JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

The Glycosyltransferase Involved in Thurandacin Biosynthesis Catalyzes Both O- and S-Glycosylation

Huan Wang, Trent J. Oman, Ran Zhang, Chantal V. Garcia De Gonzalo, Qi Zhang, and Wilfred A. van der Donk*

Howard Hughes Medical Institute and Roger Adams Laboratory, Department of Chemistry, University of Illinois at Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, United States

Supporting Information

ABSTRACT: The S-glycosyltransferase SunS is a recently discovered enzyme that selectively catalyzes the conjugation of carbohydrates to the cysteine thiol of proteins. This study reports the discovery of a second S-glycosyltransferase, ThuS, and shows that ThuS catalyzes both S-glycosylation of the thiol of cysteine and O-glycosylation of the hydroxyl group of serine in peptide substrates. ThuS-catalyzed S-glycosylation is more efficient than O-glycosylation, and the enzyme demonstrates high tolerance with respect to both nucleotide sugars and peptide substrates. The biosynthesis of the putative products of the *thuS* gene cluster was reconstituted *in vitro*, and the resulting S-glycosylated peptides thurandacin A and B exhibit highly selective antimicrobial activity toward *Bacillus thuringiensis*.

Glycosylation represents one of the most common post-translational modifications of proteins.[1−6] Naturally occurring glycopeptides most commonly involve an O- or N-glycosidic linkage between the saccharide moiety and the side chain of an appropriate amino acid residue. In contrast, examples of naturally occurring S-linked glycopeptides or glycoproteins are very rare.[7−11] To date, only two bacterial S-linked glycopeptides have been reported, which both display antimicrobial activities, sublancin 168 produced by *Bacillus subtilis* 168 and glycocin F produced by *Lactobacillus plantarum* KW30.[12,13] Sublancin contains a glucose moiety linked to Cys, which is installed by the S-glycosyltransferase SunS.[12] Glycocin F contains two sugars, an *N*-acetylglucosamine *β*-O-linked to Ser18 and an *N*-acetylglucosamine S-linked to the C-terminal Cys43 (Figure S1). The glycocin F biosynthetic gene cluster contains one gene encoding a putative glycosyltransferase GccA suggesting that perhaps the enzyme installs both sugars (Figure 1A), but its function has not been confirmed.[13] In this study, we identified the glycosyltransferase ThuS from *Bacillus thuringiensis serovar andalousiensis* BGSC 4AW1 and reconstituted its enzymatic activity *in vitro*. We show that ThuS glycosylates its peptide substrate ThuA at both Ser19 and Cys28, and therefore ThuS represents the first glycosyltransferase that catalyzes both O- and S-glycosylation of proteins *in vitro*. ThuS also displays high promiscuity toward peptide substrates and nucleotide sugars. The biosynthesis of thurandacin A and B, the putative products of the biosynthetic gene cluster that includes *thuS*, was reconstituted *in vitro*, and the resulting glycosylated products were shown to be highly selective in their antimicrobial activity, targeting only *B. thuringiensis*.

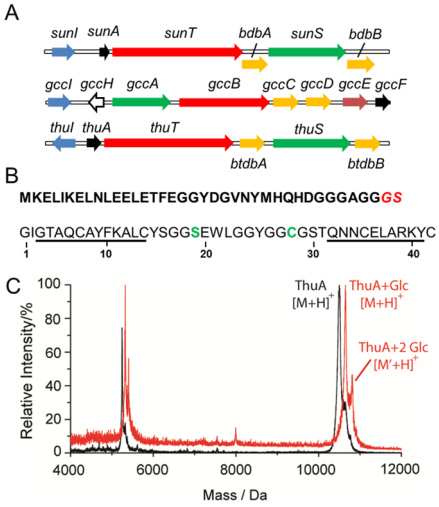

Figure 1. (A) Biosynthetic gene clusters of sublancin and glycocin F and the putative biosynthetic gene cluster that includes *thuS* in *B. thuringiensis* BGSC 4AW1. (B) Sequence of the ThuA precursor peptide. Leader sequence is in bold, and the double-glycine-type proteolytic cleavage site is in red. The Ser and Cys residues glycosylated by ThuS are in green. The predicted helical segments are underlined. (C) MALDI-TOF MS spectra of His_6-ThuA (10 μM) before (black) and after (red) incubation with His_6-ThuS (2 μM), UDP-Glc (1 mM), and Mg^{2+} (1 mM) in 50 mM Tris buffer, pH 7.5 for 10 h at room temperature.

As part of our efforts to mine genomes for new antibiotics, we detected a gene cluster in the genome of *B. thuringiensis* BGSC 4AW1 that showed similarity to the biosynthetic gene clusters of sublancin and glycocin F (Figure 1A). This gene cluster contains genes for a putative precursor peptide ThuA, a glycosyltransferase ThuS, an ABC-transporter ThuT, two thiol-disulfide oxidoreductaseses BtdbA and BtdbB, and a putative immunity protein ThuI. Bioinformatic analysis shows that ThuS shares 39% sequence identity with SunS and belongs to the glycosyltransferase family A. ThuA consists of a 38-residue leader sequence and a 42-residue core peptide, separated by a Gly-Ser motif, which is a double-glycine type[14] proteolytic cleavage site (Figure 1B). Similar to the sublancin precursor

Received: November 4, 2013
Published: December 10, 2013

   dx.doi.org/10.1021/ja411159k | *J. Am. Chem. Soc.* 2014, 136, 84−87

peptide SunA, ThuA contains five Cys residues in its core peptide. Secondary structure prediction tools (PSIPRED)[15] suggest that the peptide contains two α-helical segments spanning residues 3−14 and 32−41 (Figure 1B). The prediction that four of the five Cys residues reside in helical structures is consistent with the NMR structure of glycocin F.[16]

To investigate the function of ThuS, the *thuS* and *thuA* genes were cloned and expressed in *Escherichia coli* as N-terminal fusion proteins with a hexahistidine tag (His_6-ThuS and His_6-ThuA). Upon purification by immobilized metal affinity chromatography, His_6-ThuS was incubated with the purified precursor peptide His_6-ThuA. Addition of uridine diphosphate α-D-glucose (UDP-Glc) and Mg^{2+} resulted in conversion of ThuA to two products with mass increases of 162 and 324 Da, as determined by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Figure 1C), suggesting mono- and bisglucosylation of ThuA. Tandem MS analysis revealed that both Ser19 and Cys28 were glucosylated in the bisglucosylated ThuA (Figure S2), but no glycosylation was observed for Cys7, Cys14, Cys35, or Cys42. The lack of glycosylation of these cysteine residues that are located in the likely helical regions is similar to the site-selectivity that is observed with SunS.[12] The glycosylation of Ser19, however, was surprising since SunS displays high chemo- and regioselectivity toward Cys22 of its peptide substrate SunA and does not modify Ser22 in the SunA-C22S mutant.[12] To the best of our knowledge, ThuS is the first glycosyltransferase that catalyzes both S-glycosylation of cysteine and O-glycosylation of the hydroxyl group of serine in proteins *in vitro*. This finding also provides support that GccA may install both *N*-acetylglucosamine groups in glycocin F.

We next investigated the efficiency of the O- and S-glycosylation reactions catalyzed by ThuS. Incubation of equimolar amounts of ThuA and UDP-Glc (20 μM) in the presence of ThuS and Mg^{2+} resulted in glycosylation of only Cys28, as determined by MALDI-TOF MS and tandem MS (Figure S3), suggesting that S-glycosylation is favored. Indeed, installation of a second glucose on Ser19 of the monoglucosylated peptide (hereafter termed ThuA-Cys28-Glc) was significantly slower than the glycosylation of Cys28 in unmodified ThuA, as indicated by a coupled enzymatic assay (Supporting Information and Table S2).[17] In this assay, the nucleotide diphosphate generated upon glycosyl transfer reactions was coupled to NADH oxidation via pyruvate kinase (PK) and lactate dehydrogenase (LDH) (Scheme S1). Under appropriate conditions such that reactions catalyzed by PK and LDH are much faster than nucleotide diphosphate release catalyzed by ThuS, the observed rate of NADH consumption equals the rate of nucleotide diphosphate generation and can be continuously monitored at 340 nm.

To further investigate the selectivity of ThuS, Cys28 in the ThuA peptide was mutated to Ser to investigate whether the preferential glucosylation of Cys28 over Ser19 is the result of site-selectivity or chemoselectivity. In the presence of 0.5 μM ThuS, 100 μM UDP-Glc and 30 μM peptide substrate (maximum peptide solubility), the modification rate of ThuA-C28S was about 25% of the modification rate of ThuA (Table S3), suggesting that Cys glucosylation is faster than Ser glucosylation. This increased activity with thiols was also observed when another ThuA mutant, ThuA-S19C, was incubated with ThuS and UDP-Glc. The first glucosylation occurred with similar rates for ThuA and ThuA-S19C, as determined by the coupled enzymatic assay (Table S3). However, extended incubation time lead to more bisglucosylated product with ThuA-S19C compared to the assay with wild type ThuA (Figures 1C, S4a). Interestingly, glucose was present on either Cys19 or Cys28 in the monoglucosylated ThuA-S19C, as determined by tandem MS (Figure S4b). Although the MSMS data could not quantify the selectivity of ThuS for Cys19 or Cys28 in this mutant, the intensity of the fragment ions indicated that the enzyme did not display high preference for either. Collectively, these results show that ThuS is chemoselective for the thiol of cysteines over the hydroxyl groups in serines but displays relatively relaxed site-selectivity.

To investigate the substrate tolerance of ThuS toward sugar donors, the enzyme was incubated with His_6-ThuA and Mg^{2+} in the presence of UDP-α-D-*N*-acetylglucosamine (UDP-GlcNAc), UDP-α-D-galactose (UDP-Gal), and guanosine diphosphate α-D-mannose (GDP-Man). All three nucleotide sugars proved to be good substrates (K_m values <10 μM), but the glycosyl transfer reaction with UDP-Glc was at least 5-fold more efficient than reactions with other sugar donors, as indicated by MALDI-TOF-MS analysis and the enzymatic coupled assay (Figures S5, S6, Table S4). Hence, ThuS likely installs a glucose onto its substrate peptide under physiological conditions.

We next attempted to reconstitute the biosynthesis of the product(s) encoded by the gene cluster *in vitro*. First, a factor Xa cleavage site was engineered into the ThuA substrate at the junction of the putative leader and core peptide. The resulting His_6-ThuA-Xa peptide was expressed, purified from *E. coli*, and glucosylated by ThuS, yielding a mixture of mono- and bisglucosylated peptides. These peptides were then treated with a mixture of oxidized and reduced glutathione and EDTA to form disulfide cross-links.[18] Finally, the leader peptide was removed by the addition of the protease factor Xa, and each of the resulting products was purified by RP-HPLC (Figures 2A, S7, S8). Mass spectrometric analysis showed that the mass of both mono- and bisglucosylated products were 4 Da lower than the mass of the corresponding glucosylated ThuA-Xa core peptides. Furthermore, incubation with triscarboxyethyl phosphine increased the mass of both peptide products by 4 Da, indicating the successful formation of two disulfide bonds in both samples. Trypsin digestion of the purified bisglucosylated product yielded a peptide fragment spanning residues 12−39 with one disulfide bond as well as a peptide fragment containing residues 1−11 and residues 40−41 connected by one disulfide bond, as determined by MALDI-TOF MS (Figure S9a), indicating that one disulfide cross-link was between Cys14 and Cys35 and a second disulfide was formed between Cys7 and Cys42. Similarly, trypsin digestion of the monoglucosylated product also resulted in a peptide fragment spanning residues 12−39 with one disulfide bond, indicating the formation of the same disulfide pattern (Figure S9b).

To examine the stereochemistry of the glycosidic linkages, bisglucosylated ThuA-C28S peptide, which contains a glucose moiety on both Ser19 and Ser28, was treated with β-glucosidase. Subsequent MALDI-TOF MS analysis revealed that two glucoses were released by β-glucosidase, indicating that both glucose moieties on ThuA-C28S were β-linked (Figure S10). Therefore, ThuS is an inverting glycosyltransferase. Collectively, these results show that the *in vitro* generated glycopeptides have β-linked glucose moieties and a nested disulfide pattern similar to the S-glycopeptide antibiotics sublancin 168 and glycocin F (Figures 2A, S1).[12,13,16]

The glucosylated and oxidatively folded peptides were then tested for antimicrobial activity against a number of Gram-

dx.doi.org/10.1021/ja411159k | J. Am. Chem. Soc. 2014, 136, 84−87

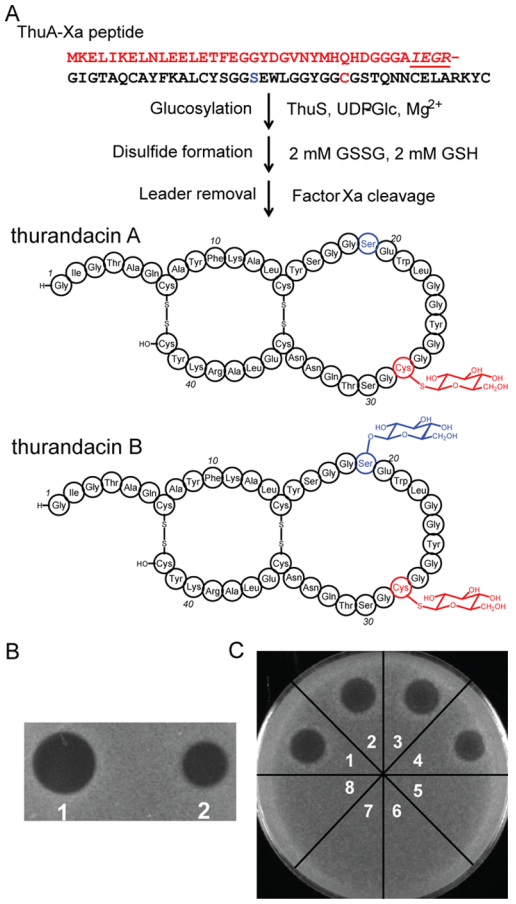

Figure 2. (A) *In vitro* reconstitution of thurandacin biosynthesis. GSSG: oxidized glutathione; GSH: reduced glutathione. (B) Agar diffusion assay of thurandacin A and B against *B. thuringiensis* BGSC 4CC1. Samples were spotted on LB agar in a volume of 5 μL: (1) 3 μM thurandacin A; (2) 3 μM thurandacin B. (C) Agar diffusion assay of thurandacin A and its analogs against *B. thuringiensis* BGSC 4CC1. The following samples were spotted on LB agar in a volume of 5 μL: (1) 3 μM thurandacin A; (2) 3 μM thurandacin A-Man analog; (3) 3 μM thurandacin A-GlcNAc analog; (4) 3 μM thurandacin A-Gal analog; (5) 3 μM oxidatively folded ThuA-Xa-Cys28-Glc peptide; (6) 3 μM ThuA-Xa-Cys28-Glc; (7) 3 μM ThuA-Xa core peptide; (8) 10 mM Tris buffer, pH 7.5.

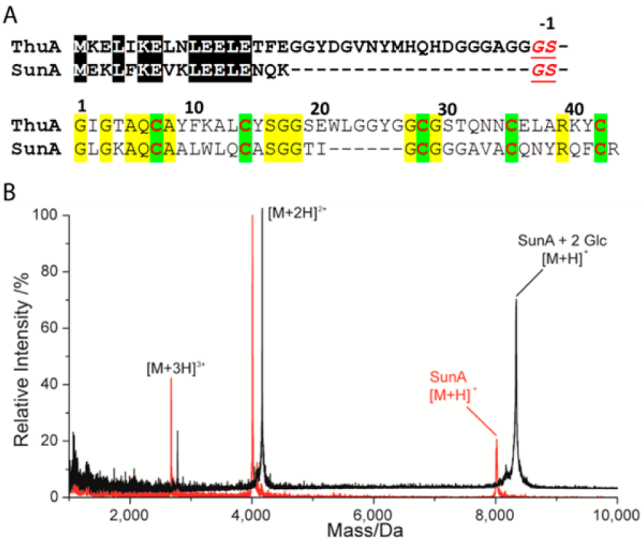

Figure 3. (A) Sequence alignment of ThuA and SunA peptides. Leader sequences are in bold, and putative double glycine type proteolytic cleavage sites are in red. Numbering shown is for the ThuA peptide. (B) MALDI-TOF MS spectra of His_6-SunA before (black) and after (red) incubation with His_6-ThuS, UDP-Glc, and Mg^{2+}.

positive bacteria including *Lactococcus lactis* HP, *Bacillus subtilis* ATCC 6633, *Bacillus halodurans* C125, and *B. thuringiensis serovar pulsiensis* BGSC 4CC1. The antimicrobial activities of *in vitro* generated peptides were determined by agar diffusion assays, and their potency was estimated from the diameter of the inhibition zone, assuming that the number of sugar modifications do not significantly alter their diffusion behavior in agar.

Both mono- and bis-glycosylated peptides exhibited potent inhibitory activity against *B. thuringiensis* BGSC 4CC1 (Figure 2B), however, very low or no inhibitory activity toward other strains (Table S6, Figure S11a). Hence, we have named the *in vitro* generated mono- and bis-glycosylated peptides from the gene cluster of *B. thuringiensis serovar andalousiensis* 4AW1 thurandacin A and B, respectively (Figure 2A). The minimum inhibitory concentration (MIC) of thurandacin A against *B. thuringiensis* BGSC 4CC1 was determined to be 0.6 μM in liquid LB medium (Figure S11b); the quantities of thurandacin B were insufficient for MIC determination in liquid culture. Thurandacin A displayed somewhat higher potency than thurandacin B in agar diffusion growth inhibition assays (Figures 2B, S12).

Glucosylation at Cys28 and removal of the leader peptide were strictly required for bioactivity of thurandacin A (Figure 2C). Thurandacin A analogs with different sugar moieties were also prepared by using GDP-Man, UDP-Gal, and UDP-GlcNAc in the ThuS catalyzed glycosylation following the *in vitro* reconstitution procedure described above. Interestingly, all thurandacin analogs displayed similar growth inhibitory activity, as indicated by agar diffusion assay (Figure 2C). Thus, the stereochemistry of the hexose does not appear to be critical for bioactivity.

Agar diffusion assays suggest that thurandacin A is more potent than sublancin 168 against *B. thuringiensis* BGSC 4CC1 (Figure S13). Sublancin from *B. subtilis* 168 displayed antimicrobial activities mainly against *B. subtilis*,[19] and glycocin F produced by *L. plantarum* KW30 has inhibitory activity limited to other lactobacilli.[20] Therefore, S-glycopeptide antibiotics discovered so far, while similar in structure (Figure S1),[16] all display narrow antimicrobial spectrum against bacteria that are closely related to their producing organisms.

We next explored the substrate specificity of ThuS with respect to peptide substrates. ThuA shares high homology with SunA at the N-terminus of the leader peptide but low sequence similarity in the core peptide (Figure 3A). Surprisingly, ThuS was able to modify the SunA peptide in the presence of UDP-Glc and Mg^{2+}, efficiently generating its bisglucosylated product (Figure 3B). In contrast, SunS was not able to modify the ThuA peptide, demonstrating that the peptide sequence selectivity of ThuS is more relaxed. Tandem mass spectrometry analysis revealed that Cys22 of bisglycosylated SunA was modified by ThuS, with the second glucosylation occurring on Ser16, Thr19 or Cys29 in a mutually exclusive manner (Figure S14). ThuS was also able to glucosylate the SunA core peptide without the leader peptide attached, however the modification was less efficient than that of full length SunA and only monoglycosylation was observed by MALDI-TOF MS (Figure S15). These results suggest that the leader peptide of SunA contributes to substrate recognition by ThuS but is not essential.

dx.doi.org/10.1021/ja411159k | J. Am. Chem. Soc. 2014, 136, 84−87

Table 1. Partial Sequences of ThuA and SunA Mutant Peptides (Residues 11−34)[a]

Substrate	Partial sequence (residues 11-34)
ThuA-CREB	KALCYSGGSE**CATLAQVCMP**DQNN
SunA-CREB	WLK**CATLAQVCM**DGGAVACQNYRQ
SunA-Syn	WLQCA**EQANAVCEAVVS**ACQNYRQ

[a]Sequences of CREB and Syn replacements of the native loop sequences in SunA and ThuA are in bold and underlined. Target Cys residues are highlighted in red.

Since ThuS was able to glucosylate SunA despite the significant differences in the sequences of SunA and ThuA, we decided to investigated whether ThuS could modify ThuA and SunA peptides with non-natural sequences inserted in the loop between the internal disulfides. S-glycosidic linkages have been shown to have significantly improved chemical and biological stability compared to O-glycosidic linkages,[21−24] and hence *S*-linked analogs of O-linked glycopeptides could be useful tools, for instance, for antibody generation. To investigate the potential use of ThuS, ThuA-CREB and SunA-CREB peptides were created that contain short sequences of the transcription factor cyclic AMP-response element binding protein (CREB), which bears an O-GlcNAc modification site on a Ser residue within the sequence (Table 1).[25−27] We envisioned generating their S-linked analogs with ThuS, and hence we mutated the Ser to Cys in these insertions (indicated in red in Table 1). In the presence of UDP-Glc and Mg^{2+}, ThuS indeed modified the ThuA-CREB and SunA-CREB peptides at the desired Cys, as determined by MALDI-TOF MS and tandem MS (Figures S16, S17). Similarly, a SunA-Syn peptide was generated that contains a short sequence of γ-synuclein that includes a Ser that is an O-GlcNAc modification site.[28] Mutation of this Ser to Cys and treatment with ThuS again resulted in the glucosylation of the inserted Cys (Table 1 and Figure S18). Given the very low sequence similarity between the loops in ThuA and SunA and the sequences of the inserted CREB and γ-synuclein peptides, ThuS displays extraordinary tolerance toward different peptide substrates. At present, attempts to glycosylate the CREB and γ-synuclein peptides with UDP-GlcNAc to generate the S-GlcNAc analogs of the natural O-GlcNAc modification have been unsuccessful. Structural information on ThuS may allow identification of potential mutations that would allow the use of UDP-GlcNAc.

In summary, we have discovered and reconstituted a glycosyltransferase ThuS, which *in vitro* installs β-linked sugars on both cysteine and serine, suggesting that GccA may install both GlcNAc sugars in glycocin F. ThuS catalyzes S-glycosylation more efficiently than O-glycosylation and displays high tolerance toward peptide substrates and nucleotide sugars, rendering it a useful synthetic tool. The biosynthesis of the putative products thurandacin A and B from the *thuS* biosynthetic gene cluster has been reconstituted *in vitro*, and both products contain a β-S-linked glucose and have a sublancin-like structure with two disulfide cross-links. Thurandacins exhibit highly selective antimicrobial activity toward *B. thuringiensis*.

■ ASSOCIATED CONTENT

Supporting Information

Description of procedures, purifications, and figures. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

vddonk@illinois.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the U.S. National Institutes of Health (GM58822 to W.A.v.d.D.) and fellowships from the NIGMS-NIH Chemistry-Biology Interface Training Grant Program (5T32-GM070421 to C.G.D.G) and the Cellular and Molecular Biology Training Program (T32 GM007283 to T.J.O.). Mass spectra were recorded in part on an instrument purchased with grant S10RR027109-01 from the U.S. National Institutes of Health.

■ REFERENCES

(1) Weerapana, E.; Imperiali, B. *Glycobiology* **2006**, *16*, 91R.
(2) Nothaft, H.; Szymanski, C. M. *Nat. Rev. Microbiol.* **2010**, *8*, 765.
(3) Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O. *Annu. Rev. Biochem.* **2011**, *80*, 825.
(4) Nothaft, H.; Szymanski, C. M. *J. Biol. Chem.* **2013**, *288*, 6912.
(5) Helenius, A. Q.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019.
(6) Peter-Katalinic, J. *Methods Enzymol.* **2005**, *405*, 139.
(7) Weiss, J. B.; Lote, C. J.; Bobinski, H. *Nat. New Biol.* **1971**, *234*, 25.
(8) Lote, C. J.; Weiss, J. B. *FEBS Lett.* **1971**, *16*, 81.
(9) Lote, C. J.; Weiss, J. B. *Biochem. J.* **1971**, *123*, 25P.
(10) Olsen, E. H.; Rahbek-Nielsen, H.; Thogersen, I. B.; Roepstorff, P.; Enghild, J. J. *Biochemistry* **1998**, *37*, 408.
(11) Elsayed, S.; Bennich, H. *Scand. J. Immunol.* **1975**, *4*, 203.
(12) Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. *Nat. Chem. Biol.* **2011**, *7*, 78.
(13) Stepper, J.; Shastri, S.; Loo, T. S.; Preston, J. C.; Novak, P.; Man, P.; Moore, C. H.; Havlicek, V.; Patchett, M. L.; Norris, G. E. *FEBS Lett.* **2011**, *585*, 645.
(14) Dirix, G.; Monsieurs, P.; Marchal, K.; Vanderleyden, J.; Michiels, J. *Microbiology* **2004**, *150*, 1121.
(15) McGuffin, L. J.; Bryson, K.; Jones, D. T. *Bioinformatics* **2000**, *16*, 404.
(16) Venugopal, H.; Edwards, P. J.; Schwalbe, M.; Claridge, J. K.; Libich, D. S.; Stepper, J.; Loo, T.; Patchett, M. L.; Norris, G. E.; Pascal, S. M. *Biochemistry* **2011**, *50*, 2748.
(17) Gosselin, S.; Alhussaini, M.; Streiff, M. B.; Takabayashi, K.; Palcic, M. M. *Anal. Biochem.* **1994**, *220*, 92.
(18) Fuller, E.; Green, B. R.; Catlin, P.; Buczek, O.; Nielsen, J. S.; Olivera, B. M.; Bulaj, G. *FEBS J.* **2005**, *272*, 1727.
(19) Paik, S. H.; Chakicherla, A.; Hansen, J. N. *J. Biol. Chem.* **1998**, *273*, 23134.
(20) Kelly, W. J.; A., R. V.; Huang, C. M. *J. Appl. Bacteriol.* **1996**, *81*, 657.
(21) Baran, E.; Drabarek, S. *Pol. J. Chem.* **1978**, *52*, 941.
(22) Jahn, M.; Chen, H.; Mullegger, J.; Marles, J.; Warren, R. A.; Withers, S. G. *Chem. Commun.* **2004**, 274.
(23) Jahn, M.; Marles, J.; Warren, R. A.; Withers, S. G. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 352.
(24) Driguez, H. *ChemBioChem* **2001**, *2*, 311.
(25) Carlezon, W. A., Jr.; Thome, J.; Olson, V. G.; Lane-Ladd, S. B.; Brodkin, E. S.; Hiroi, N.; Duman, R. S.; Neve, R. L.; Nestler, E. J. *Science* **1998**, *282*, 2272.
(26) Lonze, B. E.; Ginty, D. D. *Neuron* **2002**, *35*, 605.
(27) Rexach, J. E.; Clark, P. M.; Mason, D. E.; Neve, R. L.; Peters, E. C.; Hsieh-Wilson, L. C. *Nat. Chem. Biol.* **2012**, *8*, 253.
(28) George, J. M. *Genome Biol.* **2002**, *3*, REVIEWS3002.