REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
07-07-2012	Final	March 1, 2009 to February 28, 2012

4. TITLE AND SUBTITLE
Development of Novel Antibiotics for the Treatment of Acinetobacter and Related Pathogens

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-09-1-0240

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Hergenrother, Paul J.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Board of Trustees of the University of Illinois
506 S. Wright Street
Urbana, IL 61801-3620

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; distribution is Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This report describes the progress made toward the project objectives. Progress includes the identification of a novel antibacterial agent, and the development of a new paradigm for the synthesis of complex and diverse small molecules, one that can now be applied to the creation and discovery of novel antibacterial agents.

15. SUBJECT TERMS
antibiotics, compound screening, complex small molecules

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Paul J. Hergenrother
U	U	U			19b. TELEPONE NUMBER (*Include area code*) 217-333-0363

Professor Paul J. Hergenrother
hergenro@uiuc.edu
University of Illinois at Urbana-Champaign
Department of Chemistry
600 S. Mathews
Urbana, IL 61801
Project Title: "Development of Novel Antibiotics for the Treatment of *Acinetobacter* and Related Pathogens"
ONR Award #: N000140910240
Grant Period: 3/1/2009-2/28/2012
FINAL REPORT

Project objectives

Our objectives were to identify novel antibacterial agents and strategies for the treatment of problematic bacterial pathogens. Specifically, we were interested in the presence of the genes encoding toxin-antitoxin systems in *Acinetobacter* clinical isolates, and in the discovery of novel antibacterial agents through high-throughput compound generation and screening.

Brief summary of results

The project was successful on all fronts. We determined that the genes for toxin-antitoxin systems were not present in *Acinetobacter* clinical isolates, thus toxin activation strategies will not be useful against this pathogen. We discovered a potent and novel antibiotic, called ABTZ-1, and synthesized a family of related compounds and evaluated them for their spectrum of activity. A major manuscript describing this work was published as a cover article in *ChemBioChem.* Finally, we devised a novel paradigm for the creation of complex and diverse compounds, agents that are more natural product like in their structure and thus have the propensity to kill bacterial cells. A major manuscript describing this work is under review at *Nature*, and a Gram-negative targeted version of this project is currently funded under one year ONR award N000141210616.

Full summary of results – ABTZ-1

We screened through a collection of ~155,000 compounds (through the UIUC High-Throughput Screening Facility), and, through a series of tiered assays of increasing complexity, whittled these compounds down to one compound, called ABTZ-1. Below is described the antibacterial spectrum, killing kinetics, synthesis and structure-activity relationship, its effect on membranes, analysis of resistance mutants, and its effect on macromolecular synthesis.

Antibacterial spectrum. We evaluated ABTZ-1 versus a variety of Gram-positive and Gram-negative organisms. As shown in Figure 1, this compound has considerable activity versus Gram-positive

organisms, including clinical isolates of *Staphylococcus aureus.* It is not active versus Gram-negative organisms, likely due to lack of cell penetration; the highly penetrable *E. coli* mutant MC1061 is effectively killed by ABTZ-1.

Organism	ABTZ-1 MIC [μg mL^{-1}][a]	Organism	ABTZ-1 MIC [μg mL^{-1}]
Staphylococcus aureus		*Enterococcus faecalis* JH2-2	32
ATCC 12608	2		
ATCC 12608 + 10% serum	2	*Bacillus subtilis*	
ATCC 12608 + 50% serum	8	ATCC 6633	16
ATCC 25923	4		
ATCC 13709	8	*Bacillus cereus*	8
ATCC 29213	4	ATCC 11778	
ATCC 12598	2		
		Listeria monocytogenes	
Staphylococcus aureus		ATCC 19115	8
Methicillin-resistant			
NRS382 (USA100)	4	*Streptococcus pneumonia*	
NRS383 (USA200)	16	ATCC BAA-255	8
NRS384 (USA300)	4		
NRS22 (USA600)	4	*Escherichia coli*	
NRS3	4	DH5α	32
NRS4	4	MC1061	2
NRS21	2		
NRS74	8	*Acinetobacter baumannii*	
		ATCC 19606	> 128
Enterococcus faecium			
Vancomycin-resistant		*Salmonella cholerasuis*	
U503	16	ATCC 14028	> 128
B21190	16		
		Pseudomonas aeruginosa (CI01)	> 128

[a] Minimum Inhibitory concentration (MIC) for ABTZ-1 was determined against liquid culture of each bacterial strain as recommended by the CLSI guidelines.[16]

Figure 1. Antibacterial activity of ABTZ-1 versus Gram-positive and Gram-negative organisms.

Killing kinetics. To determine the rate of cell death induced by ABTZ-1, its killing kinetics on *S. aureus* were investigated. These results are shown in Figure 2.

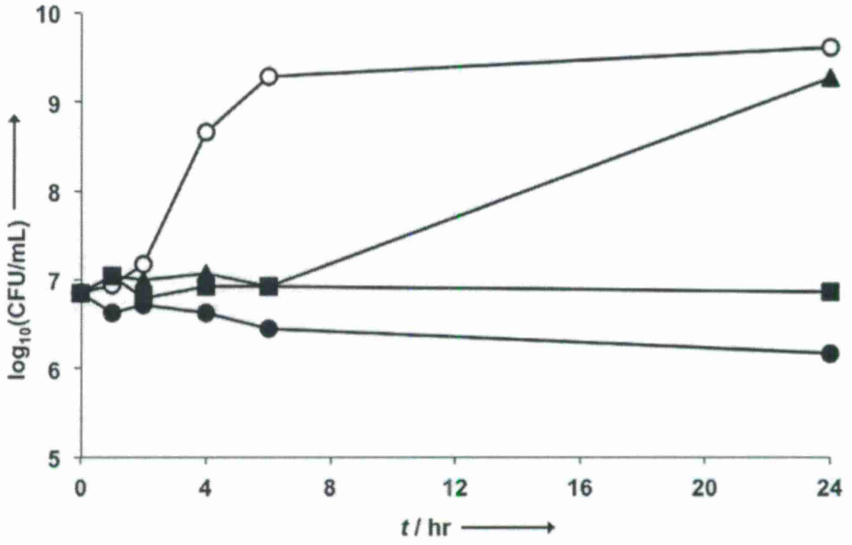


Figure 2. ABTZ-1 killing kinetics. The killing kinetics of **ABTZ-1** against *S. aureus* (ATCC 12608) were evaluated to determine if this compound exerts its antibacterial activity in a bacteriostatic or bactericidal manner. *S. aureus* (ATCC 12608) culture containing 1×10^7 cfu/mL was treated with DMSO (open circle) or **ABTZ-1** at concentrations equal to 1 (triangle), 2 (square), or 4 (closed circle) times its MIC (2, 4, and 8 μg/mL, respectively). Cultures were incubated at 37°C, and at defined time-points, 100 μL aliquots were removed from each treated culture, diluted in sterile saline, and plated on nonselective media to determine the number of viable bacterial cells in the culture. After a 24-hour incubation with *S. aureus* at 4 times its MIC, **ABTZ-1** induced a 1.3-log decrease in the number of viable cells in treated culture. As bactericidal compounds are indicated by CFU/mL reductions of 3-log or greater, **ABTZ-1** was determined to be bacteriostatic against *S. aureus*.

Synthesis and structure-activity relationship. We made a large effort to understand the structure-activity relationship of ABTZ-1. Shown in Figure 3A is the basic synthetic route to this family of compounds, and then is shown the various compounds synthesized in this program and their antibacterial activity.

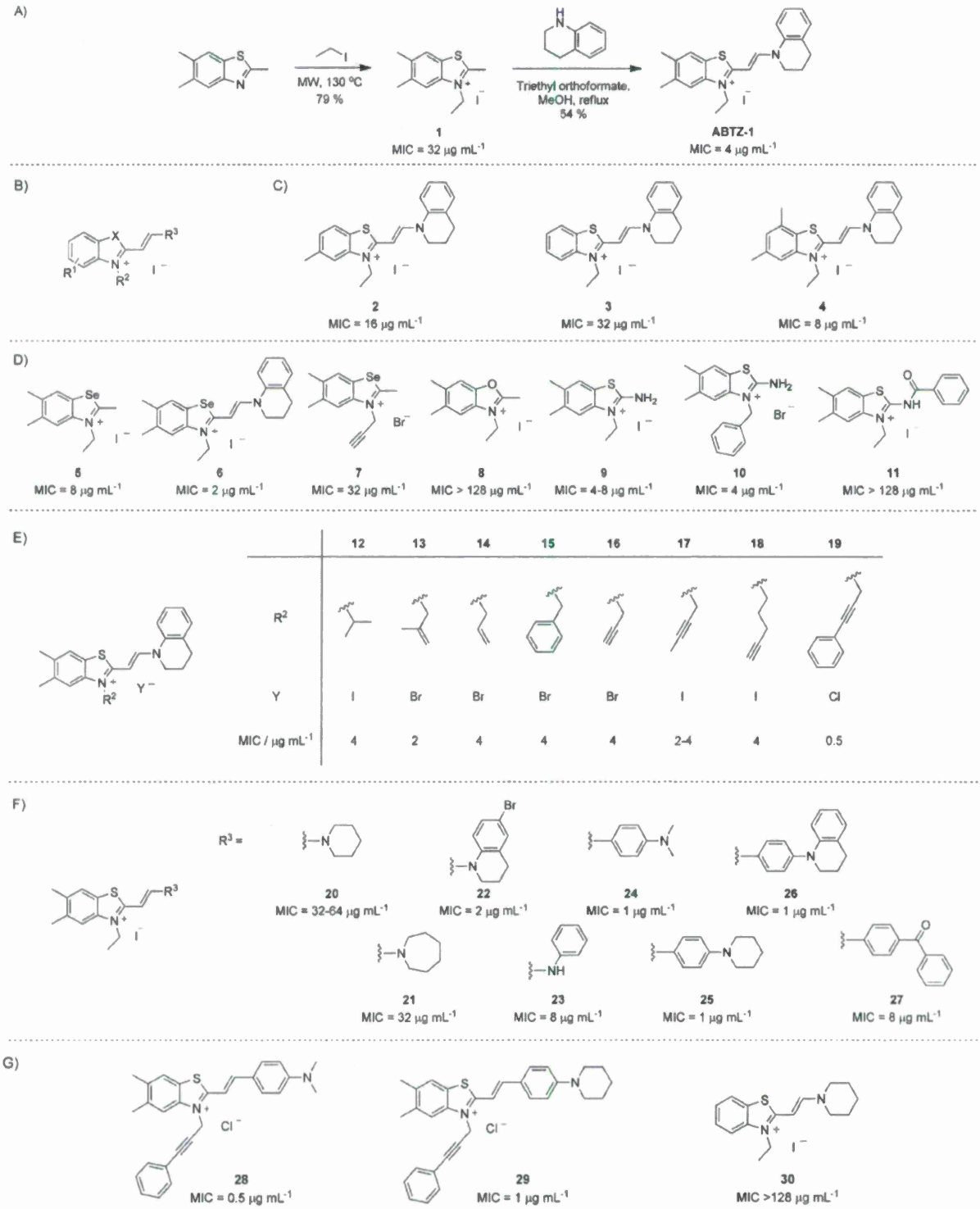


Figure 3. Synthesis and derivatization of **ABTZ-1**. **ABTZ-1** and **ABTZ-1** derivatives were synthesized and assessed for antibacterial activity against *S. aureus* (ATCC 25923). Determined MIC values for each compound against this strain are indicated. A) Synthetic route for **ABTZ-1**. Derivatives were synthesized in a similar manner. See supplementary information for details. B) Derivatives were synthesized that included altered methyl substitution patterns (R_1) on the benzothiazole core, *N*-alkylation with a variety of

saturated and unsaturated hydrocarbons (R_2), replacement of tetrahydroquinoline substitution at the 2-position (R_3), and the replacement of sulfur in the benzothiazole core with selenium and oxygen (X). These derivatives were synthesized as halide salts with either iodide, bromide, or chloride as the counteranion (Y). C) Derivatives modified at R_1. D) Derivatives synthesized using starting materials other than 2,5,6-trimethylbenzothiazole E) Derivatives modified at R_2 with counter-ion Y. F) Derivatives modified at R_3. G) Derivatives designed using structure-activity relationship data gathered from compounds **1-27**. Structural elements known to enhance antibacterial activity were incorporated into compounds **28** and **29**, and those known to attenuate antibacterial activity were incorporated into compound **30**.

Effects on prokaryotic and eukaryotic membranes. The effect of **2121** on the membrane integrity of *S. aureus* and human erythrocytes was determined by measuring the leakage of intracellular potassium and hemolysis, respectively. Treatment of *S. aureus* with **2121** at four times its MIC induced the leakage of only 1.7% of its total intracellular potassium after 1 hour, and only 2.8% hemolysis relative to a positive control was observed when erythrocytes were treated with 128 µg/mL **2121** for 2 hours. Thus, it was determined that **2121** disrupted neither bacterial nor eukaryotic membranes.

Resistant mutant analysis. Resistant mutants of *S. aureus* were spontaneously generated in the presence of 8 µg/mL **2121** at frequencies ranging from 2.7×10^{-9} to 4.7×10^{-9}. In an effort to characterize the mechanism of action for **2121**, resistant strains were treated with a collection of antibiotics with well-defined modes of action. While most antibiotics were equally active against **2121**-resistant and -susceptible strains of *S. aureus*, **2121**-resistant strains were 8- to 16-fold less susceptible to both fusidic acid and quinupristin-dalfopristin. In addition, resistant strains of *E. coli* MC1061 generated at a frequency of 1.3×10^{-6} were found to be cross-resistant to a variety of translational inhibitors including fusidic acid, quinupristin-dalfopristin, tetracycline, erythromycin, and linezolid.

In an effort to determine how fusidic acid and streptogramin resistance were conferred to these **2121**-resistant strains, a PCR screen was performed to identify common resistance determinants for these antibiotics on the genome of **2121**-resistant *S. aureus* (18, 21, 24, 39, 41). As shown in Figure 4, the fusidic acid resistance determinant, *fusB*, and streptogramin resistance determinants from the *erm* and *vat* gene classes were not amplified from the genome of **2121**-resistant *S. aureus*, suggesting that resistance to these antibiotics was not conferred via acquisition of these genes. Additionally, the *fusA* gene encoding the target for fusidic acid and the genes encoding ribosomal proteins L4, L10, L11, L16, and L22 were each amplified by PCR from the genomes of **2121**-resistant and -susceptible *S. aureus*. The sequence for each product amplified from the genome of **2121**-resistant *S. aureus* was identical to that amplified from the genome of the corresponding parent strain; this suggests that fusidic acid- and streptogramin-resistance are also not conferred to **2121**-resistant strains via target mutation.

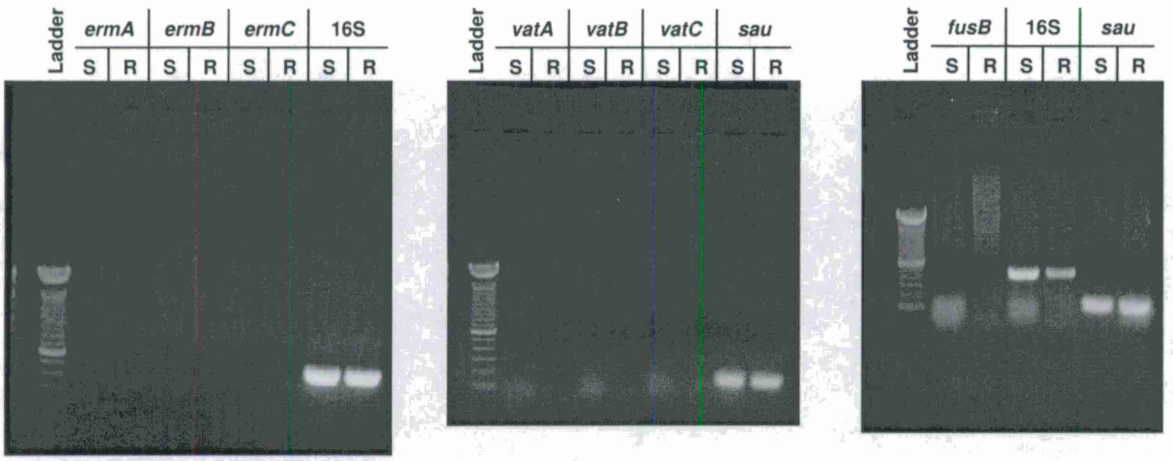


Figure 4. PCR amplification of fusidic acid and streptogramin resistance determinants from the genome of 2121-resistant *S. aureus*. Attempts were made to amplify streptogramin resistance determinants *ermA*, *ermB*, *ermC*, *vatA*, *vatB*, and *vatC* and fusidic acid resistant determinant *fusB* from the genome of 2121-resistant *S. aureus* (ATCC 12608) and that of its parent strain. PCR was performed as previously described (7, 20, 21, 35) using published primer sets specific to these genes and to those encoding the *S. aureus* 16S ribosomal subunit and a *S. aureus*-specific protease (*sau*). Reactions were loaded in 40 µL volumes onto 1% agarose gels containing 0.5 µg/mL ethidium bromide. A 10 µL volume of 1 kb ladder (Promega) was also loaded into the first well of each gel. Gels were run for 45 min at 120V and visualized on a UV light box. Reactions containing genomic DNA from 2121-susceptible *S. aureus* and 2121-resistant *S. aureus* are labeled "S" and "R", respectively. Successful amplification of positive control genes (16S and *sau*) was indicated by the visualization of bands corresponding to 420 and 108 bp products, respectively. No amplification products were detected in PCR reactions containing primers specific for resistance determinants, indicating that these genes are not present on the genome of either strain.

Effects on macromolecular synthesis. Further attempts to characterize the mechanism of action for **2121** included efforts to assess its effects on global biosynthetic pathways by measuring the incorporation of [^{3}H]-labeled precursors into corresponding macromolecules. However, these efforts were complicated by the interference of **2121** in the detection of radiation by liquid scintillation counters. As **2121** is bright yellow in color, this interference was attributed to color quenching, an effect observed when photons produced by the scintillation fluid in response to radiation are absorbed or scattered by highly colored test compounds. The employment of color quenching correction settings on the liquid scintillation counter did not ameliorate the effect; therefore, as derivative **2** was found to lack these color quenching properties, this compound was evaluated in macromolecular synthesis assays. As shown in Figure 5, dose-dependent inhibition of both DNA replication and translation was observed in the presence of compound **2**, while transcription in *S. aureus* was unaffected by this compound.

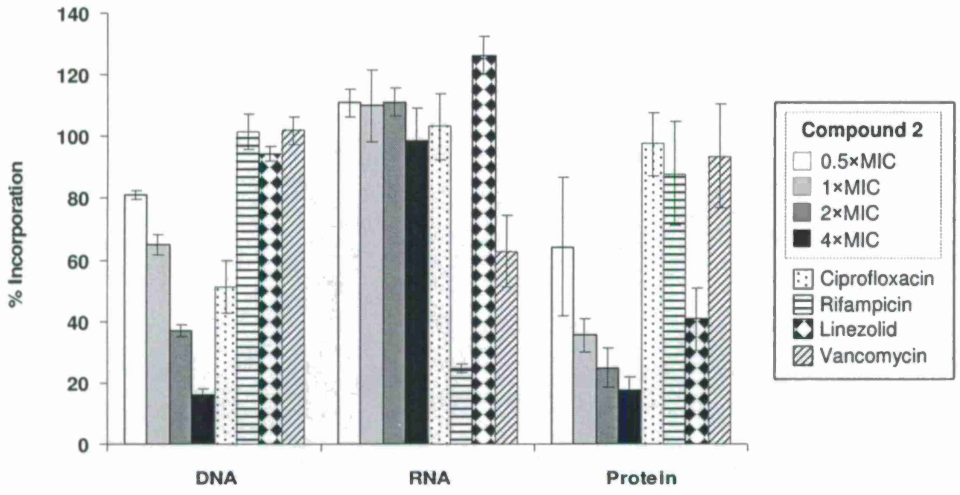


Figure 5. Evaluation of compound 2 in macromolecular synthesis assays. To determine the effect of compound 2 on the incorporation of [^{3}H]-thymidine, [^{3}H]-uridine, and [^{3}H]-leucine into DNA, RNA, and protein, *S. aureus* (ATCC 12608) cultures were co-treated with the appropriate radiolabeled precursor and either DMSO or compound 2 at 8, 16, or 32 µg/mL. Additional cultures were also treated with ciprofloxacin, rifampicin, linezolid, or vancomycin at four times their MICs. Treated cultures were incubated at 37°C for 10 min, and macromolecules then precipitated with 10% TCA. Samples were incubated on ice for 30 min to allow complete macromolecular precipitation and subsequently passed through glass microfiber filters. Filtered macromolecules were washed twice with 1 mL ethanol and then allowed to dry overnight. To determine levels of radiolabel incorporation, filters were submerged in scintillation fluid and analyzed by a liquid scintillation counter. Indicated are the percent CPMs observed for each sample relative to a DMSO-treated control.

The effect of compound **2** on translation agreed well with the observed cross-resistance phenotype for **2121**-resistant bacterial strains; therefore, efforts were made to confirm the inhibition of translation by **2121**. **2121** was evaluated in a protein synthesis assay in which the incorporation of [^{35}S]-methionine into protein was analyzed by SDS-PAGE followed by autoradiography. In this assay, **2121** was found to inhibit protein synthesis at least as potently as several known antibiotic inhibitors of translation (Figure 6). In addition, an inactive derivative of **2121**, compound **3**, was found to have no effect on translation, suggesting a link between protein synthesis inhibition and antibacterial activity for this class of compounds. This link was further substantiated by the observation that **2121** has a diminished effect on the incorporation of [^{35}S]-methionine into protein in **2121**-resistant *S. aureus* than in corresponding parent strains (Figure 7).

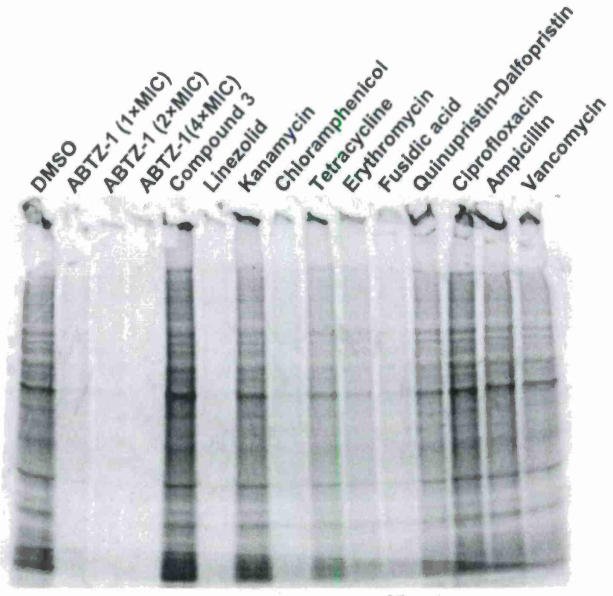


Figure 6. Evaluation of 2121 (also called ABTZ-1) in an [^{35}S]-methionine incorporation assay. *S. aureus* (ATCC 25923) cultures were co-treated with 100 µCi/mL [^{35}S]-methionine and either DMSO or ABTZ-1 at 4, 8, or 16 µg/mL. Additional cultures were treated similarly with control antibiotics at four times their MICs. Treated cultures were incubated at 37°C for 10 minutes. Casamino acids, cold methionine, and cold cysteine were subsequently added to each sample to final concentrations of 1 mg/mL, 0.1 mg/mL, and 0.1 mg/mL, respectively, and samples were then incubated at room temperature for 5 minutes. Proteins were precipitated on ice with 7.5% TCA for 30 minutes and isolated by centrifugation. Levels of [^{35}S]-methionine incorporation were determined by analysis of isolated proteins by SDS-PAGE followed by autoradiography. Lane 1: 1% DMSO. Lane 2-4: ABTZ-1 at 4, 8, and 16 µg/mL, respectively. Lane 5: 16 µg/mL compound 3. Lane 6: 8 µg/mL linezolid. Lane 7: 16 µg/mL kanamycin. Lane 8: 32 µg/mL chloramphenicol. Lane 9: 4 µg/mL tetracycline. Lane 10: 2 µg/mL erythromycin. Lane 10: 2 µg/mL fusidic acid. Lane 11: 4 µg/mL quinupristin-dalfopristin. Lane 12: 1 µg/mL ciprofloxacin. Lane 13: 0.5 µg/mL ampicillin. Lane 14: 8 µg/mL vancomycin.

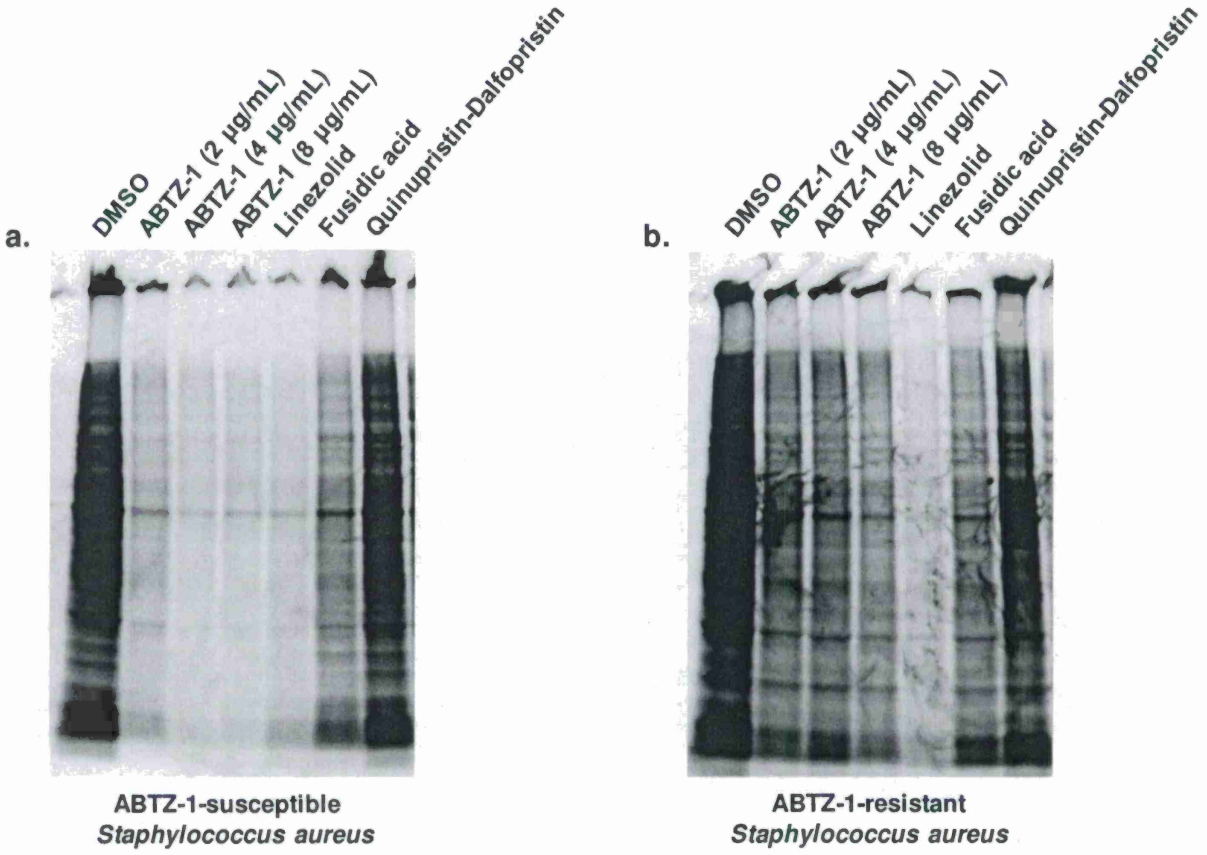


Figure 7. The effect of 2121 (also called ABTZ-1) on protein synthesis in ABTZ-1-resistant and -susceptible *S. aureus* strains as assessed by an [^{35}S]-methionine incorporation assay. Liquid cultures of ABTZ-1-resistant and ABTZ-1-suceptible *S. aureus* (ATCC 12608) were co-treated with 100 µCi/mL [^{35}S]-methionine and either DMSO or ABTZ-1 at 2, 4, or 8 µg/mL. Additional cultures were treated similarly with control antibiotics linezolid, fusidic acid, and quinupristin-dalfopristin at 8, 0.5, and 2 µg/mL, respectively. Treated cultures were incubated at 37°C for 10 minutes. Casamino acids, cold methionine, and cold cysteine were subsequently added to each sample to final concentrations of 1 mg/mL, 0.1 mg/mL, and 0.1 mg/mL, respectively, and samples were then incubated at room temperature for 5 minutes. Proteins were precipitated on ice with 7.5% TCA for 30 minutes and isolated by centrifugation. Levels of [^{35}S]-methionine incorporation were determined by analysis of isolated proteins by SDS-PAGE followed by autoradiography. To facilitate comparison of gel images for ABTZ-1-resistant and ABTZ-1-susceptible strains, the brightness and contrast of that for the ABTZ-1-susceptible strain were adjusted such that the signal intensity observed for the DMSO control for ABTZ-1-susceptible *S. aureus* matched that for the ABTZ-1-resistant strain. *a.* Evaluation of DMSO, ABTZ-1, and control antibiotics on protein synthesis in ABTZ-1-susceptible *S. aureus* (ATCC 12608). Shown here is the altered gel image. *b.* Evaluation of DMSO, ABTZ-1, and control antibiotics on protein synthesis in ABTZ-1-resistant *S. aureus* (ATCC 12608). Lane 1: 1% DMSO. Lane 2-4: ABTZ-1 at 4, 8, and 16 µg/mL, respectively. Lane 5: 8 µg/mL linezolid. Lane 6: 0.5 µg/mL fusidic acid. Lane 7: 2 µg/mL quinupristin-dalfopristin.

We are excited about ABTZ-1 and derivatives, and have been evaluating their tolerability to mice, as a prelude to an *in vivo* efficacy study.

Full summary of results – Complexity-to-Diversity

Through the generosity of the Office of Naval Research we were able to conduct a high-throughput screen and identify ABTZ-1 as a promising agent for the treatment of Gram-positive infections. We conducted analogous screens versus Gram-negative organisms, and these returned no hit compounds. As we began to investigate we learned that for a compound to penetrate a Gram-negative organism it needs to have very specific physio-chemical properties, especially with respect to molecular weight, polarity, and three-dimensionality. Interestingly, compounds in screening collections simply do not fit these parameters, so it is of no surprise that our screens (and analogous ones performed by the pharmaceutical industry) returned no hits.

So, how do we identify novel agents to combat Gram-negative bacterial infections? We have begun a program in which we start with complex natural products and convert them (through <6 chemical steps) into diverse compounds but of equal complexity. We have now fully developed this Complexity-to-Diversity (CtD) paradigm and are applying it broadly. Below is described this process, our results, and some guiding principles.

Below is described a new approach for the rapid creation of complex and diverse small molecules. In this process structurally complex natural products are converted, in an average of 3 chemical steps, to markedly different core scaffolds that are distinct from each other and from the parent natural product. Using chemoselective reactions, the core ring structures of readily available natural products are systematically altered via ring system distortion reactions, i.e., ring cleavage, ring expansion, ring fusion, ring rearrangements, or combinations thereof (Figure 8A). Importantly, this method stands in contrast to traditional optimization campaigns whose goals are to enhance the inherent biological activity of a natural product (e.g., erythromycin to azithromycin, penicillin to amoxicillin, etc). To demonstrate this scaffold diversification approach, we have selected three readily available natural products from different structural classes: gibberellic acid (diterpene), adrenosterone (steroid), and quinine (alkaloid) (Figure 8). However, dozens of readily available natural products could be converted into diverse and complex molecules using this strategy. This method takes inspiration from the manner in which nature creates certain complex natural products, using a common intermediate to generate scores of compounds that are very different from one another.

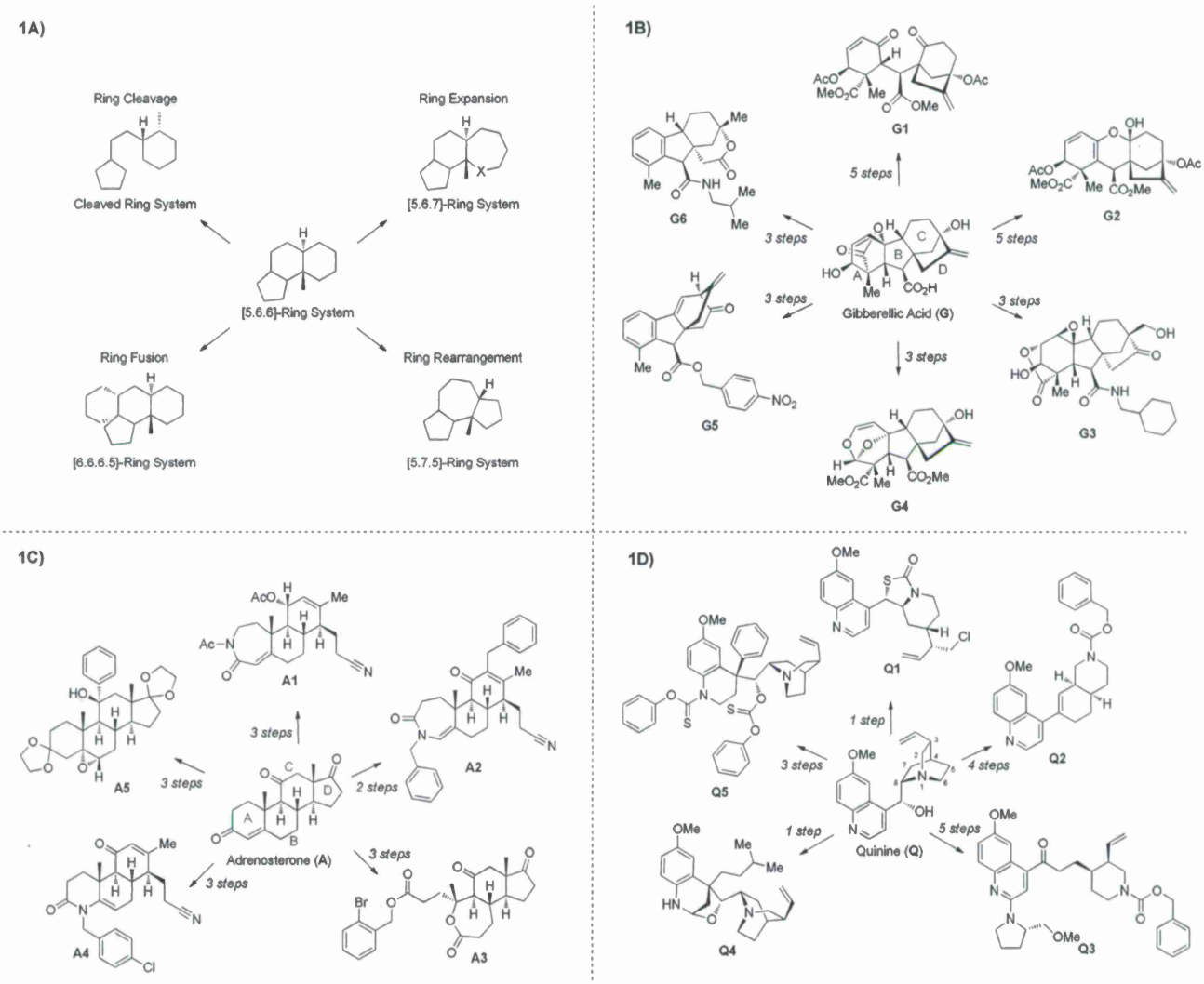


Figure 8. Natural products are readily converted to complex and diverse scaffolds. A) General transformations, B) structures produced from gibberellic acid, C) adrenosterone, and D) quinine.

Gibberellic acid

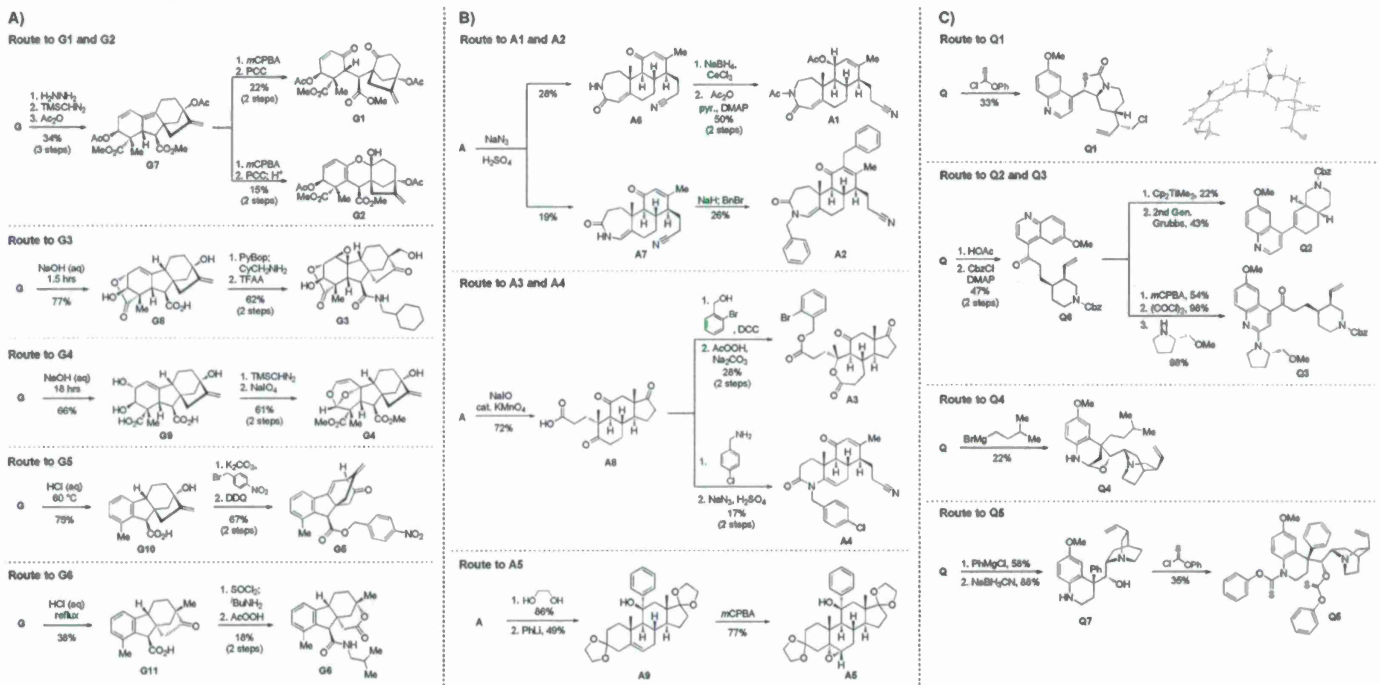
Gibberellic acid (**G**, Figure 8B) is a plant hormone isolated from *Gibberella fujikuroi* and produced industrially on the ton scale. Gibberellic acid contains a tetracyclic diterpene core with a fused lactone, two allylic alcohols, an exocyclic olefin, and a carboxylic acid, enabling selective and independent functionalization of each ring of the core structure via a variety of ring system distortion reactions. We have exploited these structural features in concert with known degradation reactions of **G** in the construction of complex and diverse scaffolds in 3-5 steps from gibberellic acid (Figure 8B, **G1**-**G6**).

Hydrazine-promoted elimination of the lactone on **G**, followed by methylation and acetylation, affords triene **G7** (Scheme 1A). Treatment with *m*CPBA yields an intermediate epoxide with complete selectivity for the tetrasubstituted olefin, which, when subjected to oxidative cleavage conditions using PCC, produces diketone **G1**. Following exposure of **G1** to silica or acid, the A-ring ketone tautomerizes

and collapses onto the C-ring ketone to form ketal **G2**; this can be achieved directly from the epoxide precursor using PCC and an acidic workup.

Exposure of gibberellic acid to basic conditions leads to lactone rearrangement and the generation of alkene **G8**. Amidation of **G8**, followed by treatment with trifluoroperacetic acid provides **G3** via epoxidation of both alkenes and Wagner-Meerwein rearrangement to afford the primary alcohol. Prolonged exposure of **G** to base leads to the cleavage of the lactone ring to provide diol **G9**. Methylation of the carboxylic acids, followed by oxidative cleavage of the diol with sodium periodate and intramolecular [4+2] cycloaddition provides acetal **G4**.

Treatment of gibberellic acid with dilute hydrochloric acid results in the elimination of the lactone and decarboxylation to aromatize the A-ring, enabling the isolation of *allo*-gibberic acid (**G10**). Esterification followed by oxidative rearrangement with DDQ gives [2.2.2]-bicycle **G5**. Exposure of gibberellic acid to refluxing hydrochloric acid results in aromatization and a Wagner-Meerwein rearrangement, forming gibberic acid (**G11**). Amidation through an intermediate acyl chloride followed by Baeyer-Villiger oxidation produces lactone **G6**.



Scheme 1. Synthesis of complex and diverse compounds.

Adrenosterone

Adrenosterone (**A**, Figure 8C) is a steroid hormone that is produced in the adrenal cortex of mammals. Adrenosterone's structurally complex steroidal framework contains five contiguous stereogenic centers; in addition, each of the four individual carbocyclic rings of adrenosterone is functionalized with an enone or ketone. Though embedded in the A-ring, the enone is also connected to the B-ring as an exocyclic double bond, while the C- and D-rings are each functionalized with a ketone.

These key functional groups provide synthetic handles that can be strategically manipulated to synthesize novel, diverse, and complex chemical scaffolds in three or fewer steps (Figure 8C, **A1-A5**).

During these synthetic investigations of adrenosterone a novel, substrate-dependent Schmidt reaction was discovered that effected both ring expansion and ring cleavage in a single synthetic transformation. Subjecting adrenosterone to Schmidt conditions for one hour gives two constitutional isomers (**A6** and **A7** in Scheme 1B) resulting from a tandem D-ring cleavage and A-ring expansion. Final dehydration of the subsequent primary amide in concentrated sulfuric acid results in the observed cyano groups in **A6** and **A7**.

Enamide **A6** and lactam **A7** were each elaborated to novel complex molecular scaffolds. Lactam **A6** undergoes a stereoselective Luche reduction of the C-ring enone, which following treatment with acetic anhydride in pyridine with catalytic DMAP affords **A1**. Treatment of enamide **A7** with sodium hydride followed by benzyl bromide results in enone **A2**.

The oxidative cleavage of adrenosterone's A-ring with $NaIO_4$ and $KMnO_4$ gives acid **A8**, which upon treatment with 2-bromobenzyl alcohol and DCC provides the corresponding ester. A selective ring expansion at the B-ring ketone using a Baeyer-Villiger reaction with peracetic acid provides lactone **A3**. Acid **A8** condenses with 4-chlorobenzylamine upon heating in ethanol to provide the corresponding A-ring substituted enamide. A final ring cleavage reaction of the D-ring using our Schmidt protocol yields **A4**. Adrenosterone undergoes a double ring fusion (at the A- and D-rings) reaction upon treatment with ethylene glycol and catalytic *p*-TsOH. The resulting ketone intermediate reacts with phenyllithium to give **A9**. Final treatment of **A9** with *m*CPBA results in epoxide ring fusion at the B-ring to yield **A5**.

Quinine

Quinine (**Q**, Figure 8D), an alkaloid isolated from the bark of the genus *Cinchona*, is used commercially as an anti-malarial therapeutic and food additive. Unlike other natural products employed herein, quinine is composed of two discrete ring systems; however, the stereochemical complexity and diverse functionality (a tertiary amine, a secondary alcohol, an olefin, and a quinoline) of quinine make it amenable to selective ring system distortion to create diverse molecular scaffolds (Figure 8D, **Q1-Q5**).

Ring cleavage of the quinuclidine by *O*-phenyl thionochloroformate occurs selectively at N1-C2. In addition to the expected ring cleavage and chloride addition, this reaction also leads to diastereoselective rearrangement of the free alcohol and thiocarbamate to form *S*-thiocarbamate **Q1** (Scheme 1C) as a single diastereomer, as confirmed by x-ray crystallography.

In contrast to the selectivity observed with *O*-phenyl thionochloroformate, acid-catalyzed Hoffman-type elimination of quinine occurs exclusively at N1-C8, and addition of benzyl chloroformate to the crude degradation product results in ketone **Q6**, which was elaborated to form two unique structures (**Q2-Q3**). Petasis methylenation of ketone **Q6** followed by 1,2-ring fusion via ring-closing metathesis using second generation Grubbs catalyst forms [4.4.0]-bicycle **Q2**. Quinoline *N*-oxidation of **Q6** using *m*CPBA, chlorination with oxalyl chloride, and nucleophilic displacement of the chloride by (*S*)-2-(methoxymethyl)pyrrolidine provides amine **Q3**.

Ring system distortions of the quinoline ring were accomplished through the addition of Grignard reagents. Exposure of quinine to isoamylmagnesium bromide in toluene results in nucleophilic addition to

the quinoline ring followed by hemiaminal ether formation to provide **Q4** as a single diastereomer. Alternatively, reduction of the hemainal ether formed through the addition of phenylmagnesium chloride to quinine with sodium cyanoborohydride provides tetrahydroquinoline **Q7**. Treatment of **Q7** with *O*-phenyl thionochloroformate results in bis-acylation to form **Q5** as the major product with no observed chlorine incorporation.

Compound Analysis

In contrast to most standard small molecule library constructions in which simple starting materials are built up into more complex products, an important consequence of starting with natural products is that all intermediates are structurally complex and worthy of inclusion in the final library in their own right. For example, in the synthesis of the compounds depicted in Figures 8B, 8C, and 8D, there are 19, 14, and 12 complex structures produced, respectively. Thus, detailed in Figure 1 is the synthesis of 45 structurally and stereochemically complex small molecules from three readily available natural products. In addition, each of these compounds possesses sites for diversification, allowing for the facile and rapid creation of hundreds of complex and diverse compounds.

Advances in chemoinformatics have enabled evaluation of massive chemical and biological data sets, allowing for a rough determination of the structural features of small molecules that correlate with biological activity. It is apparent that many compounds in screening collections have non-trivial liabilities, including non-specific reactivity and a propensity to aggregate, leading to false positives and complicating the development of a hit into a drug. When specific disease areas are examined, the problem is more acute. For example, analysis of compounds that kill Gram-negative bacterial pathogens shows an average ClogP of -0.1, a realm occupied by vanishingly few compounds in commercial screening collections.

In an attempt to quantify the structural complexity and diversity of the novel compounds created through this paradigm, structural features known to track with biological activity were analyzed. A recent study has examined eight structural parameters (molecular weight, ClogP, polar surface area, rotatable bonds, hydrogen bond donors/acceptors, complexity, and fraction sp^3 carbons (Fsp3)) of compounds synthesized by medicinal chemists over the last 50 years, and then compared them to marketed drugs. One of the most important conclusions from this analysis is that medicinal chemists are creating compounds with lower-than-ideal Fsp3, and with ClogP values that are higher than ideal. We have calculated Fsp3 and ClogP of the compounds described herein and have compared them to compounds in large screening collections. For this analysis we have used a 150,000-member compound collection from the ChemBridge MicroFormat Library – a standard commercial screening collection and one used by others in comparison analyses.

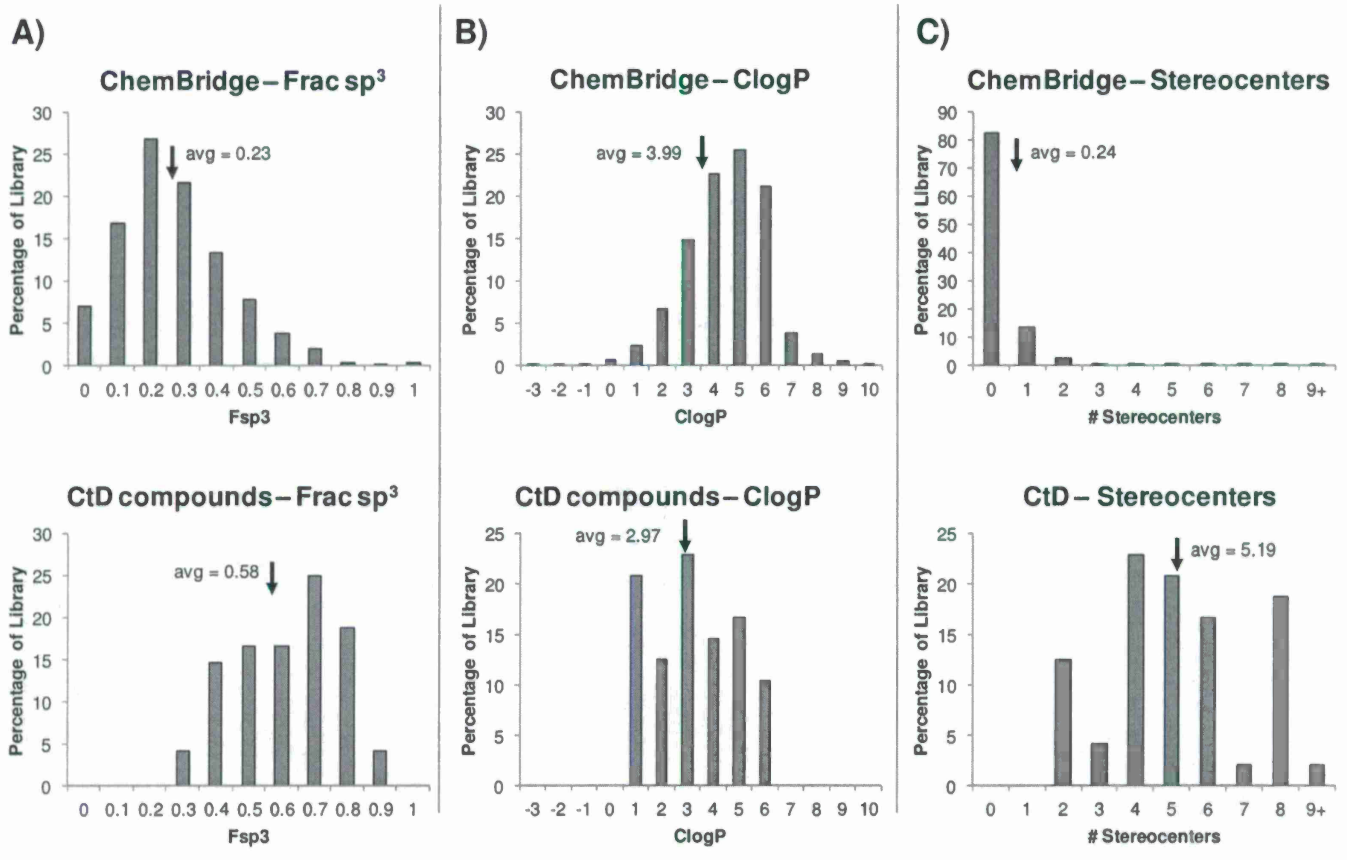


Figure 9. Compounds created through the complexity-to-diversity (CtD) method have markedly different properties from those in commercial screening collections. For this analysis, a 150,000 compound collection from ChemBridge Corporation was utilized and compared to the 45 compounds described herein for A) fraction of sp^3-hybridized carbons (Fsp3), B) calculated logP (ClogP), and C) number of stereogenic centers per compound.

Fsp3 is the number of sp^3-hybridized carbon atoms in a compound divided by the sum of sp^3- and sp^2-hybridized carbon atoms. Studies have shown the benefits of higher Fsp3, including lower melting points and enhanced aqueous solubility, and have revealed that "discovery" compounds have lower Fsp3 than actual drugs (0.36 vs 0.47). The analysis of medicinal compounds synthesized over the last 50 years has shown that average Fsp3 is declining, a result attributed in part to the increasing ease of sp^2-sp^2 coupling reactions. ClogP is often used as a rough measure of lipophilicity; among other things, compounds with higher ClogP values tend to have non-ideal solubility, promiscuity, and off-target toxicity. Analysis has shown that the average ClogP for all medicinal compounds synthesized since 1985 has gone up significantly, and is higher than the average for marketed drugs. Indeed, a survey of 18 pharmaceutical companies from 2000-2010 shows the majority are still synthesizing compounds with a mean ClogP over 4. As shown by the histograms in Figure 9A, the new compounds described herein have an average Fsp3 of 0.58, which is considerably higher than those in the commercial collection (avg = 0.23). In a similar vein, the average ClogP for these compounds is a factor of 10 lower than those in the commercial screening set (2.97 vs. 3.99; Figure 9B).

The presence of stereogenic centers in a compound can also be used as a surrogate for molecular complexity. Compounds with stereogenic centers may interact more specifically with their chiral receptors, and compounds with low or no stereogenic centers are more prone to attrition during the various stages of drug discovery. Commercial screening collections are dominated by achiral compounds; for example, of the 150,000 compound ChemBridge collection, 82% have no stereogenic centers and 14% have a single stereocenter, leaving only 4% of these compounds with multiple stereogenic centers (Figure 2C). Obviously, the synthesis of complex and diverse compounds using natural products as input materials offers a tremendous advantage in this regard. Of the 45 compounds disclosed herein, all have two or more stereogenic centers, with the median number being five (Figure 9C).

While visual inspection of the structures in Figure 8 readily reveals considerable structural diversity, we have applied a similarity metric to make more quantitative comparisons. For this, Tanimoto coefficients were generated in Discovery Studio (Accelrys) using ECFP_6 molecular fingerprints. Structures in Figure 8B-D were used as the reference input for every other compound in the set, and a similarity score was obtained for each pair on a scale from 0 to 1, with 1 representing perfect similarity. As shown by the data in Figure 10, the compounds depicted in Figure 1B-D are very different both from one another and from the parent natural products. The G, A, and Q compound sets are expected to be quite different from one another; however, even within the sets the compounds show low Tanimoto coefficients (average for G set = 0.15, average for A set = 0.15, average for Q set = 0.22), indicating considerable structural diversity. For calibration purposes, this analysis was also performed on structures representing simple modifications to the parent compounds. As expected, these minor structural changes afford higher similarity scores (average of 0.7), consistent with work of others using Tanimoto coefficients.

	G	G1	G2	G3	G4	G5	G6	A	A1	A2	A3	A4	A5	Q	Q1	Q2	Q3	Q4	Q5	
G	1.00	0.14	0.16	0.18	0.38	0.09	0.14	0.10	0.09	0.07	0.11	0.07	0.10	0.06	0.06	0.06	0.06	0.06	0.06	G
G1	0.14	1.00	0.29	0.09	0.17	0.09	0.11	0.11	0.13	0.07	0.10	0.08	0.06	0.06	0.07	0.07	0.06	0.06	0.07	G1
G2	0.16	0.29	1.00	0.09	0.19	0.10	0.11	0.08	0.11	0.06	0.09	0.06	0.08	0.06	0.06	0.06	0.06	0.06	0.06	G2
G3	0.18	0.09	0.09	1.00	0.15	0.07	0.14	0.09	0.08	0.07	0.09	0.06	0.08	0.04	0.05	0.05	0.05	0.05	0.04	G3
G4	0.38	0.17	0.19	0.15	1.00	0.10	0.12	0.09	0.09	0.06	0.09	0.06	0.11	0.07	0.07	0.07	0.07	0.07	0.07	G4
G5	0.09	0.09	0.10	0.07	0.10	1.00	0.16	0.08	0.10	0.08	0.12	0.09	0.05	0.05	0.07	0.12	0.10	0.06	0.06	G5
G6	0.14	0.11	0.11	0.14	0.12	0.16	1.00	0.10	0.11	0.10	0.13	0.09	0.09	0.06	0.08	0.08	0.07	0.09	0.07	G6
A	0.10	0.11	0.08	0.09	0.09	0.08	0.10	1.00	0.17	0.16	0.31	0.13	0.09	0.05	0.06	0.06	0.04	0.05	0.05	A
A1	0.09	0.13	0.11	0.08	0.09	0.10	0.11	0.17	1.00	0.23	0.12	0.24	0.07	0.06	0.08	0.10	0.08	0.07	0.08	A1
A2	0.07	0.07	0.06	0.07	0.06	0.08	0.10	0.16	0.23	1.00	0.12	0.26	0.10	0.05	0.07	0.11	0.10	0.06	0.08	A2
A3	0.11	0.10	0.09	0.09	0.09	0.12	0.13	0.31	0.12	0.12	1.00	0.10	0.09	0.05	0.07	0.11	0.09	0.08	0.07	A3
A4	0.07	0.08	0.06	0.06	0.06	0.09	0.09	0.13	0.24	0.26	0.10	1.00	0.07	0.05	0.07	0.07	0.06	0.06	0.05	A4
A5	0.10	0.06	0.08	0.08	0.11	0.05	0.09	0.09	0.07	0.10	0.09	0.07	1.00	0.06	0.05	0.07	0.06	0.06	0.10	A5
Q	0.06	0.06	0.06	0.04	0.07	0.05	0.06	0.05	0.06	0.05	0.05	0.05	0.06	1.00	0.30	0.26	0.20	0.28	0.30	Q
Q1	0.06	0.07	0.06	0.05	0.07	0.07	0.08	0.06	0.08	0.07	0.07	0.07	0.05	0.30	1.00	0.23	0.18	0.17	0.13	Q1
Q2	0.06	0.07	0.06	0.05	0.07	0.12	0.08	0.06	0.10	0.11	0.11	0.07	0.07	0.26	0.23	1.00	0.34	0.13	0.15	Q2
Q3	0.06	0.06	0.06	0.05	0.07	0.10	0.07	0.04	0.08	0.10	0.09	0.06	0.06	0.20	0.18	0.34	1.00	0.14	0.16	Q3
Q4	0.06	0.06	0.06	0.05	0.07	0.06	0.09	0.05	0.07	0.06	0.08	0.06	0.06	0.28	0.17	0.13	0.14	1.00	0.26	Q4
Q5	0.06	0.07	0.06	0.04	0.07	0.06	0.07	0.05	0.08	0.08	0.07	0.05	0.10	0.30	0.13	0.15	0.16	0.26	1.00	Q5
	G	G1	G2	G3	G4	G5	G6	A	A1	A2	A3	A4	A5	Q	Q1	Q2	Q3	Q4	Q5	

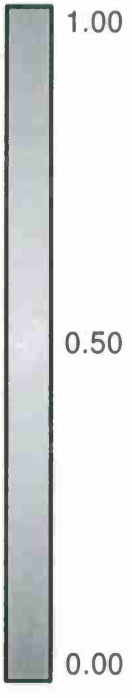


Figure 10. Tanimoto similarity coefficients for compounds from Figure 8B-D relative to the three natural products and to each other, where 1.0 represents perfect similarity.

Guidelines

This new complexity-to-diversity approach has been demonstrated with three complex natural products; however, the same logic and methods can be applied to a multitude of other natural products. In general, compounds most amenable to diversification through this method will be available in suitable quantities (either from commercial sources or through isolation) and possess orthogonal functional groups allowing for ring system distortion and diversification through chemoselective reactions. As exemplified with gibberellic acid, adrenosterone, and quinine, certain common ring distortion strategies facilitate the rapid diversification (in ≤5 synthetic steps) of complex natural products:

1) *Ring cleavage reactions.* Ring cleavage reactions enable dramatic structural changes in one chemical step, and is the most utilized tactic in our complexity-to-diversity approach. A benefit of ring cleavage reactions is that they typically provide new functional groups that can be further diversified. Examples include the base-promoted hydrolysis of the lactone on allogibberellic acid (**G9**), oxidative cleavage on adrenosterone (**A8**), and N-C cleavage on quinine (**Q1** & **Q6**).

2) *Ring expansion reactions.* Ring expansion reactions are useful in forming novel ring skeletons or as a prelude to ring cleavage reactions. There are several options for chemical reactions that induce ring expansion, with the Baeyer-Villiger and Schmidt reactions being powerful methods to target ketone and enone functionalities for ring expansion. Ring expansion, as a ring system distortion tactic, has been successfully applied in the synthesis of target structures **G6**, **A3**, **A6**, and **A7**.

3) *Ring fusion reactions.* Ring fusion reactions can provide further diversification by connecting disparate structural elements in the pre-existing ring system or simple addition of a new, constrained ring to the ring system. Various modes of ring fusion were used to demonstrate this tactic on each natural product. For example, the ring-closing metathesis product **Q2** (1,2-ring fusion), formation of the [4+2] cycloaddition product **G4**, and formation of bis-ketal **A9** (1,1-ring fusion) result from ring fusion reactions. *Ring substitution reactions* can also be an example of ring fusion. In the example of ring substitution, the composition of the ring changes without altering the ring size, as exemplified by the creation of the A-ring enamide in **A4**.

4) *Ring rearrangement reactions.* Ring rearrangement reactions that dramatically reorganize the core structure will be dictated by the natural product and are thus applicable on a case-by-case basis. This tactic is illustrated with gibberellic acid in the Wagner-Meerwein rearrangements of **G3** and **G11**, or the DDQ oxidation of **G10**. These transformations are facilitated by the propensity of the tertiary alcohol in gibberellic acid's C-ring to form a ketone upon carbon migration, altering the molecular topology.

Historically, natural products or their close analogues have been considered as end points in the drug discovery process. Indeed, this thinking has been quite fruitful; for example, 41% of anticancer drugs and 65% of antibacterial drugs are natural products or very close derivatives. The features that make natural products different from most synthetic compounds (e.g., high Fsp3, low ClogP, presence of stereogenic centers) give these compounds a propensity to bind to their macromolecular target with high affinity and specificity, while retaining the solubility and cell permeability needed for a therapeutic agent. The approach described herein uses natural products not as the end point, but as the starting point for the discovery process, starting with compounds inherently biased for biological success and systematically transforming them into diverse compounds of equal complexity.

Certain chemical properties, such as molecular complexity and multiple stereogenic centers, are extremely difficult to build in when producing large collections of compounds for high-throughput screening. The systematic application of ring distortion reactions on appropriate natural product starting materials offers a convenient approach to rapidly generate large numbers of complex and diverse small molecules. These compounds possess a high degree of molecular complexity, as shown by examination of Fsp3 and number of stereogenic centers, and are structurally diverse, as indicated by Tanimoto similarity analysis. Depending on the exact application, specific structural features (e.g., MW, ClogP, H-bond donors/acceptors, etc) can be programmed in by careful selection of diversification reactions and building blocks. This method to construct complex and diverse small molecules can rapidly provide compounds with properties suitable for a wide variety of biological and medicinal applications.

As part of ONR award N000141210616 (start date of May 1, 2012) we have expanded this CtD concept to create libraries that are biased for penetration of Gram-negative bacteria, and are developing the associated antibacterial screens.

We are thankful to the ONR for support of our work, and we look forward to working with ONR again in the future.