., and Galinski, E.A. (1994) Compatible solutes in new moderately halophilic isolates. FEMS Microbiol Lett 122: 165–172.

Dutnall, R.N., Tafrov, S.T., Sternglanz, R., and Ramakrishnan, V. (1998) Structure of the histone acetyltransferase Hat1. Cell 94: 427–438.

Elbein, A.D. (2003) New insights on trehalose: a multifunctional molecule. Glycobiology 13(4): 17R–27.

Empadinas, N., and Costa, M.S.D. (2008) Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. Int Microbiol 11: 151–161.

Favrot, L., Blanchard, J.S., and Vergnolle, O. (2016) Bacterial GCN5-related N-Acetyltransferases: from resistance to regulation. Biochemistry 55: 989–1002.

Forouhar, F., Lee, I.-S., Vujic, J., Vujic, S., Shen, J., Vorobiev, S.M., et al. (2005) Structural and functional evidence for Bacillus subtilis PiaA as a novel N1-spermidine/spermine acetyltransferase. J Biol Chem 280: 40328–40336.

Hou, M., Zhuang, J., Fan, S., Wang, H., Guo, C., Yao, H., et al. (2019) Biophysical and functional characterizations of recombinant RimI acetyltransferase from Mycobacterium tuberculosis. Acta Biochim Biophys Sin 51: 960–968.

Jeong, C.-S., Hwang, J., Do, H., Cha, S.-S., Oh, T.-J., Kim, H.J., et al. (2020) Structural and biochemical analyses of an aminoglycoside 2’-N-acetyltransferase from Mycolicium smegmatis. Sci Rep 10: 21503.

Jiang, K., Xue, Y., and Ma, Y. (2015a) Identification of N-(alpha)-acyl-alpha-lysine as a probable thermolyte and its accumulation mechanism in Salinicoccus halodurans H3B36. Sci Rep 5: 18518.

Jiang, K., Xue, Y., and Ma, Y. (2015b) Complete genome sequence of Salinicoccus halodurans H3B36, isolated from the Qaidam Basin in China. Stand Genomic Sci 10: 116.

Joghee, N.N., and Jayaraman, G. (2014) Metabolomic characterization of halophilic bacterial isolates reveals strains synthesizing rare diaminoacids under salt stress. Biochimie 102: 102–111.

Kets, E.P., Galinski, E.A., De Wit, M., De Bont, J.A., and Heipieper, H.J. (1996) Mannitol, a novel bacterial compatible solute in Pseudomonas putida S12. J Bacteriol 178: 6665–6670.

KöCher, S., Averhoff, B., and Müller, V. (2011) Development of a genetic system for the moderately halophilic bacterium Halobacillus halophilus: generation and characterization of mutants defect in the production of the compatible solute proline. Environ Microbiol 13: 2122–2131.
Lentzen, G., and Schwarz, T. (2006) Extremolytes: natural compounds from extremophiles for versatile applications. *Appl Microbiol Biotechnol* 72: 623–634.

Liszczak, G., Arnesen, T., and Marmorstein, R. (2011) Structure of a Ternary Naa50p (NAT5/SAN) N-terminal acetyltransferase complex reveals the molecular basis for substrate-specific acetylation. *J Biol Chem* 286: 37002–37010.

Majorek, K.A., Kuhn, M.L., Chruszcz, M., Anderson, W.F., and Minor, W. (2013) Structural, functional, and inhibition studies of a Gcn5-related N-Acetyltransferase (GNAT) superfamily protein PA4794: a new C-terminal lysine protein acetyltransferase from *Pseudomonas aeruginosa*. *J Biol Chem* 288: 30223–30235.

Martin, D.D., Ciulla, R.A., Robinson, P.M., and Roberts, M.F. (2001) Switching osmolyte strategies: response of *Methanococcus thermolithotrophicus* to changes in external NaCl. *Biochim Biophys Acta* 1524: 1–10.

Muller, S., Hoffmann, T., Santos, H., Saum, S.H., Bremer, E., and Muller, V. (2011) Bacterial abl-like genes: production of the archaeal osmolyte N(epsilon)-acetyl-beta-lysine by homologous overexpression of the yodP-kamA genes in *Bacillus subtilis*. *Appl Microbiol Biotechnol* 91: 689–697.

O’Flynn, B.G., Prins, K.C., Shepherd, B.A., Forbrich, V.E., Suarez, G., and Merker, D.J. (2020) Identification of catalytically distinct aminomutase and beta-lysine acetyltransferase genes of methanogenic archaea are salt induced and are essential for the biosynthesis of *N*epsilon-acetyl-beta-lysine and growth at high salinity. *Microbial Biotechnology* 13: 15895–15901.

Pfugler, K., Baumann, S., Gottschalk, G., Lin, W., Santos, H., and Muller, V. (2003) Lysine-2,3-aminomutase and beta-lysine acetyltransferase genes of methanogenic archaea are salt induced and are essential for the biosynthesis of *N*epsilon-acetyl-beta-lysine and growth at high salinity. *Appl Environ Microbiol* 69: 6047–6055.

Roberts, M.F. (2005) Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline Syst* 1: 5.

Salah Ud-Din, A.I.M., Tikhomirova, A., and Roujeunikova, A. (2016) Structure and functional diversity of GCN5-related N-acetyltransferases (GNAT). *Int J Mol Sci* 17: 1080.

Santos, H., Lamosa, P., Faria, T.Q., Borges, N. and Neves, C. (2007) The physiological role, biosynthesis, and mode of action of compatible solutes from (hyper)thermophiles. In *Physiology and Biochemistry of Extremophiles* C. Gerday & N. Glansdorff (eds). Washington, DC, USA: ASM Press.

Sowers, K.R., Robertson, D.E., Noll, D., Gunsalus, R.P., and Roberts, M.F. (1990) N epsilon-acetyl-beta-lysine: an osmolyte synthesized by methanogenic archeabacteria. *Proc Natl Acad Sci USA* 87: 9083–9087.

Vetting, M.W., Bareich, D.C., Yu, M., and Blanchard, J.S. (2008) Crystal structure of Riml from Salmonella typhimurium LT2, the GNAT responsible for (Nalpha)-acetylation of ribosomal protein S18. *Protein Sci* 17: 1781–1790.

Vetting, M.W., de Carvalho, L.P., Roderick, S.L., and Blanchard, J.S. (2005a) A novel dimeric structure of the RimL Nalpha-acetyltransferase from Salmonella typhimurium. *J Biol Chem* 280: 22108–22114.

Wohlfarth, A., Severin, J., and Galinski, E.A. (1993) Identification of N-Delta-acetylornithine as a novel osmolyte in some Gram-positive halophilic eubacteria. *Appl Microbiol Biotechnol* 39: 568–573.

Wright, G.D., and Ladak, P. (1997) Overexpression and characterization of the chromosomal aminoglycoside 6’-N-acetyltransferase from *Enterococcus faecium*. *Antimicrob Agents Chemother* 41: 956–960.

Wu, H., Moshkina, N., Min, J., Zeng, H., Joshua, J., Zhou, M.M., and Plotnikov, A.N. (2012) Structural basis for substrate specificity and catalysis of human histone acetyltransferase 1. *Proc Natl Acad Sci USA* 109: 8925–8930.

Xavier, T.-M., Xavier, V., and Galinski, E.A. (2011) Osmoadaptive accumulation of N-acetylated lysine in green sulfur bacteria and *Bacillus cereus*. *FEMS Microbiol Lett* 159: 167.

Xie, L., Zeng, J., Luo, H., Pan, W., and Xie, J. (2014) The roles of bacterial GCN5-related N-acetyltransferases. *Crit Rev Eukaryot Gene Expr* 24: 77–87.

Yoshikawa, A., Isono, S., Sheback, A., and Isono, K. (1987) Discovery of a novel N’-acetyltransferase. *Microbial Biotechnology* 1665.