Activity-Based Substrate Profiling for Gcn5-related N-acetyltransferases: The Use of Chloroacetyl-Coenzyme A to Identify Protein Substrates

Michael Yu, **Luiz Pedro Sorio de Carvalho**, **Guangxing Sun**, and **John S. Blanchard**
Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461.

The first members of what is now termed the GNAT superfamily of *N*-acetyltransferases were identified as aminoglycoside *N*-acetyltransferases that regioselectively acetylate one of the primary or secondary amine functions of this class of antibiotics[1], resulting in substantially decreased binding to the ribosomal A site in the 30S ribosomal subunit[2]. In 1992, the yeast transcriptional regulator, Gcn5, was shown to have sequence homology with these bacterial enzymes[3], and shortly thereafter, a *Tetrahymena* homologue of Gcn5 was shown to catalyze histone acetylation[4]. The Gcn5-related *N*-acetyltransferases comprise one of the largest enzyme superfamilies, with over 10,000 identified GNAT members in sequenced genomes[5]. GNAT superfamily members can be identified in all kingdoms of life, including archea, bacteria, plants and mammals. While the overall sequence homology can be very low, the core monomer fold, representing the acetylCoA (AcCoA) binding domain, is extraordinarily structurally conserved[5]. Although well over two dozen structures are now known, these unfortunately provide little indication of their acetylatable substrate specificity, and GNAT family members can regioselectively acetylate antibiotics (aminoglycosides, tabtoxin), metabolic intermediates (glucosamine-6-phosphate, serotonin, spermidine) and proteins, both internally on the ε-amino group of lysine residues of histones and metabolic enzymes, and at the α-N-terminal amino acid[5].

In *E. coli*, twenty-six GNAT's can be identified in the genome. Only one has been expressed and functionally characterized as a spermine/spermidine *N*-acetyltransferase[6]. Three others, RimI, RimJ and RimL, have been proposed to be the α-N-acetyltransferases that acetylate the ribosomal proteins S18, S5 and L12, respectively[7,8]. The other twenty-two *E. coli* GNAT's have no known, or hypothesized, function. We have cloned both the *rimL* and *L12* genes from *Salmonella enterica*, heterologously expressed these proteins in *E. coli*, and shown that L12 is a bona fide substrate for RimL[9]. Having a functional *N*-acetyltransferase and its cognate protein substrate, we wished to explore the development of a reagent and method to identify the substrate for any acetyltransferase.

Using a spectrophotometric assay to measure the thiol of CoA, we demonstrated that chloroacetyl-CoA (ClAcCoA) is a substrate for RimL (Scheme 1B), exhibiting a k_{cat}/K_m value only 10 times lower than AcCoA (Figure S1,S2). We selected chloroacetylCoA over other haloacetylCoA's because we felt that the chloroacetylated product would be more stable to hydrolysis, generating the nonreactive glycolated protein. In order to confirm the identity of the acetylCoA labeled product, we incubated ClAcCoA, L12 and RimL for various times. Mass

blanchar@aecom.yu.edu.

spectrometric analysis of a ten-minute incubation (Figure S2) revealed the presence of a major peak corresponding to the chloroacetylated-L12 (mass: 13,065) and a minor peak corresponding to acetylCoA labeled L12 (mass: 13,796).

The rapid formation of the chloroacetylated-L12, and its much slower reaction with CoA, generated stoichiometrically with chloroacetylated-L12, suggested that exogenously added thiol-containing compounds could outcompete CoASH for the chloroacetylated substrates. To assess the applicability of ClAcCoA as a reagent for the identification of GNAT substrates, we tested whether the Hat1 histone acetyltransferase would use ClAcCoA as a substrate and would retain its substrate specificity. Using purified yeast Hat1 and a mixture of histones H2A, H2B, H3 and H4, chloroacetyltransfer from ClAcCoA was assessed. The chloroacetylation occurs exclusively on H4 (Figure 1), as assessed by capture of the chloroacetylated-H4 by cysteamine-TAMRA, a result fully consistent with the known preferential specificity of Hat1 for H4[10].

In order to generate a capture reagent containing both a fluorophore and an affinity tag, a peptide containing an His_8 sequence with a C-terminal cysteine residue and N-terminal fluorescein label was prepared. ClAcCoA and histidine-tagged RimL were added to crude *Salmonella* extracts and incubated for twenty minutes, followed by the addition of the fluorescein-His_8 capture peptide. After gel filtration to remove unlabeled peptide, the mixture was loaded onto a Ni-NTA column, washed and eluted with 300 mM imidazole. After SDS-PAGE, a single fluorescent band (Figure 2) was observed on the gel corresponding to the mass of peptidylated L12, as determined on the same gel after Coomassie staining. The fluorescent band was excised and subjected to trypsin digestion. The resulting tryptic fragments were analyzed by MS/MS spectra and confirmed the band as L12 with the fluorescein-His_8-GGC peptide covalently bound to its N terminus (Figures S4,5).

These experiments using purified *N*-acetyltransferases and their known cognate substrates validated the proposed chemistry and utility of ClAcCoA as a labeling reagent. However, the identification of the unknown substrates of the thousands of GNAT superfamily members requires that the labeling be specific and sensitive enough to identify substrates in cell extracts. We prepared dialyzed, centrifuged extracts of *Salmonella enterica* grown to early log phase, where the L12/L7 ratio is largest, and added purified RimL and ClAcCoA to the extract. After twenty minutes, the aminoethanethiolated derivative of TAMRA was added and incubated with the extracts. The entire extract was loaded onto an SDS-PAGE gel, and visualized fluorometrically and by staining with Coomassie Blue R-250. As seen in Figure 3, only in the presence of RimL and ClAcCoA is a fluorescent band observed at the molecular weight of the L12 monomer. These data demonstrate that RimL will correctly find its cognate substrate, L12, in a complex mixture and chloroacetylate it.

Chloroacetyl-coenzyme A appears to fulfill all of the requirements for a reagent that will be useful in identifying the substrate or substrates for the functionally unidentified GNAT superfamily members in bacteria. Both prokaryotic and eukaryotic GNAT's use this compound as substrate. The chloroacetylated products are stable enough to be hydrolyzed slowly, yet reactive enough to undergo rapid, and selective, reaction with thiol nucleophiles. Thiol-containing capture reagents can be generated with both fluorescent and affinity tags that will be useful for purifying and identifying chloroacetylated products. The method has been shown to be sensitive (10 ng of fluorescein labeled L12 can be observed in SDS-PAGE gels, Figure S6) and selective enough for the proteome-wide identification of the substrates for bacterial GNAT's. Although reversible acetylation has been most thoroughly studied in eucaryotes, there are a growing number of examples where reversible acetylation plays an important regulatory role in prokaryotes. The *Salmonella* acetyl-CoA synthase has recently been shown to be reversibly acetylated on Lys460 by a protein acetyltransferase with a concomitant loss of catalytic activity[11]. In addition, there is mounting evidence that eucaryotic histone

acetyltransferases may have broader substrate specificity than once imagined, including the catalytic capability to acetylate both histones, transcription factors and non-histone proteins[12]. A reagent, such as the one described in this report, would be extremely valuable for the potential identification of unknown acetyltransferases with known substrates or known histone acetyltransferases for whom the manifold of it's protein substrates are unknown.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

References

1. Davies J, Wright GD. Trends Microbiol 1997;5:234–240. [PubMed: 9211644]

2. Llano-Sotelo B, Azucnea EF, Kotra LP, Mobashery S, Chow CS. Chem. Biol 2002;9:455–463. [PubMed: 11983334]

3. Berger SL, Pina B, Silverman N, Marcus GA, Agapite J, Regier JL, Trienzenberg SJ, Guarente L. Cell 1992;70:251–265. [PubMed: 1638630]

4. Brownell JE, Zhou J, Ranalli T, Kobayashi R, Edmondsom DG, Roth SY, Allis CD. Cell 1996;84:843–851. [PubMed: 8601308]

5. Vetting MW, de Carvalho LP, Yu M, Hegde SS, Magnet S, Roderick SL, Blanchard JS. Arch. Biochem. Biophys 2005;433:212–226. [PubMed: 15581578]

6. Fukuchi J, Kashiwagi K, Takio K, Igarashi K. J. Biol. Chem 1994;269:22581–22585. [PubMed: 8077207]

7. Tanaka S, Matsushita Y, Yoshikawa A, Isono K. Mol Gen Genet 1989;217:289–293. [PubMed: 2671655]

8. Yoshikawa A, Isono S, Sheback A, Isono K. Mol Gen Genet 1987;209:481–488. [PubMed: 2828880]

9. Vetting MW, de Carvalho LPS, Roderick SL, Blanchard JS. J. Biol. Chem 2005;280:22108–22114. [PubMed: 15817456]

10. Parthun MR, Widom J, Gottshling DE. Cell 1996;87:85–94. [PubMed: 8858151]

11. Starai VJ, Escalante-Semerena JC. J. Mol. Biol 2004;340:1005–12. [PubMed: 15236963]

12. Glozak MA, Sengupta N, Zhang X, Seto E. Gene 2005;363:15–23. [PubMed: 16289629]
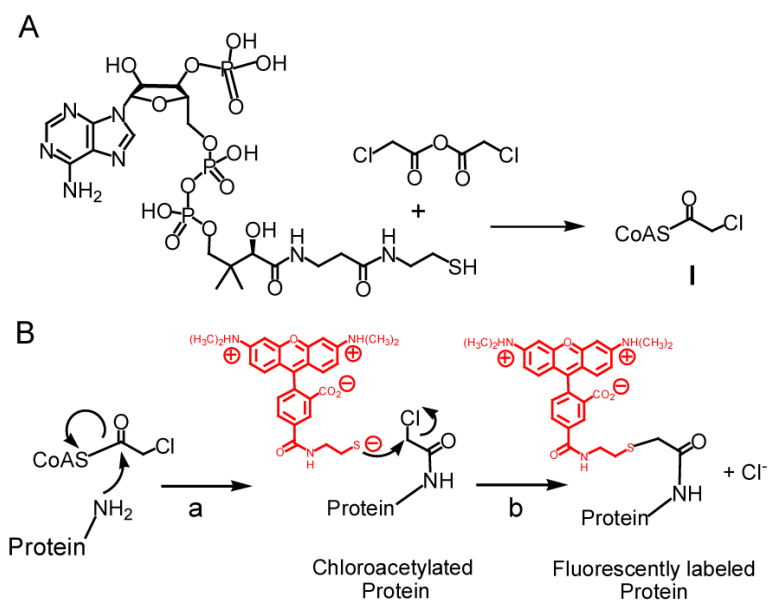
Scheme 1.
A: Synthesis of chloroacetyl-CoA. B: Enzyme-catalyzed chloroacetyltransfer from ClAcCoA to substrate (a), followed by chloride elimination to generate the fluorescently labeled protein (b).
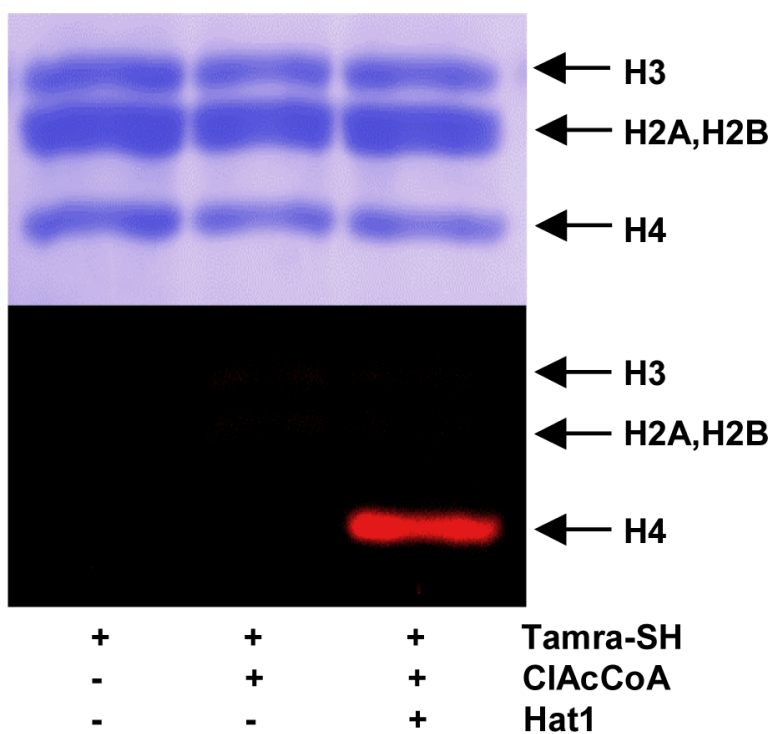
Figure 1.
Hat1-specific chloroacetylation of histone H4. Top: Coomassie stained SDS-PAGE gel of histone mixture. Bottom: Fluorogram of same gel showing specific chloroacetylation of H4 and capture by TAMRA-Cysteamine.
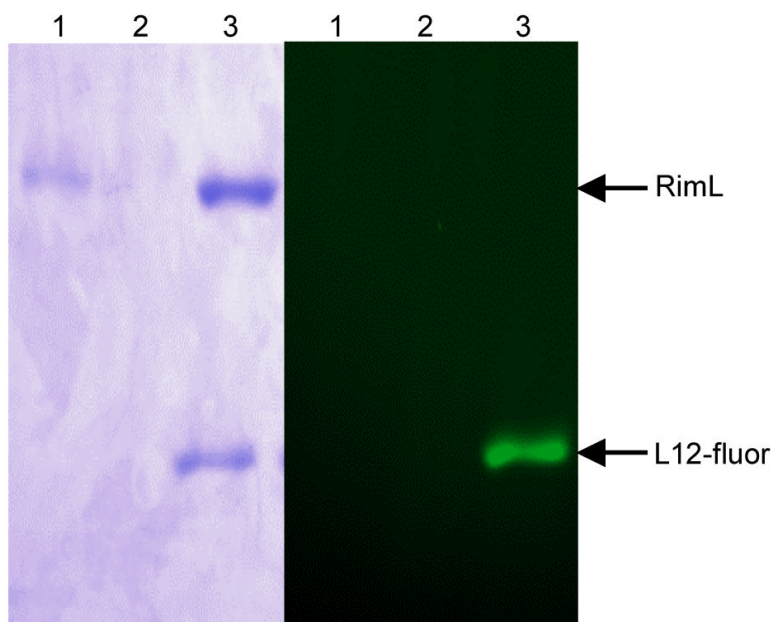
Figure 2.
Chloroacetyl-L12 capture by a fluorescein-labeled His8-GGC peptide, followed by Ni-NTA affinity purification. Lane 1:-ClAcCoA, Lane 2: -Riml, Lane 3: complete. Left: Coomassie staining, Right, Fluorometric analysis.
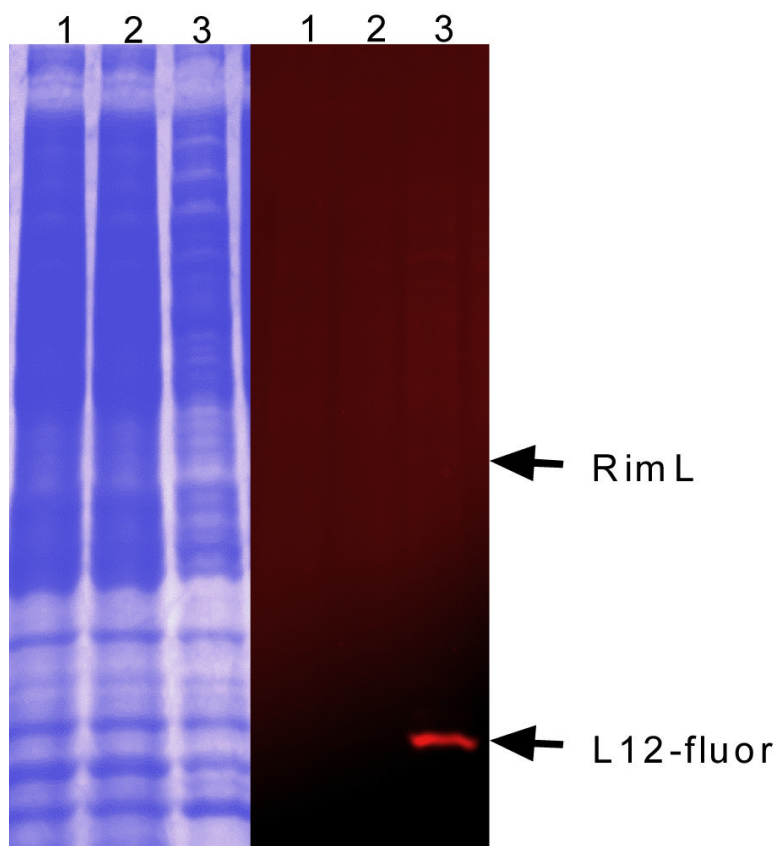
Figure 3.
Chloroacetyl-L12 capture by TAMRA-cysteamine in crude cell extracts. Lane 1:-ClAcCoA, Lane 2: -Riml, Lane 3: complete. Left: Coomassie Staining, Right: Fluorometric analysis.