

MIT Open Access Articles

An apparent binary choice in biochemistry: mutual reactivity implies life chooses thiols or nitrogen-sulfur bonds, but not both

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Petkowski, Janusz J., William Bains, and Sara Seager, "An apparent binary choice in biochemistry: mutual reactivity implies life chooses thiols or nitrogen-sulfur bonds, but not both." *Astrobiology* 19, 4 (March 2019): p. 579-613 doi 10.1089/AST.2018.1831 ©2019 Author(s)

As Published: 10.1089/AST.2018.1831

Publisher: Mary Ann Liebert Inc

Persistent URL: <https://hdl.handle.net/1721.1/124665>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



An Apparent Binary Choice in Biochemistry: Mutual Reactivity Implies Life Chooses Thiols or Nitrogen-Sulfur Bonds, but Not Both

Janusz J. Petkowski,¹ William Bains,² and Sara Seager^{1,3}

Abstract

A fundamental goal of biology is to understand the rules behind life's use of chemical space. Established work focuses on why life uses the chemistry that it does. Given the enormous scope of possible chemical space, we postulate that it is equally important to ask why life largely avoids certain areas of chemical space. The nitrogen-sulfur bond is a prime example, as it rarely appears in natural molecules, despite the very rich N-S bond chemistry applied in various branches of industry (*e.g.*, industrial materials, agrochemicals, pharmaceuticals). We find that, out of more than 200,000 known, unique compounds made by life, only about 100 contain N-S bonds. Furthermore, the limited number of N-S bond-containing molecules that life produces appears to fall into a few very distinctive structural groups. One may think that industrial processes are unrelated to biochemistry because of a greater possibility of solvents, catalysts, and temperatures available to industry than to the cellular environment. However, the fact that life does rarely make N-S bonds, from the plentiful precursors available, and has evolved the ability to do so independently several times, suggests that the restriction on life's use of N-S chemistry is not in its synthesis.

We present a hypothesis to explain life's extremely limited usage of the N-S bond: that the N-S bond chemistry is incompatible with essential segments of biochemistry, specifically with thiols. We support our hypothesis by (1) a quantitative analysis of the occurrence of N-S bond-containing natural products and (2) reactivity experiments between selected N-S compounds and key biological molecules. This work provides an example of a reason why life nearly excludes a distinct region of chemical space. Combined with future examples, this potentially new field of research may provide fresh insight into life's evolution through chemical space and its origin and early evolution. Key Words: N-S bond—Nitrogen-sulfur bond—Thiols—Chemical space. *Astrobiology* 19, 579–613.

1. Introduction

UNDERSTANDING WHY life uses the chemistry that it does is one of the fundamental questions in biology. We postulate that, given the enormous scope of possible chemical space, it is equally important to explain why life *does not* utilize certain chemical functionalities. While there are a wide range of hypotheses published as to why life uses certain specific chemical functionalities (*e.g.*, peptide bond, phosphates [Westheimer, 1987; Pace, 2001; Benner *et al.*, 2004]), there are no studies that address the question why life *does not* use certain specific chemical classes of molecules that are chemically flexible, stable, and have wide chemical and structural functionality. After a lengthy combinatorics

and database-curating exercise (Seager *et al.*, 2016), we have quantified occurrence of different bonds, molecular fragments, and molecules among life's natural products. Through this previous work, we have consolidated known, yet chemically puzzling, gaps amid the vast diversity of the chemistry of life.

One specific gap in life's use of chemistry is that life rarely makes compounds containing N-S bonds. The few papers reporting compounds that contain N-S bonds comment on the rarity of this chemistry (Blunt *et al.*, 2015) (for a detailed review, see Petkowski *et al.* [2018]). Perhaps due to their rarity among natural products, little research has been published on biochemistry involving N-S bond-containing compounds (with an exception of S-nitrosothiols,

¹Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

²Rufus Scientific, Melbourn, UK.

³Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.

utilized in redox metabolism of the cell, *e.g.*, reviewed by Broniowska and Hogg [2012] and Cortese-Krott *et al.* [2016]).

1.1. Motivation to study N-S bond-containing compounds

In this work, we pursue N-S bond-containing compounds to illustrate that the extent of life's avoidance of certain chemical functionalities can be not only quantified but also systematically explored. There are more instances of exclusions in biochemistry that go beyond N-S chemistry, and our hope is that collectively they may yield insights into the origin and evolution of life. Our aim for this work, however, is to provide the first hypothesis, and example, of a chemical explanation of why life *does not* use a chemical class of molecules that is chemically flexible, stable, and has wide chemical and structural functionality.

1.2. N-S bond-containing compounds in industry

Life's apparent avoidance of N-S bond chemistry is surprising in light of the diversity, flexibility, and utility of sulfur chemistry. Sulfur can be found in compounds stable to hydrolysis, stable to oxidation by O₂ and reduction by H₂, in all 8 of sulfur's oxidation states between -2 and +6, and in (II), (IV), or (VI)-coordinate centers supporting multiple single or double bonds, a flexibility offered by no other element. The structural and functional diversity enabled by the variety of oxidation and valence states of sulfur is reflected by the abundance of N-S chemistry in human endeavors. In modern industry, N-S bond compounds have countless applications. N-S bond use in pharmaceuticals is well known; for example, the first widely used synthetic antibiotic was a sulfonamide called prontosil (Hörlein, 1909). Over 500 of the 7200 compounds in the DrugBank database¹ contain N-S bonds, and their widespread and diverse pharmacology supports the idea that N-S bond-containing compounds can have functional and biologically useful properties and that life could use N-S bonds to a powerful effect. N-S bond-containing compounds are also used by humans as intermediates in drug syntheses (Craine and Raban, 1989; Koval, 1990, 1996a, 1996b; Petrov *et al.*, 1990; Lücking, 2013).

N-S compounds are used in a wide range of other applications beyond pharmaceuticals, such as components of lubricating greases, and dyes (Slack and Wooldridge, 1965; Fleischer *et al.*, 1983; Craine and Raban, 1989), as herbicides, fungicides (both in pharmacology and in the paint industry), pesticides, and antimicrobial agents (*e.g.*, industrial antimicrobial finishes for textiles) (Beck, 1975; Craine and Raban, 1989; Kalgutkar *et al.*, 2010; Bland *et al.*, 2014). The industrial uses are not limited to small-quantity production; some are used in 100 kiloton amounts, industrial quantities². For example, sulfenamides are essential for the modern rubber industry as vulcanization agents; in the automotive industry alone, approximately one-third of the utilized vulcanization procedures (corresponding to 50% of

the total market) depend on N-S sulfenamide chemistry (Chenier, 2002).

In addition to commonly used N-S bond-containing compounds, there is a growing set of others, including N-S(VI), N-S aromatics, N-S(IV), and N-S(II) bond-containing molecules under increased study, with uses patented (more than 30 patents over the last decade or so) (Katritzky *et al.*, 2008; Arndt *et al.*, 2009; Reitz *et al.*, 2009; Carta *et al.*, 2012; Loso *et al.*, 2012; Scozzafava *et al.*, 2013; Bowden *et al.*, 2014). Every oxidation state and valence state of S in N-S bond-containing compounds finds a use or potential use in industry and medicine, as shown by examples in Table 1.

1.3. The extent of chemical space of N-S bond-containing compounds

The hundreds of chemicals that contain N-S bonds that are found to be useful by humans are themselves a very small subset of the chemical space of possible, stable N-S compounds that can theoretically exist. The number of different N-S ring systems alone exceeds many hundreds (Katritzky and Rees, 1984; Katritzky *et al.*, 1996, 2008). To illustrate the scale of possible N-S chemistry, we generated a set of possible chemical structures that (a) are composed of up to 7 atoms selected from S, P, O, N, and C (with as many H atoms as was required to satisfy valence rules) and (b) have structures that are plausibly stable in the presence of water, using an algorithm derived from that described in the work of Bains and Seager (2012) and Seager *et al.* (2016). We found that 18% (180,397 out of 957,078) of this computationally generated set of structures contained N-S bonds.

Thus N-S chemistry provides an immense source of potential functionalities that life could exploit, illustrated both by the possible theoretical space of N-S chemistry and wide use of N-S chemistry in industry. Yet the literature implies that life does not use N-S chemistry as a potential source of chemical functionality (Petkowski *et al.*, 2018).

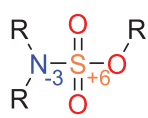
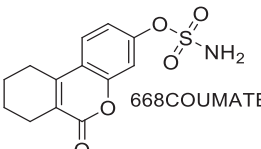
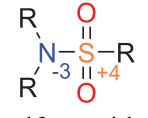
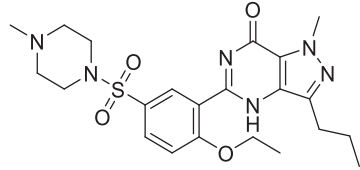
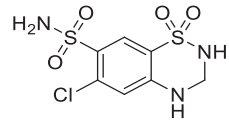
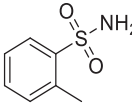
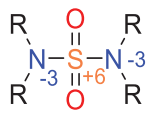
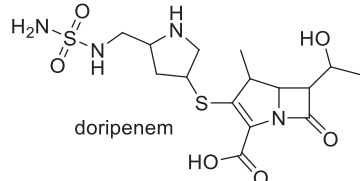
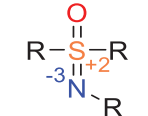
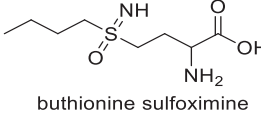
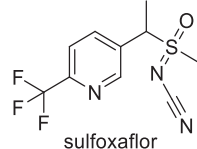

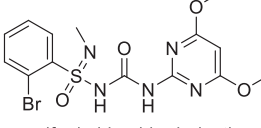
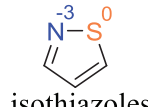
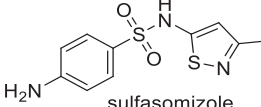
One might wonder whether the high contrast between the very small number of N-S compounds produced by life and the very large number of pharmacologically active and industrial N-S compounds created by humans has any meaning. The conventional thinking is that there is no relationship at all, in that the chemical versatility of certain chemical groups (*e.g.*, N-S bond-containing functional groups), their biological activity, and their potential usefulness in industry have no correlation to what is possible biochemically. Typically, the scientific community adopts the thinking that in biological systems there are a limited number of chemical "building blocks," and an aqueous, oxygenic environment will inevitably constrain the types of bonds that can be constructed. The thinking continues that with the precursors available and under endogenous conditions, certain types of biochemistry typically predominate over others, unlike in organic chemistry in general, where a much wider range of conditions and starting materials are available.

We do not accept this explanation. First, the precursors for making an N-S bond, that is, chemical "building blocks," are various sulfur-containing compounds and nitrogen-containing compounds (most commonly amines) that are very common

¹<https://www.drugbank.ca>

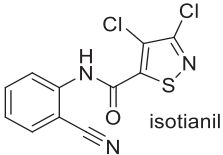
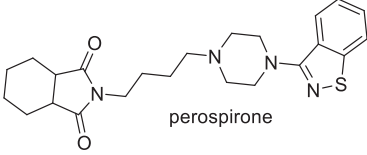
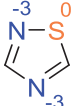
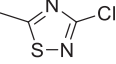
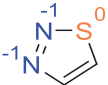
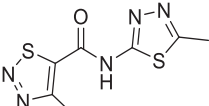
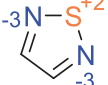
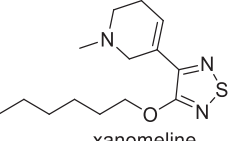
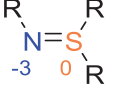
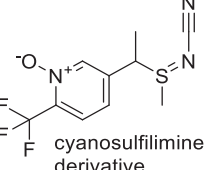
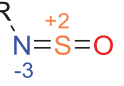
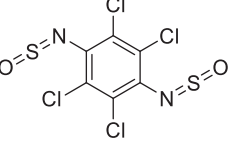
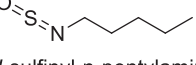
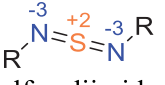
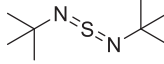
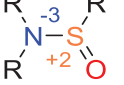
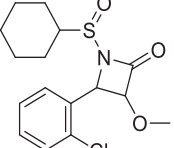
²EG entries 202-409-1 and 202-411-2 in the EU's REACH database of industrial chemical safety data, both listed as being used in the 10–100 kiloton category in Europe alone.

TABLE 1. EXAMPLES OF INDUSTRIALLY IMPORTANT N-S COMPOUNDS

Structural Class/Type	Example of Chemicals Used (or of Potential Use) in Human Industry	Industrial or Medicinal Use	Reference
 sulfamates	 668COUMATE	potential anti-cancer drugs	(Winum <i>et al.</i> , 2005)
 sulfonamides	 viagra	erectile dysfunction treatment	(Steinhagen, 2011)
	 hydrochlorothiazide	diuretic drug	(Steinhagen, 2011)
	 toluene sulfonamide	industrial plasticizer	(Bergen and Craver, 1947)
 sulfamides	 doripenem	broad spectrum antibiotic	(Mazzei, 2010)
 sulfoximines	 buthionine sulfoximine	adjunct in chemotherapy	(Defty and Marsden, 2012)
	 sulfoxaflor	insecticide	(Tomizawa and Casida, 2003)
 sulfonimidamides	 sulfonimidamide derivative	herbicide	(Hillemann, 1986)
 isothiazoles	 sulfasomizole	broad spectrum antibiotic	(Adams <i>et al.</i> , 1960)

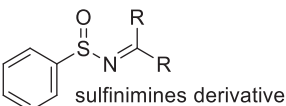
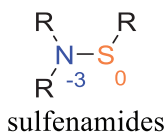
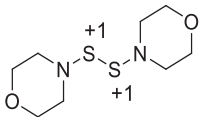
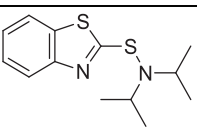
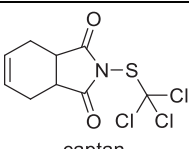
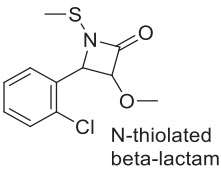
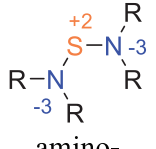
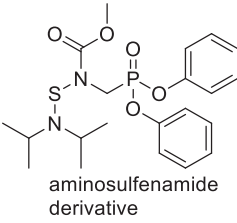
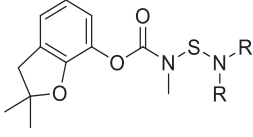
(continued)

TABLE 1. (CONTINUED)

		 isotianil	fungicide	(Hitoshi <i>et al.</i> , 1993)
		 perospirone	anti-psychotic drug	(de Paulis, 2002)
	 1,2,4-thiadiazoles	 3-chloro-5-methyl-1,2,4-thiadiazole	dye component	(Bradbury <i>et al.</i> , 1996)
	 1,2,3-thiadiazoles	 1,2,3-thiadiazole derivative	fungicide	(Katritzky <i>et al.</i> , 2008)
	 1,2,5-thiadiazoles	 xanomeline	anti-psychotic drug	(Lieberman <i>et al.</i> , 2008; Heinrich <i>et al.</i> , 2009)
N-S(IV)	 sulfilimines	 cyanosulfilimine derivative	insecticide	(Bland <i>et al.</i> , 2014)
	 sulfinyl-amines	 aryl-N-sulfinylamine	broad spectrum biocide	(Hooks and Ottmann, 1966)
		 N-sulfinyl-n-pentylamine	gasoline antiknock agent	(Licke, 1971)
	 sulfur-diimides	 di-t-butylsulfurdiimide	reagents for industrial and laboratory organic synthesis	(Kresze and Wucherpfennig, 1967)
	 sulfenamides	 sulfenamide beta-lactam	antibiotic	(Heldreth <i>et al.</i> , 2006)

(continued)

TABLE 1. (CONTINUED)

		 sulfinimines derivative	reagents for industrial and laboratory organic synthesis	(Davis, 2006; Morton and Stockman, 2006)
N-S(II)	 sulfenamides	 <i>N,N</i> -di-thiobis(morpholine)	vulcanization agent	(Craine and Raban, 1989)
		 benzothiazole sulfenamide	component of lubricants and greases	(Craine and Raban, 1989)
		 captan	fungicide	(Craine and Raban, 1989)
		 N-thiolated beta-lactam	antibiotic	(Revell <i>et al.</i> , 2007)
	 amino-sulfenamides	 aminosulfenamide derivative	herbicide	(Sikorski and Hoobler, 1984)
		 benfuracarb and furathiocarb	pesticide	(Sannino <i>et al.</i> , 2004; Yi <i>et al.</i> , 2015)

in biochemistry. Second, although industrial processes do often use non-aqueous solvents, exotic and toxic metal catalysts, and high temperatures to make chemicals, including those containing N-S bonds, living cells could still find a way to synthesize N-S bond-containing compounds where functionally useful. There are many parallels in other types of chemistry. For example, it is well known that compounds such as esters, nucleic acids, and peptides are made both in industrial settings and by living cells, via completely different routes. Living cells use alternative routes to synthetic methods in industry, using different catalysts to make the same chemical structures in water and at moderate temperatures.

The fact that life *can* make any N-S bonds at all from the plentiful precursors available, and has evolved the ability to do so independently several times (discussed below), sug-

gests that the restriction on life's use of N-S chemistry is not in its synthesis.

In this work, we first describe our methods (Section 2); next we quantify the rarity of N-S bonds in biochemistry and further characterize the relative rarity of different classes of N-S bond in natural products (Section 3). This leads us to propose that N-S bond chemistry is fundamentally incompatible with a key aspect of biochemistry, namely thiols (Section 3). Thiols are ubiquitous and essential components of life. A chemical species that is inherently reactive toward them, in an unconstrained, unregulated fashion, will be, in consequence, incompatible with biochemistry. We argue that reactivity toward thiols is the reason that N-S compounds are rare in biology (Section 3), show why exceptions fit with this argument, and discuss other potential explanations and challenges to our hypothesis (Section 4).

2. Materials and Methods

2.1. Creation of the natural products database

We assembled a large database of chemicals made by life (“natural products”), by expanding the part of our previously developed database³ focused on volatile molecules produced by life (Seager *et al.*, 2016), with an extensive literature search and by trawling publicly available online natural product repositories (see Appendix Table A2). Compiling a complete list of all that is known about each chemical made by life was a surprisingly challenging task. Although many databases of compounds made by life are available, no individual database covers more than 20% of the known natural products (Farnsworth, 2016). Most databases are focused for drug design and include synthetic derivatives of natural products or by-products of drug metabolism that are not true natural compounds and must be manually excluded from our database. In addition, a range of format differences and coding errors mean that most data sources need extensive checking and modification. Errors we have identified and corrected for include:

- (1) Incorrect structures or SMILES strings, including impossible structures and incorrect tautomeric forms
- (2) Errors in structure determination, for example, incorrectly assigned NMR structures
- (3) Natural products containing “exotic” elements (like Si or F) as a result of “feeding” an organism with a precursor of a given biochemical pathway, for example F derivatives of plant hormones or nucleotides, isolated from the natural source upon providing fluorinated derivative (biotransformation products)
- (4) Artificial compounds accumulating in plants and animals and subsequently isolated as potential natural products, for example, many chiral Cl-containing pesticides
- (5) Synthetic drugs, drug-like derivatives of natural products, drug metabolites, or products of semisynthesis.

In addition, most available natural products databases do not contain information on the extent of chemical modification of “polymers of life” (RNA, DNA, proteins), including natural, modified and unmodified, nucleotides and amino acids found only in biopolymers (*e.g.*, post-translational modifications of amino acids in proteins).

All of the above problems motivated us to curate our own natural products database (see Appendix Table A2 and Fig. A2). Our natural products database contains only true natural products—that is, compounds that are a result of natural biochemical processes of a living organism. It is as complete as is practical to make, and it is unique because it has biological sources identified for every entry (*i.e.*, a list of the organism(s) from which the natural product was isolated). Our database also includes representations of chemotypes found in biopolymers. Because of access limitations on much of the material in this database, including terms of access for the authors’ institution to databases under commercial license, this database cannot be made available to the wider community. The authors instead welcome opportunities to collaborate on use of this database.

2.2. Experimental testing of reactivity of N-S compounds with L-cysteine and other common biochemicals

To test our hypothesis presented in Section 3, we react N-S compounds with a set of chemicals that provide a proxy of cellular metabolism. We use an NMR time-course assay to show whether or not an N-S compound reacted with the proxy compounds and, if it did, to assess if an N-S bond was cleaved in the presence of the proxy compounds. Experimental procedures presented in this paper were performed by Organix Inc. (<http://www.organixinc.com>) according to established procedures.

We used N-S compounds belonging to two distinct classes of chemicals, sulfenamides (N-S(II)) and sulfonamides (N-S(VI)).

For the sulfenamide N-S(II) compounds, we chose 1-(methylsulfanyl)pyrrolidin-2-one and 1-(methylsulfanyl)pyrrolidine, which were synthesized as described, for example, in the work of Nelsen *et al.* (1982). The synthesized sulfenamides were made into stock solutions in the solvent CD₃OD/d₆-DMSO (20 μmol/1.5 mL). This solvent system provided the required stability of 1-(methylsulfanyl)pyrrolidine and an improved solubility of the amino acids (see below).

The proxy for cellular metabolism are the following compounds: L-alanine, L-methionine, L-cysteine, L-serine, D-fructose, D-glucose, and D-glucose-6-phosphate. These compounds were obtained from Sigma (<https://www.sigmaaldrich.com>). The procedure involved addition of 20 μmol of each compound of interest to 1.5 mL of the stock solution.

The NMR spectra (16 scans) of the resulting solutions/suspensions were recorded immediately after each was prepared (5–10 min actual time) and every 45–50 min thereafter for a period of 7 h. Finally, the mixture was analyzed after 24 h. The NMR reactivity time-course assay was performed on a JEOL Eclipse 300 NMR spectrometer (JEOL USA Inc.) running at 300 MHz (1H) with the aid of Delta NMR Software (JEOL USA Inc.).

For the sulfonamide N-S(VI) compounds, we chose 1-methanesulfonylpyrrolidin-2-one and 1-methanesulfonylpyrrolidine. These compounds were obtained from Sigma (<https://www.sigmaaldrich.com>). The sulfonamides were prepared into stock solutions as described above. Due to financial limitations, and based on our hypothesis (Section 3), the sulfonamide derivatives were only reacted with L-cysteine. Twenty micromoles of L-cysteine was added to 1.5 mL of the stock solution.

Nuclear magnetic resonance spectra (16 scans) of the resulting suspension were recorded immediately after it was prepared (5–10 min actual time), every 30 min thereafter for a period of 10 h, and after 16 h.

As with the sulfenamides, the NMR reactivity time-course assay was performed on a JEOL Eclipse 300 NMR spectrometer (JEOL USA Inc.) running at 300 MHz (1H) with the aid of Delta NMR Software (JEOL USA Inc.).

3. Hypothesis and Supporting Data-Driven Arguments and Experiments

We start with a statement of our hypothesis (3.1), followed by quantification of occurrence of different structural types of N-S bond-containing compounds in biochemistry (3.2), followed by experimental support for our hypothesis

³www.allmols.org—a repository of potential biosignature gases for astrobiological research and applications for terrestrial biochemistry

(3.3), and concluding with further support of the hypothesis from a literature review (3.4).

3.1. N-S-thiol incompatibility hypothesis

We postulate that life's near exclusion of N-S bond-containing compounds is due to the fact that a large fraction of N-S bond types are incompatible with life's biochemistry. Specifically, we argue that N-S bonds are avoided by life because they are highly reactive in the presence of *thiols*, despite being stable in other aqueous environments. Thiols are -SH-group-containing compounds that form the foundation of metabolism and are common in cells. It would not be an exaggeration to say that almost every protein has thiols (in the form of essential amino acid cysteine). In fact, thiols are ubiquitously present in the cell not only as proteins but also as small-molecule metabolites. Because thiols are central to biochemistry, life cannot easily tolerate N-S bond-containing compounds, which is the reason for their rarity in biochemistry.

N-S(II) bonds, and to a lesser extent N-S(IV) bonds and N-S aromatic compounds, are directly forbidden by their reactivity to thiols. N-S(VI) bonds are themselves unreactive to -SH groups, but their plausible biosynthetic precursors, breakdown or reduction products are likely to include N-S(II) or N-S(IV) compounds, which themselves are reactive to thiols (see Section 4.2.4 for detailed discussion of the reasoning behind N-S(VI) exclusion by life).

Our hypothesis already had hints in previously published works, via scattered literature mentions of the reactivity of N-S bonds with thiols, involving both *in vitro* and *in vivo* studies. An intriguing hint that fueled this observation was a buried result in papers on the activity and mode of action of sulfenamide (N-S(II)) antibiotics (Revell *et al.*, 2007), which briefly stated that sulfenamides and sulfenamide N-S(IV) compounds react exclusively with thiol-group-containing biochemicals, like cysteine and co-enzyme A, leading to bactericidal effects, while not reacting with other components of the cell (Heldreth *et al.*, 2006; Revell *et al.*, 2007). Notably, N-S(VI) sulfonamide analogs of these antibiotics had no effect on the bacterial cells, hinting at a lack of reactivity to thiols and other functional groups (Heldreth *et al.*, 2006).

We searched the literature to assess the reactivity of each of the N-S bond structural classes and types with thiols (Table 2). While systematic studies of reaction rates have not been published, experimental results are reported qualitatively as to whether a compound (N-S bond) survives in the presence of thiols or not.

3.2. Low occurrence of N-S compounds in natural products

There has previously been no systematic study of occurrence of N-S bond-containing natural products. We therefore exhaustively surveyed the occurrence of N-S bonds in the context of the chemical space explored by living organisms (Petkowski *et al.*, 2018). Our goal is to quantify occurrence of N-S bond-containing compounds of different structural types and assess their reactivity with thiols.

From our Database of Natural Products (described in Section 2.1; Appendix Table A2 and Fig. A2), we have found there are only ~100 natural products (NPs) con-

taining the N-S bond, out of a total of ~200,000 known, unique, natural products. While it is recognized that N-S bonds are unusual in natural products, neither the rarity of the bonds nor the occurrence of the many different types of N-S bond-containing compounds has previously been described systematically. We confirmed that the publicly available databases we drew from were not under-representing reports of natural products containing N-S bonds with an exhaustive literature review, reported elsewhere (Petkowski *et al.*, 2018).

To illustrate the discrepancy between the degree of utilization of N-S bonds in synthetic organic chemistry and biochemistry, we have compared the fraction of N-S bond-containing compounds in 200,000 natural products with the fraction of N-S bond-containing compounds in a series of chemical databases containing a diverse collection of chemicals (see Appendix Table A1), representing respectively research, specialist industrial, and large-scale industrial use. N-S bonds are notably underrepresented among natural products as compared to all these collections of chemicals (Figs. 1 and 2).

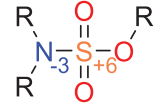
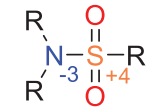
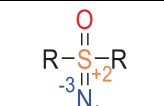
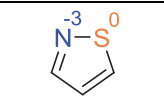
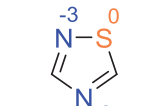
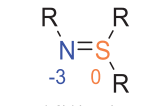
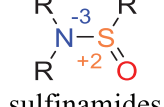
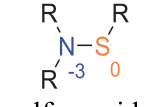
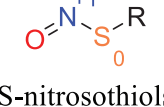
The degree of utilization of N-S chemistry by life heavily depends on the valence state of the sulfur atom. The great majority of natural products containing N-S bonds fall into the category of N-S(VI) (*e.g.*, sulfamates and sulfonamides); ~80% of all N-S bond-containing natural products fall into this category. The remaining ~10% of the N-S bond-containing natural products are dominated by N-S aromatic rings (isothiazoles and thiazoles). N-S(IV) and N-S(II) structural types, for example, sulfilimines and sulfenamides and others, occur one to a few times, constituting only a few known examples (Fig. 3, Table 2 and Appendix Figs. A1 and A5). It is notable that the underrepresentation of N-S bond-containing compounds in biochemistry, compared to industrial chemistry, holds true for all major N-S structural classes and valence states of sulfur atoms (N-S(VI), aromatic N-S, N-S(IV), and N-S(II)) (Fig. 1).

We found that the N-S bond structural types that are the most common in life are the least reactive to thiols (Fig. 4) (Table 2). More specifically, close to 80% of all known N-S bond-containing compounds produced by life belong to the S(VI) sulfamate/sulfonamide which are reported to be unreactive to thiols. The second-largest group is populated by N-S bond aromatic systems which are in general, as a structural class, much less prone to N-S bond cleaving nucleophilic attack by thiols than the sulfenamide N-S(II) bonds or N-S(IV) compounds (Table 2). Compounds belonging to the N-S(II) or N-S(IV) structural families are much more vulnerable to thiolysis and therefore, we propose, are heavily underrepresented in natural products (Table 2).

The occurrence pattern of N-S-bond-containing structural types in biochemistry supports our hypothesis. The N-S bond chemical functionality is very diverse, which allows for vast differences between chemical reactivity of different N-S functional groups. However, what is common for different N-S species in the same structural class is their reactivity to thiols (Fig. 4) (Table 2). This reactivity of different N-S functional groups toward thiols is reflected in the occurrence pattern of N-S structural types in biochemistry (Fig. 4).

One of our main findings is that the most common N-S compounds in natural products (N-S(VI)) are those that are the least reactive to thiols.

TABLE 2. TOTAL OCCURRENCE OF N-S BONDS IN NATURAL PRODUCTS (NPs) AND REACTIVITY TO CELLULAR THIOLS

Structural Class/Type ^a	Total occurrence in NPs	N-S-Thiol Reactivity	Comments	
N-S(VI)	 sulfamates	73	NO	Highly stable to cellular thiol-dependent biotransformation with impressive pharmacological properties (Winum <i>et al.</i> , 2005).
	 sulfonamides	8	NO	Highly stable to cellular thiol-dependent biotransformation with impressive pharmacological properties (Heldreth <i>et al.</i> , 2006; Kalgutkar <i>et al.</i> , 2010; Sivaramakrishnan <i>et al.</i> , 2010; Dahlin <i>et al.</i> 2015)
	 sulfoximines	3	NO	Highly stable to cellular thiol-dependent biotransformation (Frings <i>et al.</i> , 2017)
N-S (aromatic)	 isothiazoles	5	MODERATE	The majority of isothiazole N-S bonds are not affected by cellular thiols (Dalvie <i>et al.</i> , 2002; Teffera <i>et al.</i> , 2010). The degree of thiol reactivity of N-S bond in a ring is dependent on ring modifications and substitutions (<i>e.g.</i> , not fully conjugated rings or isothiazolone derivatives are highly reactive to cellular thiols) (Collier <i>et al.</i> , 1990).
	 thiadiazoles	5	MODERATE	N-S bonds of all three possible isomers of thiadiazoles are highly resistant to thiolysis (Dalvie <i>et al.</i> , 2002; Kalgutkar <i>et al.</i> , 2007). Ring modifications and loss of aromatic character might increase thiol reactivity (Dahlin <i>et al.</i> , 2015).
N-S(IV)	 sulfilimines	2	HIGH	Highly reactive to cellular thiols. Thiolytic is the main route of biotransformation and mode of action of biologically active compounds in this group (Young and Briedis, 1988; Koval, 1990).
	 sulfenamides	3	HIGH	Reactive to cellular thiols although somewhat less than their sulfilimine or sulfenamide counterparts. Thiolytic is the important route of biotransformation and mode of action of biologically active compounds in this group (Heldreth <i>et al.</i> , 2006; Sivaramakrishnan <i>et al.</i> , 2010)
N-S(II)	 sulfenamides	3	HIGH	Highly reactive to cellular thiols. Thiolytic is the main route of biotransformation and mode of action of biologically active compounds in this group (Chiu <i>et al.</i> , 1975; Umetsu <i>et al.</i> , 1980; Fukuto <i>et al.</i> , 1983; Wallace and Zerba, 1989).
	 S-nitrosothiols	3	HIGH	Highly reactive to cellular thiols. Thiolytic is the main route of biotransformation and mode of action of biologically active compounds in this group (Williams, 1996; Dicks <i>et al.</i> , 1998; Wong <i>et al.</i> , 1998; Hu and Chou, 2006).

^aSingle instances (total occurrence=1) of isothiazines, aminopolysulfides, dithiazines, dithiadiazetidines, N-thiosulfoximides N-S Structural Classes/Types among NPs were omitted due to insufficient information on thiol reactivity. Isothiazines, aminopolysulfides, dithiazines, dithiadiazetidines, N-thiosulfoximides are not studied extensively, and chemical reactivity of those N-S systems is largely unknown (reviewed in Petkowski *et al.* [2018]).

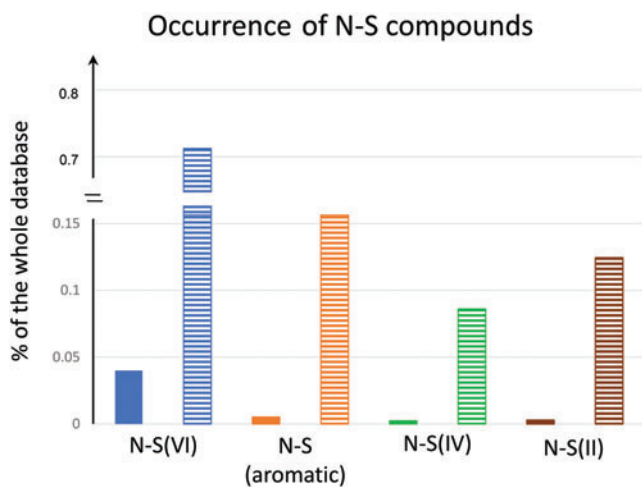


FIG. 1. Fraction of N-S bond-containing compounds known to be produced by life (solid bars) compared to the fraction of N-S bond-containing compounds in synthetic and industrial chemistry (shaded bars). The fraction produced by life comes from our Database of Natural Products (NPdb), and the fraction in synthetic and industrial chemistry comes from the Combined Chemical Dictionary (<http://ccd.chemnetbase.com>). The N-S chemistry is significantly underrepresented in natural products compared to synthetically and industrially used compounds, and holds true for all major N-S structural classes (N-S(VI), aromatic N-S, N-S(IV), and N-S(II)).

3.3. Experimental support for N-S-thiol reactivity

To support our observation that thiol reactivity correlates with N-S bond occurrence in natural products, and to specifically test reactivity to biologically relevant moieties, we perform experiments focused on two extremes in reactivity with thiols.

For the highly reactive end, N-S(II) compounds, we choose sulfenamides 1-(methylsulfanyl)pyrrolidin-2-one and 1-(methylsulfanyl)pyrrolidine. Recall that sulfenamides are very rare in life. For the non-reactive end, N-S(VI) compounds, we choose sulfonamides 1-methanesulfonylpyrrolidin-2-one and 1-methanesulfonylpyrrolidine. Recall that the majority of N-S compounds in natural products are in the N-S(VI) category.

We react the sulfenamide (1-(methylsulfanyl)pyrrolidin-2-one and 1-(methylsulfanyl)pyrrolidine) N-S compounds with seven major cellular metabolites containing all major functional groups (alcohols, amines, carboxylic acids, etc.) utilized by life: L-alanine, L-methionine, L-cysteine, L-serine, D-glucose, D-fructose, and D-glucose-6-phosphate. Only L-cysteine contains a thiol group; the others have non-thiol functional groups. Reactions with this set of compounds constitute a simplified proxy for reactivity with the components of cellular metabolism.

Under our hypothesis of N-S and thiol reactivity, the N-S(II) bond in sulfenamides (1-(methylsulfanyl)pyrrolidin-2-one and 1-(methylsulfanyl)pyrrolidine) will only be destroyed in the reactions with the compound containing the

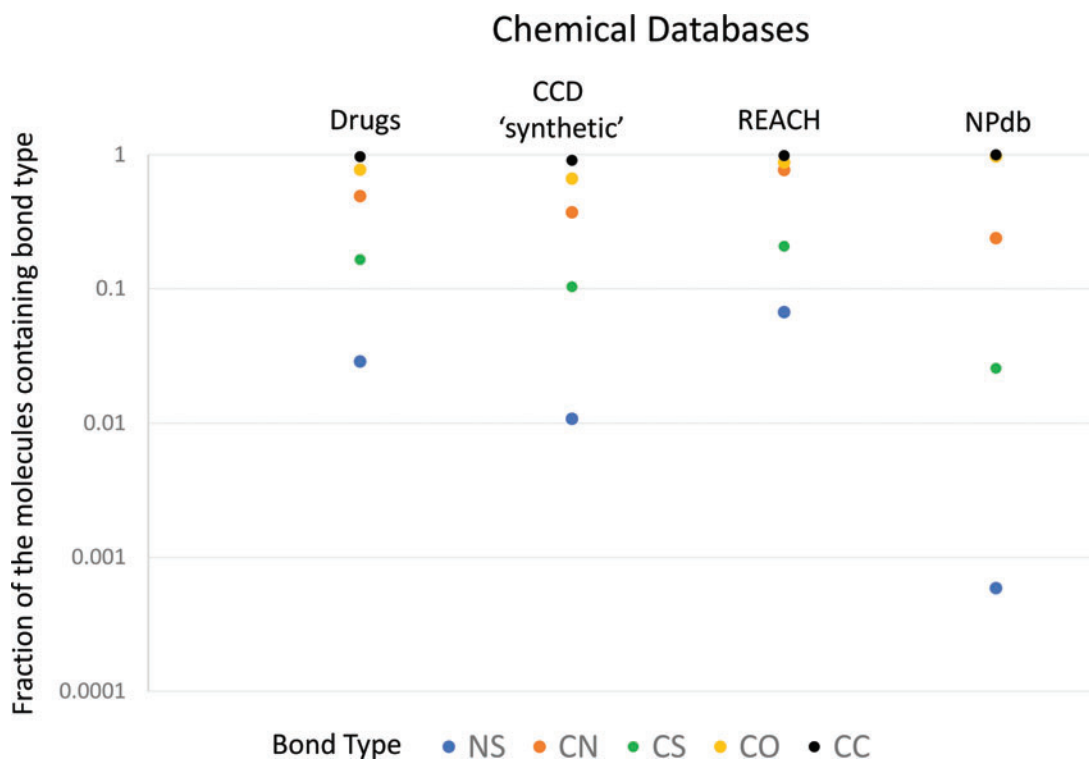


FIG. 2. N-S bond-containing compounds in industry and biochemistry - NPdb (our Database of Natural Products; see Appendix Fig. A2, Tables A1 and A2). The N-S bond is significantly underrepresented in natural products as compared to N-S bond occurrence in the chemicals in various collections of industrial chemicals. C-S bonds are much less underrepresented in the NPdb, and other bond types are of comparable frequency in industrial and NP databases. Note that the point for CO bonds in NPdb lies under the point for CC bonds, and CO bonds are extremely common in natural products.

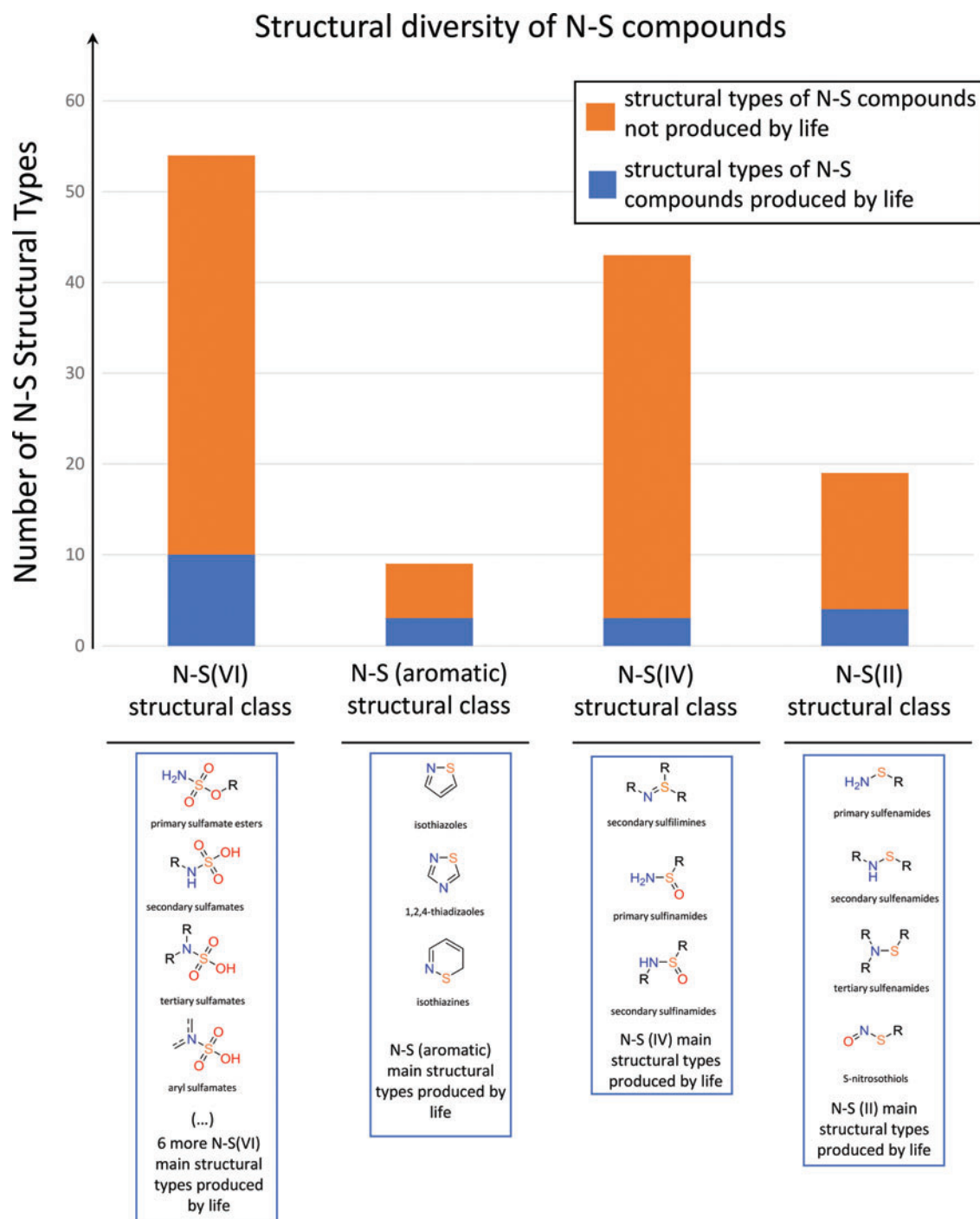


FIG. 3. The scope of chemical space of N-S bond-containing structural types. The bars are structural classes defined by the valence state of sulfur and show the number of different structural types within a class (see the appendix)—not the number of compounds. The bars show there is a much larger diversity of N-S structural types for life to potentially use (orange) than life actually explores (blue). Only the main structural types within N-S(VI), N-S (aromatic), N-S(IV), and N-S(II) structural classes of N-S chemistry are shown and counted. In more detail, only numbers for N-S structural types containing one or two nitrogen atoms and one sulfur atom are shown (with no aliphatic N-N bond-containing structural types shown). The total numbers are likely minima, and further chemical work could extend the number of known, characterized N-S structural types even more. Structural types of N-S compounds produced by life (blue bars) are only a very small fraction of all possible N-S chemical space. See Table 1 and Appendix Fig. A1 for more details.

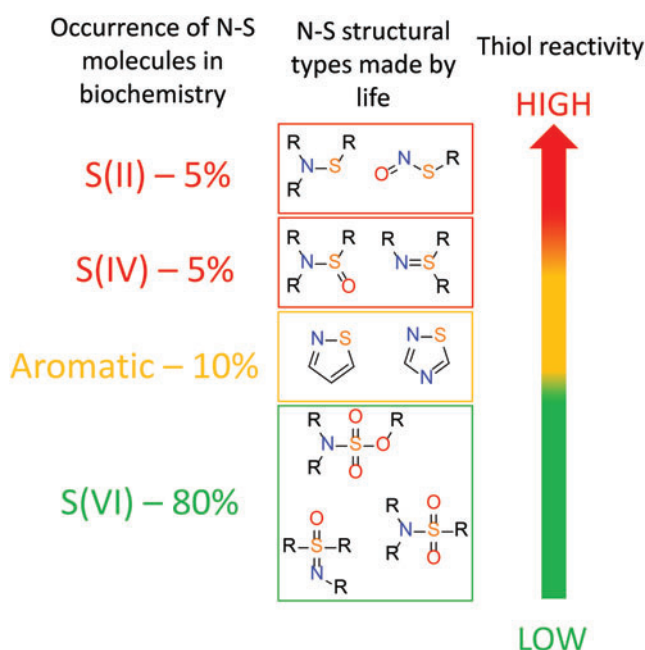


FIG. 4. Schematic representation of the correlation between the occurrence of the N-S compounds in biochemistry and their thiol reactivity. The occurrence of N-S molecules in biochemistry drops as their reactivity to thiols increases. Our N-S-thiol incompatibility hypothesis provides an explanation for the pattern of occurrence of N-S compounds in life, that is, why N-S(II), N-S(IV), and aromatic N-S compounds are rare in life and why they are underrepresented in biochemistry as compared to the number of thiol-unreactive N-S(VI). We note that N-S(VI) overall rarity in biochemistry is not directly explained by thiol reactivity but can be tied to thiol reactivity indirectly (see Section 4.2.4 for detailed discussion of the reasoning behind N-S(VI) exclusion by life).

thiol group (L-cysteine), and, in contrast, will be stable in the presence of the compounds lacking the thiol group.

The N-S(VI) sulfonamides (1-methanesulfonylpyrrolidin-2-one and 1-methanesulfonylpyrrolidine) on the other hand will be stable in the presence of all the tested compounds and will not be reactive to thiol-containing L-cysteine nor the other compounds. We tested that by reacting N-S(VI) sulfonamides with L-cysteine.

An NMR time-course reactivity assay was performed which showed that thiol containing L-cysteine was the only metabolite that reacted directly with sulfenamides (both 1-(methylsulfonyl)pyrrolidin-2-one and 1-(methylsulfonyl)pyrrolidine) resulting in the formation of S-S cysteine conjugate upon cleavage of the N-S bond in 1-(methylsulfonyl)pyrrolidin-2-one and 1-(methylsulfonyl)pyrrolidine (Figs. 5a, 5c, and 6). None of the other tested metabolites showed any reactivity with sulfenamides (see Appendix Figs. A3 and A4). Sulfenamides (1-methanesulfonylpyrrolidin-2-one and 1-methanesulfonylpyrrolidine), on the other hand, as predicted by our thiol-incompatibility hypothesis, did not show any reactivity with L-cysteine, and no cleavage of the N-S bond was observed (Fig. 5b, 5d).

In summary, the results of our experiments (presented in Figs. 5 and 6; see also Appendix Figs. A3 and A4) show that the sulfenamide N-S(II) bond is completely cleaved in the presence of the L-cysteine thiol, and the more oxidized

sulfenamide N-S(VI) variant shows no reaction. The results support our hypothesis, including the occurrence pattern observation that N-S(VI) compounds, although still rare, form the majority of all N-S bond-containing natural products.

3.4. Literature review on N-S-thiol reactivity

The results of the NMR time-course reactivity assay of N-S compounds presented above are consistent with (and complementary to) the thiol reactivity profile of N-S compounds presented in the literature. There are many isolated studies of individual N-S compound's reactivity with thiols that point toward the biologically driven destruction of N-S(II) bond-containing compounds. *In vitro* and *in vivo* experiments show that N-S(II) sulfenamides, in the presence of excess free thiol-containing biochemicals such as cysteine, glutathione, or cysteine-containing proteins, undergo a very rapid (within seconds) cleavage of the N-S bond to yield a disulfide conjugate and a corresponding amine (Nti-Addae, 2008) (see Table 3 for examples of biologically relevant thiolysis reactions of thiol-reactive N-S structural types).

The rapid N-S(II) reactivity toward cellular thiols is the basis of the mechanism of action of many antibiotics (Turos *et al.*, 2002; Turos, 2005; Revell *et al.*, 2007; Prosen *et al.*, 2011; Ramaraju *et al.*, 2012; Shang *et al.*, 2013), pesticides (Chiu *et al.*, 1975; Umetsu *et al.*, 1980; Fukuto *et al.*, 1983; Wallace and Zerba, 1989), and prodrug candidates (Olbe *et al.*, 2003; Hemenway, 2006; Hemenway *et al.*, 2007; Nti-Addae, 2008; Nti-Addae and Stella, 2011; Nti-Addae *et al.*, 2011; Proença *et al.*, 2011; Huttunen *et al.*, 2012). Similarly, the reactivity of N-S(II) compounds to other common functional groups found in the cell (other than thiols) is not detected, showing that the main reactivity of N-S(II) compounds in the cellular environment comes from -SH groups and not from others (alcohols, amines, carboxylic acids, etc.) (Heldreth *et al.*, 2006; Revell *et al.*, 2007).

4. Discussion

Our hypothesis is that N-S bond-containing compounds are rare in life due to an incompatibility with the thiol-containing compounds which are central to life's metabolism. There are a wide range of hypotheses as to why life uses the chemistry that it does, such as proposed explanations for life's use of peptides, phosphates, water, or individual elements (even carbon itself) (Westheimer, 1987; Pace, 2001; Benner *et al.*, 2004; Ball, 2008; Bains, 2014; Reich and Hondal, 2016). However, to our knowledge, our work provides the first example of a chemical explanation of why life *does not* use a chemical class of molecules that is chemically flexible, stable, and has wide chemical and structural functionality. Given the enormous scope of possible chemical space (Kirkpatrick and Ellis, 2004; Raymond *et al.*, 2012; Ruddigkeit *et al.*, 2012), explaining why life does not utilize certain chemical functionalities is as important as explaining why life does prefer certain others.

Our hypothesis that life avoids N-S bond-containing compounds actually has three layers to it. First, N-S bonds are very rare in biochemistry, leading to our general hypothesis that life has apparently made a binary choice to use thiols and not N-S bond-containing compounds (Sections 1 and 3.1). Second, the rarity of the N-S bond-containing functional groups is not uniform and follows a specific occurrence pattern (Fig. 4). The

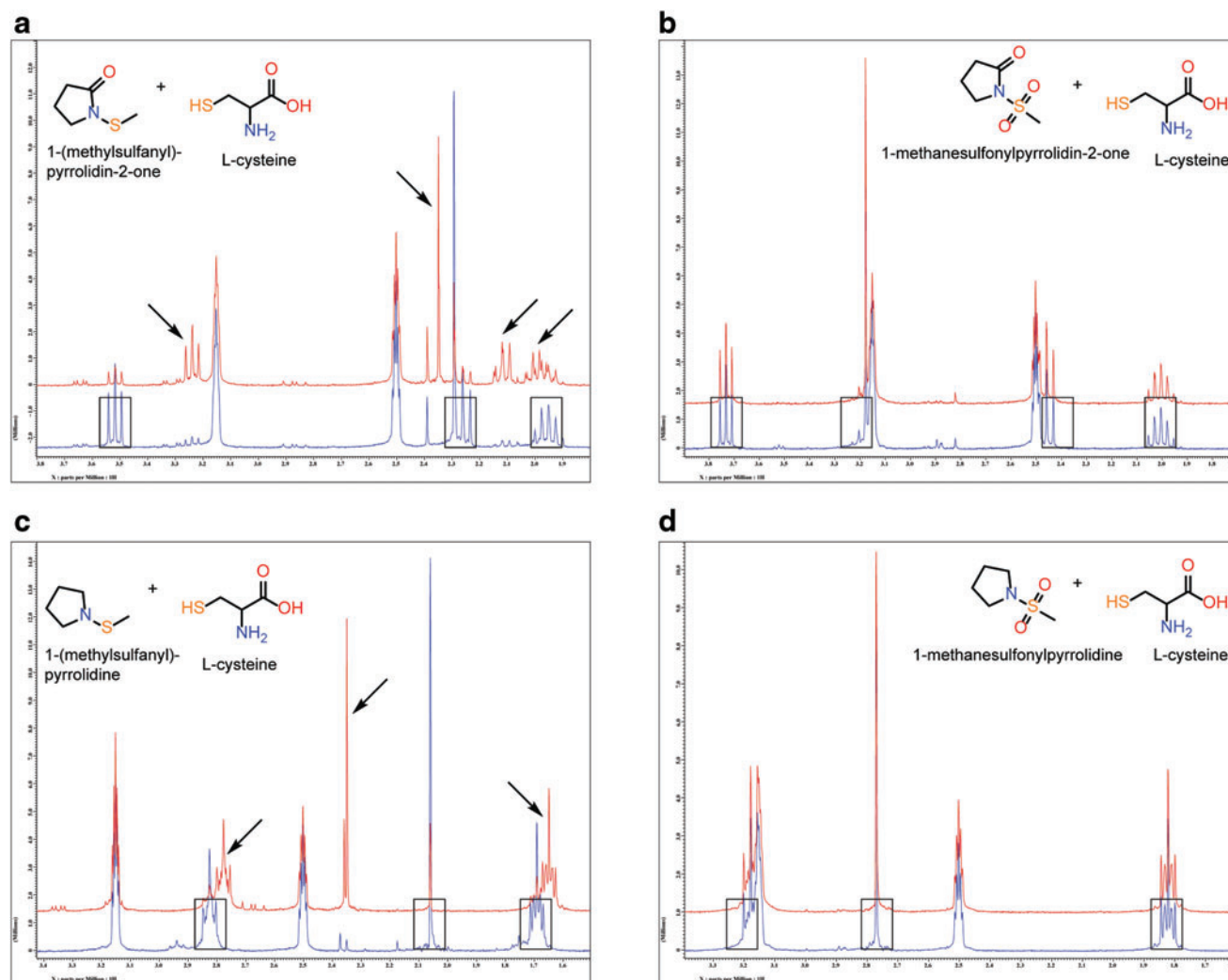


FIG. 5. NMR time-course reactivity assay of N-S compounds (a) 1-(methylsulfanyl)pyrrolidin-2-one (N-S(II)), (b) 1-methanesulfonylpyrrolidin-2-one (N-S(VI)), (c) 1-(methylsulfanyl)pyrrolidine (N-S(II)), and (d) 1-methanesulfonylpyrrolidine (N-S(VI)) with thiol-containing amino acid L-cysteine. The altered shape and position (black arrows) of the peaks show that the N-S bond in the N-S(II) compounds was cleaved in the presence of L-cysteine ([a] and [c], compare $t=0$ [blue] and $t=24$ h [red]), but the N-S bond in the N-S(VI) compounds was not affected by L-cysteine ([b] and [d]). Black rectangles denote the position of signals corresponding to tested N-S compounds. The NMR assay shows that the sulfenamide N-S(II) bond is much more reactive toward thiols than the more oxidized sulfonamide N-S(VI) variant. The N-S-thiol reactivity pattern obtained by the NMR time-course reactivity assay is consistent with the observation that the thiol-unreactive N-S(VI) compounds form the great majority of all N-S bond-containing natural products. Figures A3 and A4 in the appendix show that the N-S bond in the N-S(II) compounds is not cleaved in the presence of the other life-proxy compounds not containing thiol functional groups.

occurrence pattern of N-S compounds in biochemistry depends on the valence state of sulfur in the N-S bond and, importantly, correlates with thiol reactivity (Sections 3.2 and 3.3). Third, life finds a work-around, a way to use highly reactive N-S chemistry even in the presence of thiols. Rather than negating the main hypothesis, the examples support the hypothesis because life uses the N-S-thiol destructive reactivity under highly controlled and regulated conditions for purposeful advantages. The Discussion Section (4.1) focuses on this third layer of our hypothesis and provides arguments as to why there are *any* N-S compounds utilized by life if they are inherently incompatible with biochemistry.

The bulk of the Discussion Section (4.2) details potential challenges to our hypothesis, including potential reasons for the exclusion of the thiol-unreactive N-S(VI) compounds.

The third and final part of our Discussion Section (4.3) is more speculative, in terms of how the rarity of N-S chemistry in life can connect the N-S-thiol incompatibility hypothesis to origins-of-life research.

4.1. Why are there any N-S compounds in biochemistry?

If our N-S-thiol incompatibility hypothesis is correct, we must try to understand why, remarkably, some life actually does ultimately use thiol-reactive N-S bond-containing compounds and chemistry, although in very rare, tightly controlled, and specialized cases.

We argue that while the N-S compounds in biochemistry could be seen as counter-examples to our hypothesis,

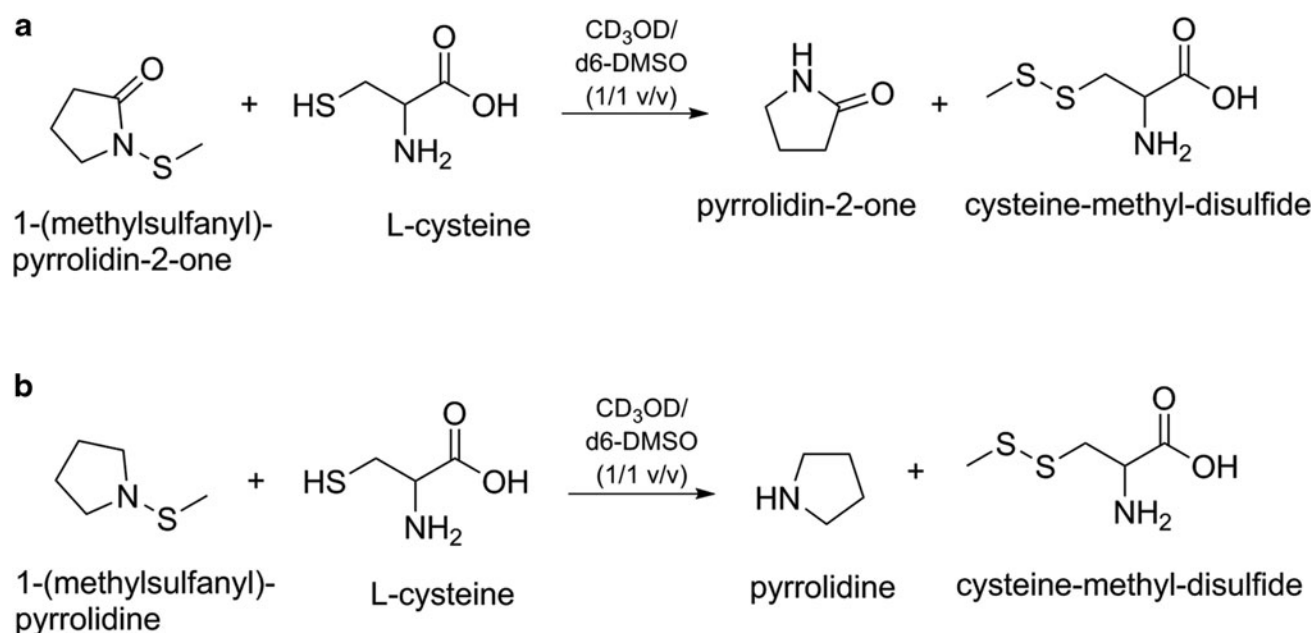


FIG. 6. Proposed thiolysis reaction of N-S(II) compounds: 1-(methylsulfanyl)pyrrolidin-2-one and 1-(methylsulfanyl)pyrrolidine with L-cysteine. Both 1-(methylsulfanyl)pyrrolidin-2-one (**a**) and 1-(methylsulfanyl)pyrrolidine (**b**) undergo cleavage of the N-S bond (as a result of reactivity with thiol group of L-cysteine) and the formation of the cysteine-methyl-disulfide. The chemical reaction explains the results of the NMR reactivity assay in Fig. 5.

such few instances of accommodation of N-S chemistry susceptible to thiolysis support our hypothesis, rather than challenge it. Some organisms have found a way to utilize selected parts of N-S chemistry and developed special means to accommodate certain N-S chemical groups despite their reactivity to cellular thiols (*e.g.*, by burying thiol-reactive N-S bonds deep within protein folds). We present one such example of utilization of N-S-thiol reactivity in detail below. For an overview of other thiol-reactive N-S compounds produced by life, see Table 4. For a detailed, case-by-case description of utilization and accommodation of N-S chemistry by life on Earth, see the work of Petkowski *et al.* (2018).

By far the most intensively studied and, hence, well-understood example of utilization of thiol-reactive N-S chemistry comes from natural cyclic sulfenamide 1,2-thiazolidin-3-one (Fig. 7) identified in several proteins as a post-translational modification involved in a cellular response to oxidative stress, both in bacteria and eukaryotes (Dubbs and Mongkolsuk, 2012; Defelipe *et al.*, 2015). In contrast to thiol-unreactive N-S bonds that are present in small-molecule natural products, 1,2-thiazolidin-3-one exists only in proteins, where it is constrained within the protein fold so as not to be able to react with either other thiols within that protein or with thiols in other proteins in the cell. We note that other natural N-S compounds that have proven thiol reactivity are generally present as post-translational modifications (PTMs) of proteins and are not produced by life as isolated small molecules (Table 4). We believe that limiting the thiol-reactive N-S species only to protein chemical modifications allows for tight regulation of their reactivity with cellular thiols. In fact, in the case of 1,2-thiazolidin-3-one the normally detrimental reactivity of

sulfenamides to thiols is utilized in a controlled environment of the active site of several catalytic cysteine-containing enzymes (Fig. 7).

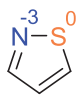
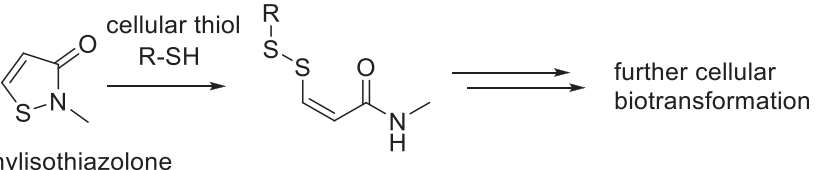
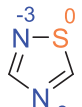
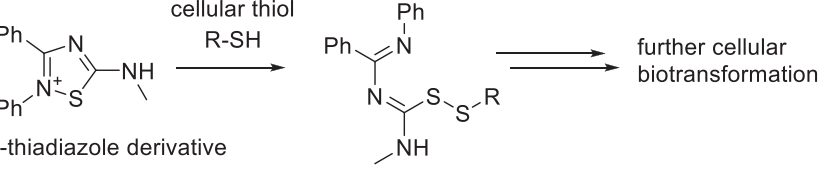
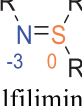
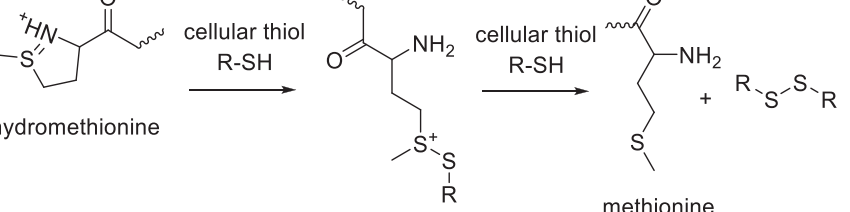
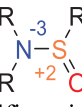
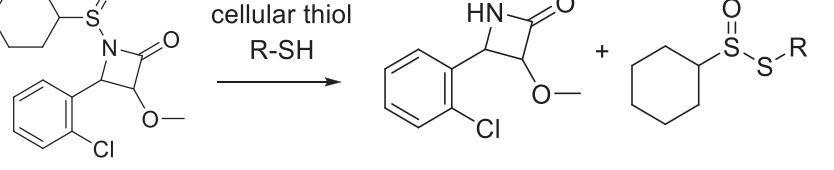
We note that many biochemicals react with thiols, and utilization of thiol-reactive chemistry is one of the pillars of the core metabolism of life. However, in such cases (*e.g.*, widespread utilization of thiol-reactivity in enzyme catalysis) the thiol reactivity is part of a normal, highly regulated, crucial physiological function in the cell. Unless the function of a molecule is to specifically react with thiols, in a controlled fashion, then a thiol-reactive compound is going to be a problem⁴.

Life does not use thiol-reactive N-S chemistry unless there is a very specific need that provides a selective pressure for adapting the chemistry to cope with (and/or use) the reactivity, and which cannot be met by other chemistry without ever more extreme adaptation. For example, a widespread use of S-nitrosothiols⁵ or high conservation of

⁴Heavy metals like Cd and Hg or arsenites are a well-known example of such detrimental thiol-reactivity. The high levels of toxicity of Cd and Hg or arsenites are mediated by their random and facile reactivity with thiols.

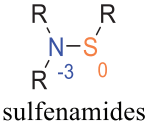
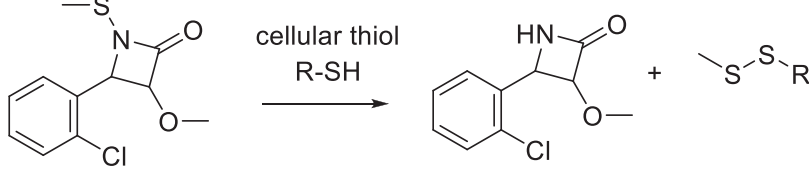
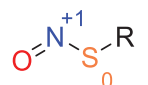
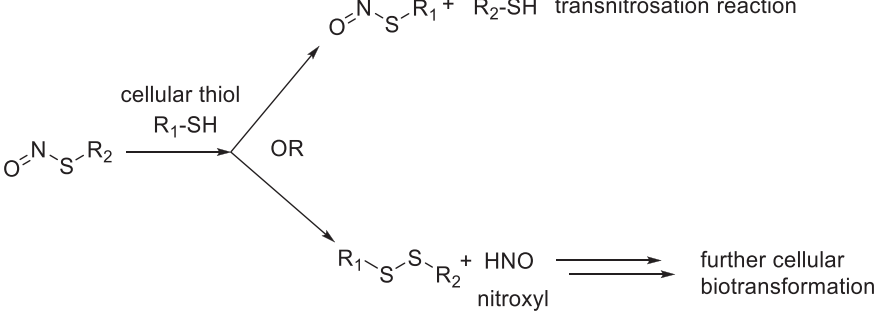
⁵We note that S-nitrosothiols may at first glance appear to be an exception to our hypothesis. S-nitrosothiols are a fascinating example of a specific, unusual structure that has been adapted widely for cellular signaling by many evolutionarily distant organisms. S-nitrosothiols have a very distinct reactivity and a specialized, highly regulated, function in the cell (*i.e.*, as an NO carrier in NO signaling) (Table 3). N-S-thiol reactivity in NO signaling likely evolved early, giving life enough time to adapt to “novel chemistry” of S-nitrosothiols. For a review on S-nitrosylated species produced by life, please see the work of Petkowski *et al.* (2018). For detailed discussion on S-nitrosothiol regulation and reactivity in the cell see, for example, the work of Benhar *et al.* (2009) and Smith and Marletta (2012).

TABLE 3. EXAMPLES OF THIOLYSIS REACTIONS AMONG DIFFERENT N-S BOND-CONTAINING FUNCTIONAL GROUPS BELONGING TO THIOL-REACTIVE STRUCTURAL CLASSES/TYPES (SEE ALSO COMMENTS IN TABLE 2)

Structural Class/Type		Examples of Thiol Reactivity of N-S Compounds
N-S (aromatic)	 isothiazoles	 N-methylisothiazolone biocide Thiol reactivity is crucial for proposed mode of action of isothiazolone biocides (Collier <i>et al.</i> , 1990)
	 thiadiazoles	 1,2,4-thiadiazole derivative Derivative of 1,2,4-thiadiazole, proposed as a histone acetyltransferase Rtt109 inhibitor, rejected as a potential drug target due to promiscuous, unspecific reactivity to cellular thiols (Dahlin <i>et al.</i> , 2015)
N-S(IV)	 sulfilimines	 dehydromethionine methionine Thiol reactivity function of dehydromethionine is unknown, possibly related to immune response (Young and Briedis, 1988; Beal <i>et al.</i> , 2009; Peskin <i>et al.</i> , 2009)
	 sulfinamides	 sulfinamide beta-lactam antibiotic Thiol reactivity is crucial for proposed mode of action of sulfinamide and sulfenamide beta-lactam antibiotics (Heldreth <i>et al.</i> , 2006)

(continued)

TABLE 3. (CONTINUED)

<p>N-S(II)</p>  <p>sulfenamides</p>		 <p>N-thiolated beta-lactam antibiotic</p> <p>Thiol reactivity is crucial for proposed mode of action of sulfenamides and sulfenamide beta-lactam antibiotics (Heldreth <i>et al.</i>, 2006; Revell <i>et al.</i>, 2007)</p>
 <p>S-nitrosothiols</p>		 <p>Two possible reaction pathways of cellular S-nitrosothiols crucial in regulation of nitric oxide signaling and metabolism (Williams, 1996; Dicks <i>et al.</i>, 1998; Wong <i>et al.</i>, 1998; Hu and Chou, 2006; Benhar <i>et al.</i>, 2009)</p>

collagen IV sulfilimines (N=S) in animal species are both examples where life went to extraordinary lengths to accommodate thiol-reactive N-S chemistry to take advantage of unique solutions it could provide (Table 4).

Now that we have reviewed examples of N-S bond-containing compounds in biochemistry and their specific functionalities, for completeness, we round out this subsection with a discussion on the three evolutionary options life has in general to utilize N-S chemistry to achieve useful function. The first, discussed in detail above, is to find a way to get around N-S-thiol incompatibility by developing specific biochemical pathways that allow for accommodation of carefully selected N-S chemistry in a thiol-rich cellular environment.

The second is to adapt a huge number of cellular processes to eliminate dependence on thiols. The second solution would require simultaneous redesign of many pathways in the cell to accommodate selected N-S based functionality (*e.g.*, limiting usage of thiol groups in all metabolic pathways in the cell simultaneously) and seems implausible.

The third is to find an alternative biochemical route to a given function that does not involve utilization of N-S chemistry. This third solution, apparently used by most organisms on Earth, is to develop biochemical functionalities that are analogous to those provided by N-S chemistry, therefore circumventing the N-S-thiol reactivity problem

(*e.g.*, widespread formation of disulfide S-S bonding in proteins instead of potential utilization of sulfenamide (N-S), sulfenamide (N-S(=O)) or sulfilimine (N=S) cross-links). For example, the organism might evolve a solution for a selective challenge relating to protein cross-links that is based on S-S bonds rather than N-S bonds (see Appendix A1 for comparison of biochemical properties of S-S disulfide and N-S(II) sulfenamide cross-links; for review on utilization of N-S cross-links in the cell, see the work of Petkowski *et al.* [2018]).

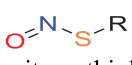

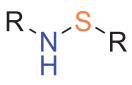
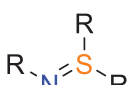
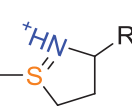
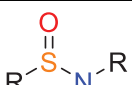
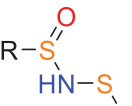
4.2. Alternative explanations and challenges to the N-S-thiol incompatibility hypothesis

There are several alternative explanations why life does not use N-S chemistry. These include thermodynamic explanations (they cost too much energy to make), kinetic explanations (they are unstable), and evolutionary explanations (their synthesis is too hard to evolve). Here we present those challenges and arguments against these alternatives.

4.2.1. N-S bonds are energetically favorable to make.

One may argue that life favors formation of certain chemical bonds for energetic reasons. However, we argue that one can rule out thermodynamic limitations to formation of N-S bonds as a general explanation. We address the energetics of formation and the biosynthesis of N-S bonds in detail below.

TABLE 4. FUNCTIONS OF THIOL-REACTIVE N-S PROTEIN POST-TRANSLATIONAL MODIFICATIONS IN CELLULAR METABOLISM

Protein or Peptide N-S Post-translational Modification	Protein Examples	Function of N-S modification and Function of thiol Reactivity (if known)	
N-S(II)	 <p>S-nitrosothiols</p>	S-nitrosocysteine in many proteins and S-nitrosohomocysteine (HcyNO) in, e.g., hemoglobin.	Regulation of nitric oxide metabolism and redox state of the cell; thiol reactivity mediates release of NO gasotransmitter and regulatory transnitrosation reactions (Broniowska and Hogg, 2012; Smith and Marletta, 2012; Gould <i>et al.</i> , 2013)
	 <p>cyclic sulfenamides</p>	cyclic sulfenamide in, e.g., PTB1B phosphatase and other proteins	Reversible, protective, inactivation of catalytic cysteines during oxidative stress; reactivity with thiol group of glutathione is required for reactivation of enzymes (Fig. 7) (den Hertog <i>et al.</i> , 2005; Sarma and Mugesh, 2007; Sivaramakrishnan <i>et al.</i> , 2010; Dubbs and Mongkolsuk, 2012; Defelipe <i>et al.</i> , 2015)
	 <p>secondary sulfenamides</p>	Lys-Cys cross-link in NemR transcriptional regulator	Regulation of cellular response to reactive chlorine species (RCS); Thiol reactivity ensures that the RCS response is reversible (Gray <i>et al.</i> , 2013, 2015)
N-S(IV)	 <p>secondary sulfilimines</p>	Hyl-Met cross-link in collagen IV	Maintaining structure of Collagen IV in extracellular matrix; Thiol reactivity limited due to inaccessibility of N=S bond to any cellular thiols (Vanacore <i>et al.</i> , 2009; Cummings, 2013; McCall <i>et al.</i> , 2014; Robertson <i>et al.</i> , 2014)
	 <p>cyclic sulfilimines</p>	N-terminal methionine residues of proteins	Formed during immune response as a result of HOCl formation; Reactive to cellular thiols, thiol reactivity function unknown, possibly related to immune response (Young and Briedis, 1988; Beal <i>et al.</i> , 2009; Peskin <i>et al.</i> , 2009)
	 <p>sulfenamides</p>	cysteine sulfenamides in, e.g., calcium binding protein S100A8 and other proteins	Formed during oxidative stress and as a natural consequence of HOCl protein oxidation during inflammation; thiol reactivity is crucial to alleviate oxidative stress conditions (Raftery <i>et al.</i> , 2001; Sivaramakrishnan <i>et al.</i> , 2010; Keceli <i>et al.</i> , 2013; Keceli and Toscano, 2015)
	 <p>N-thiosulfoximides</p>	N-thiosulfoximide cross-link with Hcy in DDAH enzyme	Result of a reaction between thiol group of catalytic cysteine of DDAH and S-nitroso group of HcyNO; Regulatory inactivation of DDAH enzyme; <i>In vivo</i> reversibility of inactivation through reactivity with cellular thiols is uncertain (Knipp <i>et al.</i> , 2005; Frey <i>et al.</i> , 2006; Braun <i>et al.</i> , 2007; Hong and Fast, 2007)

What little is known on the detailed mechanisms of biosynthesis of N-S compounds (Waldman *et al.*, 2017; Petkowski *et al.*, 2018) suggests that there are generally no thermodynamic barriers to making N-S bonds, and no N-S biosynthetic pathway requires direct coupling of ATP hydrolysis to the formation of the N-S bond (Waldman *et al.*, 2017; Petkowski *et al.*, 2018).

The biosynthesis of N-S(II) and N-S(IV) bonds often involves localized formation of the S-X halogenated Cys

or Met intermediate (where X is a halogen donated by HOX hypohalous acids generated by enzymes from peroxidase superfamily). S-X intermediate reacts with nearby amine to form N-S(II) or N-S(VI) cross-links. Such formation of N-S(II) bonds proceeds quite readily (Petkowski *et al.*, 2018).

Similarly, the formation of the majority of N-S(VI) bonds is quite robust and is likely to be catalyzed by amine sulfotransferases (EC 2.8.2.3) that utilize 3'-phospho-5'-adenylyl

Active site of PTPB1 phosphatase

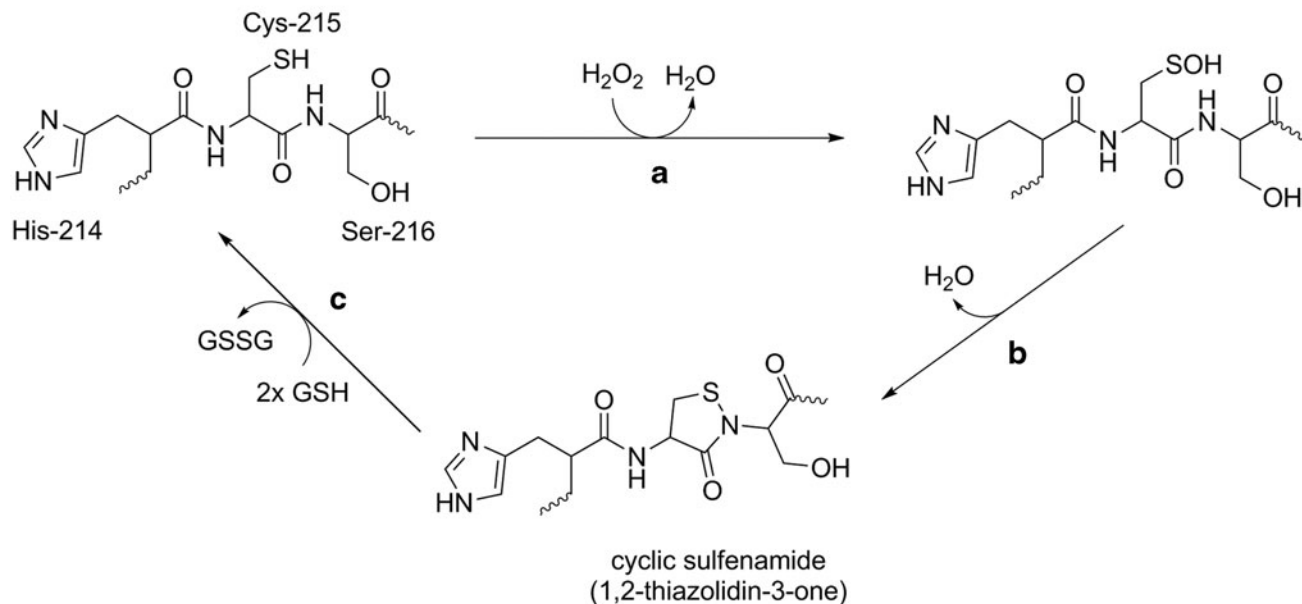


FIG. 7. Example of utilization of thiol-reactive N-S chemistry by life: role of cyclic sulfenamide post-translational modification in cellular oxidative stress response (ROS). Cyclic sulfenamide (1,2-thiazolidin-3-one) is formed as a protective measure from irreversible oxidation of catalytic cysteine residues (*e.g.*, Cys-215 in PTPB1 phosphatase active site). (a) Exposure of the catalytic Cys-215 of PTPB1 to H_2O_2 results in formation of a sulfenic acid (R-SOH), which undergoes (b) intramolecular cyclization with neighboring amino group, resulting in formation of a cyclic sulfenamide. (c) Reduction of 1,2-thiazolidin-3-one is possible through controlled reaction with cellular glutathione, first by formation of a mixed disulfide which is subsequently reduced by glutaredoxin.

sulfate (PAPS) as a sulfonylation donor (Mueller and Shafiqat, 2013; Waldman *et al.*, 2017).

Even if in certain cases the formation of N-S bonds was postulated to be difficult from the energetics standpoint, such energetic barriers are not impossible to overcome. Formation of five-membered cyclic sulfenamide rings (*e.g.*, in PTP1B) (Fig. 7) requires high energy which is associated with the breakage of the S-O bond of the intermediate sulfenic acid (Sarma and Muges, 2007). However, in reality the formation of cyclic sulfenamide is favorable to occur due to a special arrangement of the neighboring amino acids in the active site of PTP1B (*e.g.*, via H-bonding between the amide oxygen and His-214 in PTP1B). This special arrangement of residues results in enhancement of the electrophilicity of the sulfenic acid sulfur or the nucleophilicity of the amide nitrogen and allows for the easy breakage of the S-O bond in the sulfenic acid and the completion of the reaction (Sarma and Muges, 2007; Nagy and Winterbourn, 2010).

To support the above observations, we compute enthalpy of formation of simple chemical structures containing N-S bonds as individual molecules using semi-empirical quantum mechanical calculations (reviewed in Stewart [2007]), specifically the AM1 (Dewar *et al.*, 1985), PM3 (Stewart, 1989), and MNDO (Dewar and Thiel, 1977) methods implemented in GAMESS (Schmidt *et al.*, 1993). As a control and validation of our approach, we calculated energy of formation of analogous N-C bonds or C-S bonds (which are commonly used by life). Our results suggest that, within the accuracy of this approach, the formation of N-S bonds is no less energetically unfavorable than the formation of N-C, C-

S, or C-C bonds (see Appendix Fig. A6). We therefore conclude that the underrepresentation of N-S bonds in natural products is generally not due to the energy required to synthesize the various N-S bonds; in fact, their formation can be quite energetically favorable.

4.2.2. N-S bond-containing compounds are stable in the aqueous environment of the cell. The majority of life's chemical reactions take place in the aqueous environment of the cell; therefore, any assessment of life's utilization of any chemistry has to take into account chemical stability and reactivity to water.

Many chemical functional groups are excluded by life on the basis of very high reactivity to water. One of the most clear-cut examples of such a scenario are arsenate esters, including tri-, di-, and mono-esters, and pyro-arsenates. They react with water in many orders of magnitude faster than their phosphate counterparts and are unstable even in conditions where water is present at low concentrations. Thus, the inherent chemical properties of arsenate esters make it exceedingly unlikely for life to utilize organo-arsenic chemistry in ways that are analogous to phosphates (Tawfik and Viola, 2011; Benner *et al.*, 2013). The example of hydrolytically unstable arsenic esters illustrates that an aqueous environment will constrain available chemical space of life. However, there are other examples of chemistry, of which N-S chemistry is one, for which there is no such obvious constraint. We discuss the hydrolytic stability of N-S compounds next.

Calculating kinetic stability to hydrolysis for a wide range of complex molecules is not practical, and as mentioned

above, there is very limited experimental kinetic data on N-S bond reactivity. Instead, we consider the fact that a wide variety of N-S bond-containing compounds are used in aqueous environments. We note that just like with any other class of chemical compounds the N-S chemical reactivity and hydrolytic stability depend largely on the larger structural context of the molecule, in the case of the N-S compounds on the nature of the substituents on nitrogen and sulfur. Many both aryl and aliphatic N-S compounds of all valence states of sulfur find their uses in pharmacology (e.g., sulfonamide antibiotics), agriculture (e.g., sulfenamides benfuracarb and furathiocarb), and medicine (e.g., isothiazole perospirone, an antipsychotic drug), which implies sufficient water stability for them to be considered as potential biological agents in the first place. Many reactive N-S compounds such as N-S(II) sulfenamides or N-S(IV) sulfilimines are utilized as pesticides or herbicides in agriculture or considered as medicines (Beck, 1975; Casini *et al.*, 2002; Bland *et al.*, 2014; Zhou *et al.*, 2014a, 2014b, 2014c) and are water stable; in fact, their mode of action (if known) often involves rapid thiolysis in the presence of cellular -SH groups (Wallace and Zerba, 1989; Revell *et al.*, 2007; Nti-Addae, 2008) (Tables 1 and 3), which implies that they are stable in the aqueous biological environment until they encounter thiols. Similarly, N-S molecules belonging to structural classes produced by life by definition have to be water stable.

4.2.3. N-S chemistry is not inherently difficult to evolve. One may also argue that N-S chemistry is difficult to evolve and so has only evolved once and is used by a limited range of organisms. An argument of this type is used to explain evolutionary limitations of oxygenic photosynthesis or nitrogen fixation. For each of those two processes, only one biochemical mechanism was described, evolving in a single group of organisms (Raymond *et al.*, 2004; Blankenship, 2010; Bains and Schulze-Makuch, 2016). Such an argument does not explain the paucity of N-S chemistry in biochemistry. All major branches of life make N-S compounds, and most classes of N-S bond are made by more than one kingdom of life (see Appendix Fig. A5). What little is known of the biosynthesis of N-S compounds confirms different pathways and enzymes in different organisms, even sometimes different pathways leading to very similar compounds (reviewed in the work of Petkowski *et al.* [2018]). This suggests that N-S chemistry is not difficult to evolve; the evolutionary barrier is evolving mechanisms to protect the cell from the toxic result of that chemistry.

4.2.4. Possible reasons for the exclusion of thiol-unreactive N-S(VI) compounds. Our N-S-thiol incompatibility hypothesis does not directly explain why the thiol unreactive N-S(VI) compounds are rare in life. Even if N-S(VI) compounds dominate life's N-S chemistry repertoire and the pattern of occurrence of N-S compounds in biochemistry correlates with N-S reactivity to cellular thiols, the N-S(VI) compounds are still rare (numbering only around 80 compounds in ~100 total N-S compounds) (Fig. 4; Table 2). There can be reasons for their rarity, which could be indirectly connected to thiol metabolism of the cell or completely independent from it.

We present our reasoning for the N-S(VI) rarity in biochemistry below.

Biosynthetic and metabolic limitations contribute to the exclusion of N-S(VI) compounds from biochemistry: We have shown that N-S(II) bonds, N-S(IV) bonds, and to a lesser extent N-S aromatic compounds are directly forbidden by their reactivity to thiols. N-S(VI) bonds on the other hand are themselves unreactive to S-H groups, but their plausible biosynthetic precursors, breakdown or reduction products may be N-S(II) or N-S(IV) compounds, and themselves reactive to thiols. The reactivity of N-S(II) or N-S(IV) compounds (e.g., sulfenamides) with thiols likely limits the number of possible safe routes for biosynthesis of any N-S bond-containing compounds, making it exceedingly difficult for life to explore the N-S bond chemistry in general. Formation of a sulfenamide N-S(II) bond at any step in the biosynthetic pathway (even as a transient intermediate) exposes the cell to the detrimental reactivity of N-S(II) group to nearby thiols. Hence, if the cell needs to synthesize an N-S bond-containing compound, it has to do it without the N-S thiol-reactive intermediates, either by first oxidizing the sulfur atom to thiol-unreactive S(VI) (before formation of the N-S bond, e.g., in sulfamates [Alnouti and Klaassen, 2006; Sanchez and Kauffman, 2010; Carlsson and Kjellén, 2012]), through formation of less-reactive aromatic systems (e.g., in isothiazoles), or highly regulated reaction intermediate channeling (Garcia *et al.*, 2012). Any unspecific, undirected, and robust chemical reaction that interferes with essential functions of thiols in the living cell has to be either tightly regulated or, if that is impossible (e.g., due to robustness of the reaction), excluded entirely. Such unregulated reactivity will lead to sulfenamides, and by extension other N-S bond-containing molecules derived from them, to be exceedingly rare in natural products.

Overrepresentation of toxic compounds among N-S natural products: A preliminary statistical analysis of biological effects of natural compounds collected in our Natural Product Database led to an observation that the biological effect of ~62% of all N-S bond-containing natural products, that have been tested, is widely reported to be toxic against other organisms, regardless of whether their thiol reactivity is known or suspected. The toxic natural N-S compounds include broad-spectrum antibiotics, neurotoxins, herbicides, fungicides, or others (reviewed in the work of Petkowski *et al.* [2018]). It is likely that the organisms making N-S compounds are using them as broad-spectrum toxins, similar to other molecules that are highly reactive biocides, such as cyanogen bromide, superoxide, oxirane derivatives, and others (Drahl *et al.*, 2005; Gersch *et al.*, 2012; Vanelslander *et al.*, 2012). The broad biocidal properties of natural N-S compounds are also mirrored by biological activities of many synthetic N-S compounds that are commonly used in many industries, for example as wood preservatives, fungicides in paints, acaricides, insecticides, pesticides, herbicides, and antibiotics in medicine (Adams *et al.*, 1960; Kalgutkar *et al.*, 2010; Bektas and Eulgem, 2014; Kalogirou and Koutentis, 2014) (Table 1). Thus, our statistical observation on toxicity of N-S compounds supports the inherent incompatibility of N-S bond chemistry with life.

Connection between cellular oxidative stress and the exclusion of the N-S compounds from biochemistry: Formation of

N-S(VI) and other N-S compounds as a result of aberrant redox metabolism in the cell was suggested as one of the markers of detrimental effects of cellular oxidative stress (Raftery *et al.*, 2001; Nagy and Winterbourn, 2010).

Redox signaling, of which N-S chemistry is an integral part, is strictly dependent on a tightly controlled oxidation of thiol-containing cysteine residues (Nagy and Winterbourn, 2010; Paulsen and Carroll, 2013; Go *et al.*, 2015; Poole, 2015). Release of the highly oxidizing hypohalous acids, for example during immune response, results in the perturbation of the redox signaling, which in turn leads to the damaging effects of the oxidative stress. Interestingly, the biosynthesis of N-S(II) and N-S(IV) bonds often requires highly controlled and localized formation of the hypohalous acids (HOCl or HOBr) (Petkowski *et al.*, 2018). Therefore, any uncontrollable exposure to hypohalous acids often leads to an unspecific formation of a variety of N-S post-translational modifications in proteins and peptides. Such unspecific N-S post-translational modifications were linked to detrimental toxic effects of oxidative stress (Raftery *et al.*, 2001; Nagy and Winterbourn, 2010). Moreover, under prolonged oxidative stress conditions, for example, as a result of HOCl release during inflammation, various N-S species can undergo further oxidation to N-S(VI) sulfenamides. Formation of sulfenamides in the cell is deemed largely irreversible and is often viewed as a hallmark of severe oxidative damage and disease (Nagy and Winterbourn, 2010; Paulsen and Carroll, 2013; Go *et al.*, 2015).

We can therefore speculate that the connection between N-S chemistry and the detrimental effects of oxidative stress, including the formation of largely irreversible N-S(VI) species, might contribute to the exclusion of the N-S(VI) chemistry from life.

4.3. Speculation on life's exploration of chemical space

We propose that N-S-thiol incompatibility is not unique, and other incompatibilities exist between potential chemistries and the actual biochemistry of life. The incompatibility between N-S bonds and thiols, resulting in the effective near exclusion of N-S chemistry from biochemistry, is an example of a general negative constraint on how life explores chemical space. Such incompatibilities (like thiols vs. N-S) do not represent inevitable exclusions of chemical classes from a biochemistry. Rather, they represent mutually exclusive options; a self-consistent biochemistry can be built either with abundant thiols or with abundant sulfenamides (or neither), but not both. Thiols and sulfenamides represent two “islands of stability,” separated by a region where both are ubiquitously present in the cell, a situation that that is not compatible with the stability requirements for life. If life on Earth originated in an environment rich in thiols, then exclusion of N-S chemistry from terrestrial biochemistry follows from this origin. Life originating in another environment, either a “shadow biosphere” (Davies, 2011) on Earth or life on another world, could have taken a different path. We speculate more widely; there may be other mutually exclusive sets of chemical options, which wider systematic study of the chemical groups in Earth's biochemistry could reveal.

5. Summary and Conclusions

The exclusion of N-S bonds from biochemicals is surprising. N-S bonds are stable, provide a wide range of biologically relevant functionality, and in instances when N-S bonds are utilized by life on Earth their synthesis has evolved independently on multiple occasions. Why then are N-S containing molecules so rare in biochemistry?

We postulate that there are fundamental reasons why only a small fraction of the possible N-S bond chemistry is utilized by life and that biases in life's preference for a specific subset of N-S bond chemistry are rooted in the chemistry of the N-S bond itself, in particular in its reactivity toward cellular thiols. N-S-thiol incompatibility hypothesis makes testable predictions regarding which N-S bond chemicals are likely to be more common in natural products, and why. These predictions are borne out by the observed occurrence of N-S bonds in the space of known natural product chemistry, where N-S structural types reactive to thiol groups are much rarer than those that are resistant to thiolysis.

We hypothesize that life's emphasis on utilization of certain chemistry (like that of SH thiol chemistry) made it very hard to explore other chemical alternatives, essentially making significant areas of chemical space chemically incompatible with life's biochemistry (*e.g.*, extremely reactive toward each other). These mutual chemical incompatibilities put clear constraints on what chemistry life can explore and utilize for useful function. This in turn leads to a formation of a self-contained and self-compatible area of chemical space that is capable of forming a stable biochemistry.

Our N-S result is the first step in a larger effort. We aim to identify more areas of chemical space that life avoids and try to again explain why each case is so, with the hope of providing insight into life's evolution through chemical space and therefore into the origin and early evolution of life.

Appendix A.1. Comparison Between S-S (Disulfide) and N-S (Sulfenamide) Bonds and Their Cellular Reactivity

N-S(II) sulfenamide bonds are in many ways analogous to disulfide bridges (S-S bond) that are crucial for proper folding and stabilization of tertiary structures of peptides and proteins. In addition to maintaining protein structures, disulfide bonds are crucial for mediating specific protein activities including catalytic activity in enzymes. Both bacteria and eukaryotes rely on specific enzymes (disulfide isomerase in eukaryotes and thioredoxin-like enzymes in bacteria) to ensure the correct disulfide bond formation state during protein folding. It is important to note that correct disulfide bond formation is also crucial during the biosynthesis of many secondary (nonprotein) metabolites (there are approx. 1000 known natural products containing an S-S bond, not counting cysteine disulfide bonds in proteins and peptides), where S-S bonds are essential for their biological activity, often through the generation of reactive oxygen species and by inactivating target proteins through reactivity with catalytic cysteine residues via thiol-disulfide exchange reaction, a chemical activity not dissimilar to the mechanism of action of many N-S(II) sulfenamide antibiotics.

TABLE A1. DATABASES COLLECTING SYNTHETIC AND INDUSTRIAL CHEMICALS

Abbreviation	Full name	Download	Accessed	Comments
Drugs	DrugBank	https://www.drugbank.ca	May 2017	
CCD	Combined Chemical Dictionary	http://ccd.chemnetbase.com	November 2017	CCD with removed natural products. Only synthetic and industrial chemicals
REACH	European Chemicals Agency Registered Substances database	https://www.echa.europa.eu/information-on-chemicals/registered-substances	November 2017	Only substances with defined chemical formula counted
NPdb	MIT Natural Products Database		<>	manually curated

The reaction of N-S(II) bonds with thiols is in many ways analogous to biologically important thiol–disulfide exchange reaction. However, the reactivity of sulfenamide N-S(II) bonds to deprotonated thiol (thiolate (S^-)) and protonated thiol (SH) nucleophiles is generally greater than those of S-S disulfide bonds. N-S(II) bonds react not only with strong nucleophiles like thiolates (S^-) but also with protonated thiols (SH) and thioacids (Craine and Raban, 1989; Koval, 1996b). Only thiolates (S^-), not thiols (SH), can attack S-S disulfide bonds efficiently, and as a result, thiol–disulfide exchange reaction is inhibited at low pH (generally below pH 7–8). At lower pH, the protonated thiol (SH) form is favored relative to the deprotonated thiolate (S^-) form. Thiols (SH) are much weaker nucleophiles than thiolates (S^-), and (contrary to N-S(II) sulfenamides) direct reaction between protonated SH thiols and disulfides has not been observed (Singh and Whitesides, 2010). In addition, spontaneous thiol–disulfide exchange is rather

slow, and enzyme catalysis is often employed to accelerate these reactions *in vivo* (Nagy, 2013). This difference in reactivity is crucial, as it allows for tighter regulation and control over the reactivity of thiol–disulfide system, for example by limiting or facilitating the formation of thiolates (S^-) in the cell (see, *e.g.*, reviewed in Nagy [2013] and Nagy and Winterbourn [2010]); such control of reactivity is more difficult to achieve for much more reactive thiol–N-S system.

In fact, the difference in reactivity efficiency of protonated thiols (SH) toward S-S and N-S bonds is utilized in proteomic research to specifically identify endogenous N-S(II) sulfenamides, like cyclic sulfenamides, isothiazolidin-3-ones (see Section 4.1, Fig. 7). The method is based on a series of small-molecule thiols (SH) that react with N-S sulfenamides (resulting in N-S bond cleavage and formation of S-S conjugates) but not S-S disulfides (*i.e.*, without any cross-reactivity with disulfides) (Shiau *et al.*, 2006).

TABLE A2. A LIST OF SOURCES FOR NATURAL PRODUCTS COLLECTED FOR OUR DATABASE (NPDB)^a

Source Name	Number of entries	Description	Web page	Reference
Dictionary of Natural Products	~185000		http://dnp.chemnetbase.com	(DNP, 2016)
UNPD	~210000			(Gu <i>et al.</i> , 2013)
Other Sources:	~70000			
KnapSack	~52000	A Comprehensive Species-Metabolite Relationship Database	http://kanaya.naist.jp/KNapSack	(Nakamura <i>et al.</i> , 2015)
Pherobase	~6000	Database of Pheromones and Semiochemicals	http://www.pherobase.com	(El-Sayed, 2016)
mVOC	~800	Volatile Natural Products	http://bioinformatics.charite.de/mvoc	(Lemfack <i>et al.</i> , 2014)
Streptomedb	~4000	Natural Products isolated from various species of <i>Streptomyces</i>		(Klementz <i>et al.</i> , 2016)
Mitishamba	~1000	Plant Natural Products from Kenya	http://mitishamba.uonbi.ac.ke	(Derese <i>et al.</i> , 2015)
Herbal Ingredients' Targets	~500	Chinese Natural Herbal Medicine Active ingredients		(Ye <i>et al.</i> , 2011)
NuBBE	~1500	Natural Products from Brazil	http://nubbe.iq.unesp.br/nubbeDB.html	(Valli <i>et al.</i> , 2013)
SANC	~600	South African Natural Product Database	https://sancdb.rubi.ru.ac.za	(Hatherley <i>et al.</i> , 2015)
ASM	~600	Gaseous Molecules produced by Life	http://www.allmols.org	(Seager <i>et al.</i> , 2016)
Literature collection	~3000	Natural Products collected by manual literature search		

^aWe note that natural products from newer natural product databases, published in 2017 and 2018, *e.g.*, NANPDB (Ntie-Kang *et al.*, 2017), NPCARE (Choi *et al.*, 2017), NPASS (Zeng *et al.*, 2018), as well as from tgsc (<http://www.thegoodscentcompany.com>) are not counted in Table A2 or in Figure A2 and are currently in the process of manual curation.

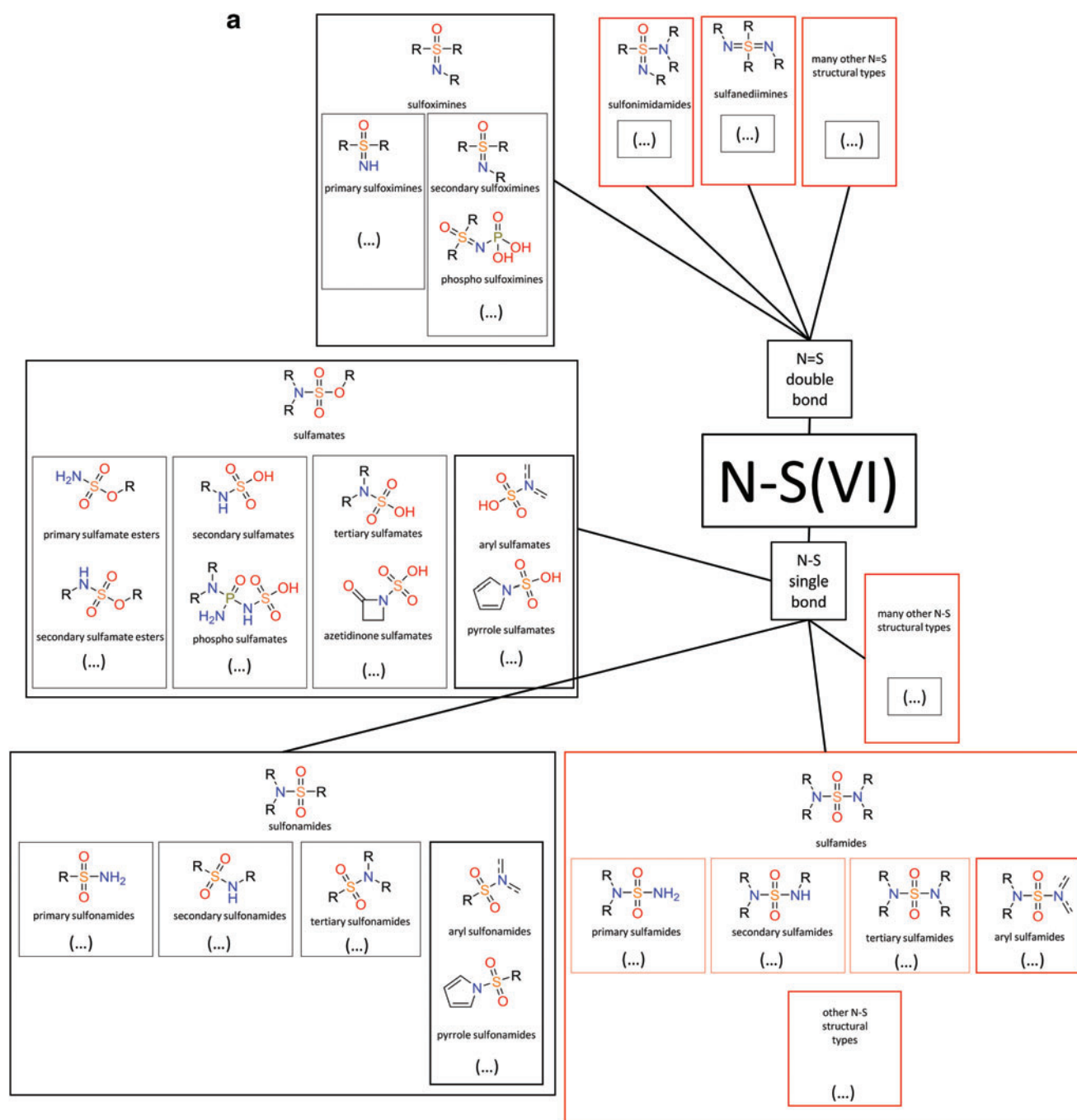


FIG. A1a. Example of structural diversity of N-S compounds. An overview of selected structural types within (a) N-S(VI), (b) N-S(IV), (c) N-S (aromatic), and (d) N-S(II) structural classes of N-S bond-containing compounds. Structural types of compounds produced by life (black contoured boxes) and examples of industrially relevant structural types of N-S compounds not produced by life (red contoured boxes) are shown. For detailed examples of industrial utilization of N-S compounds belonging to different N-S structural types, see Table 1.

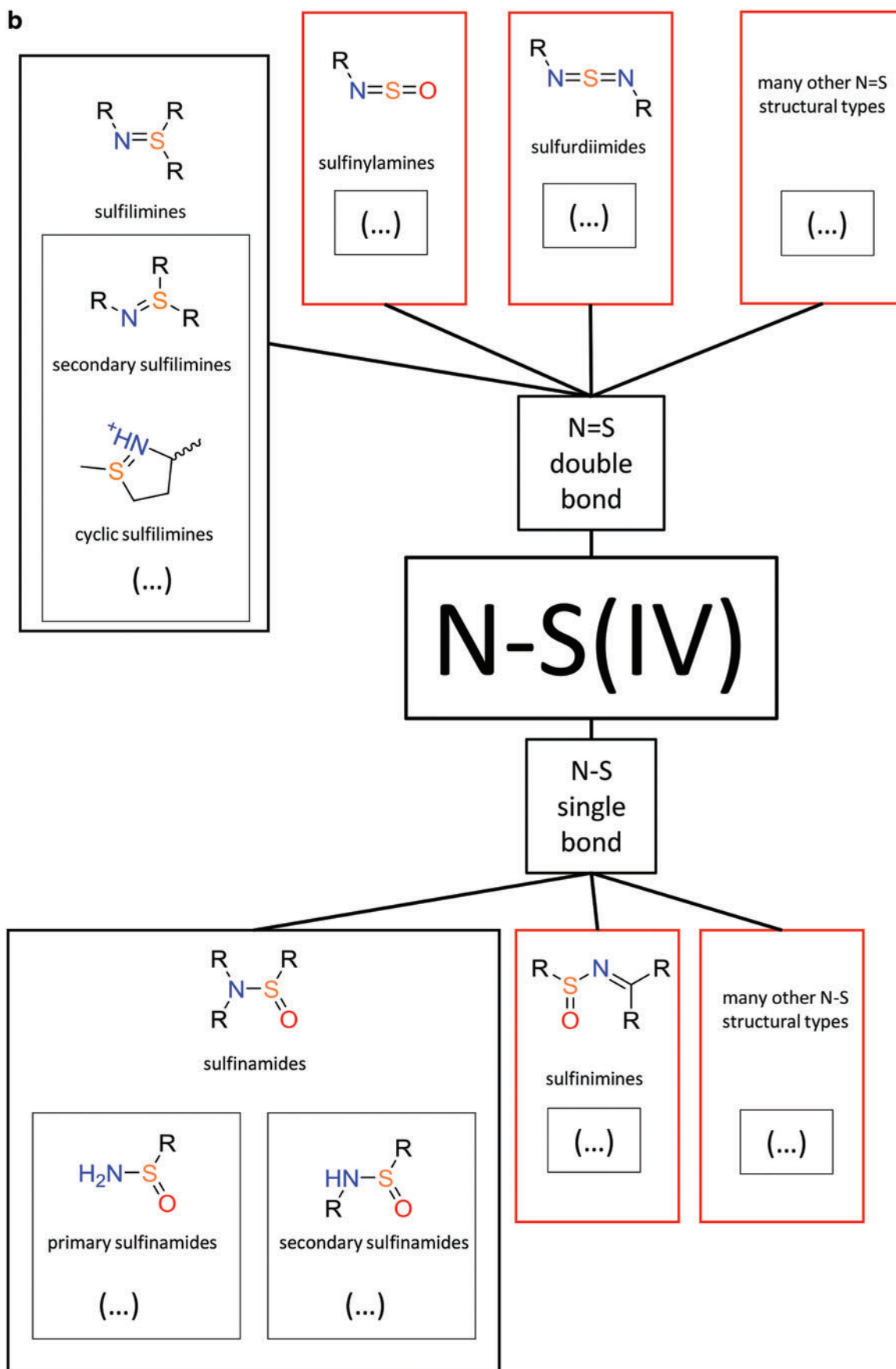
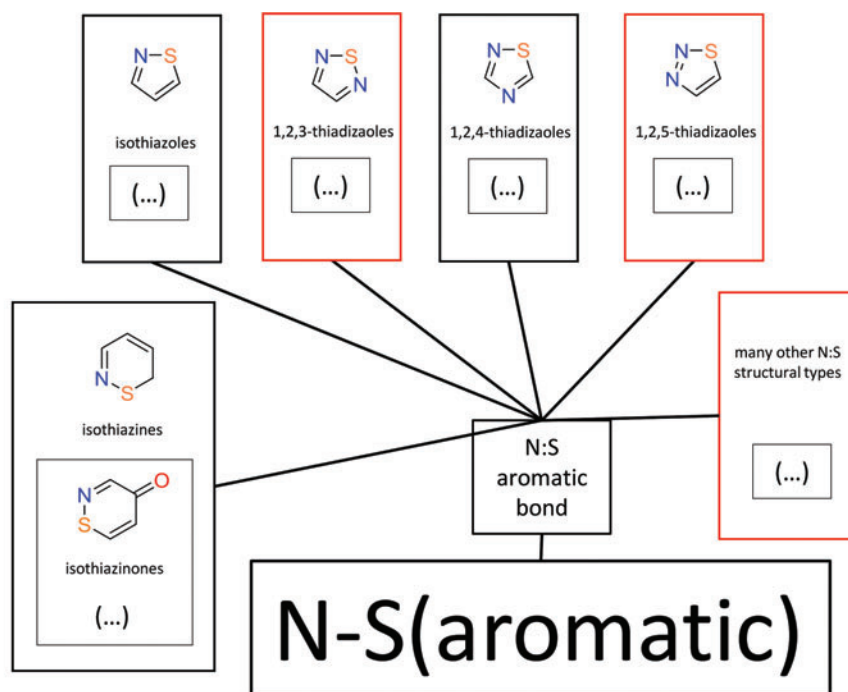


FIG. A1b. (Continued).

c



d

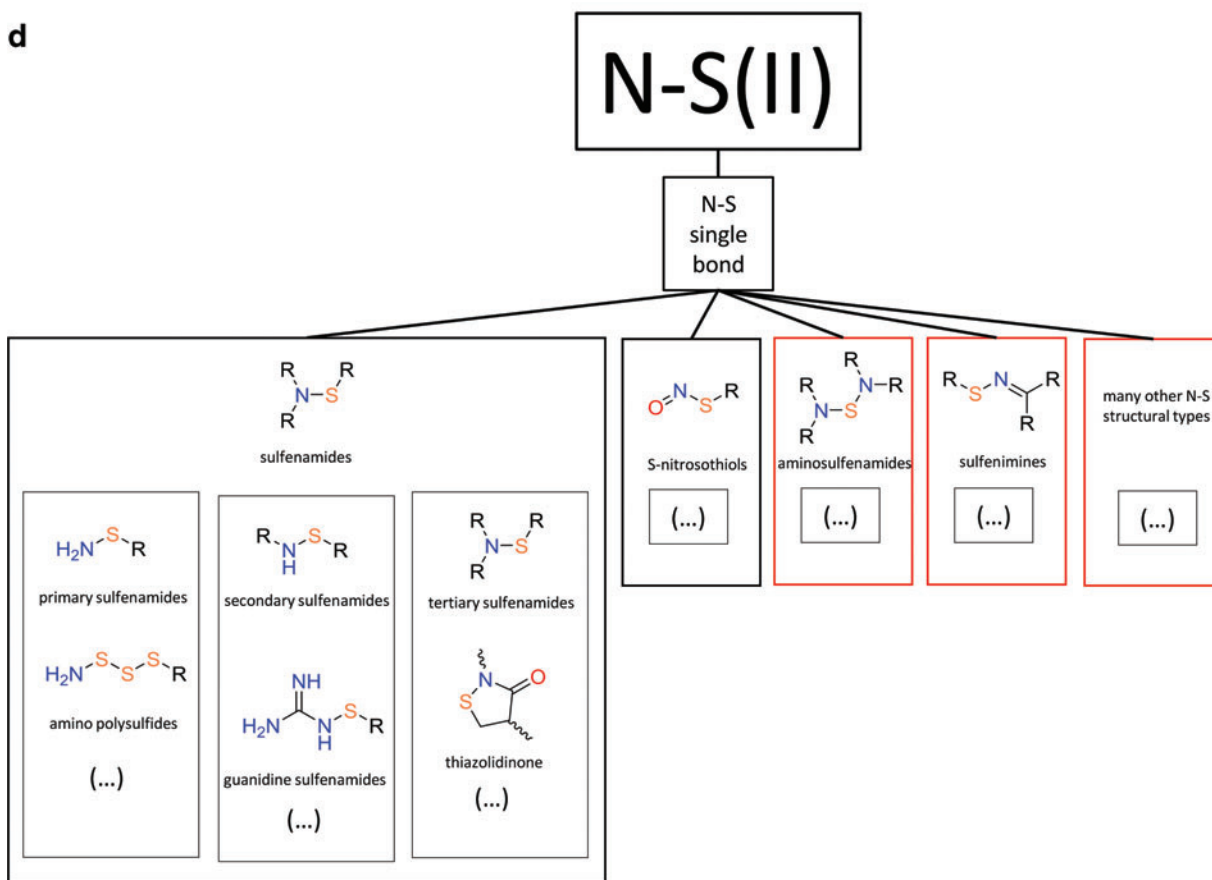


FIG. A1c,d. (Continued).

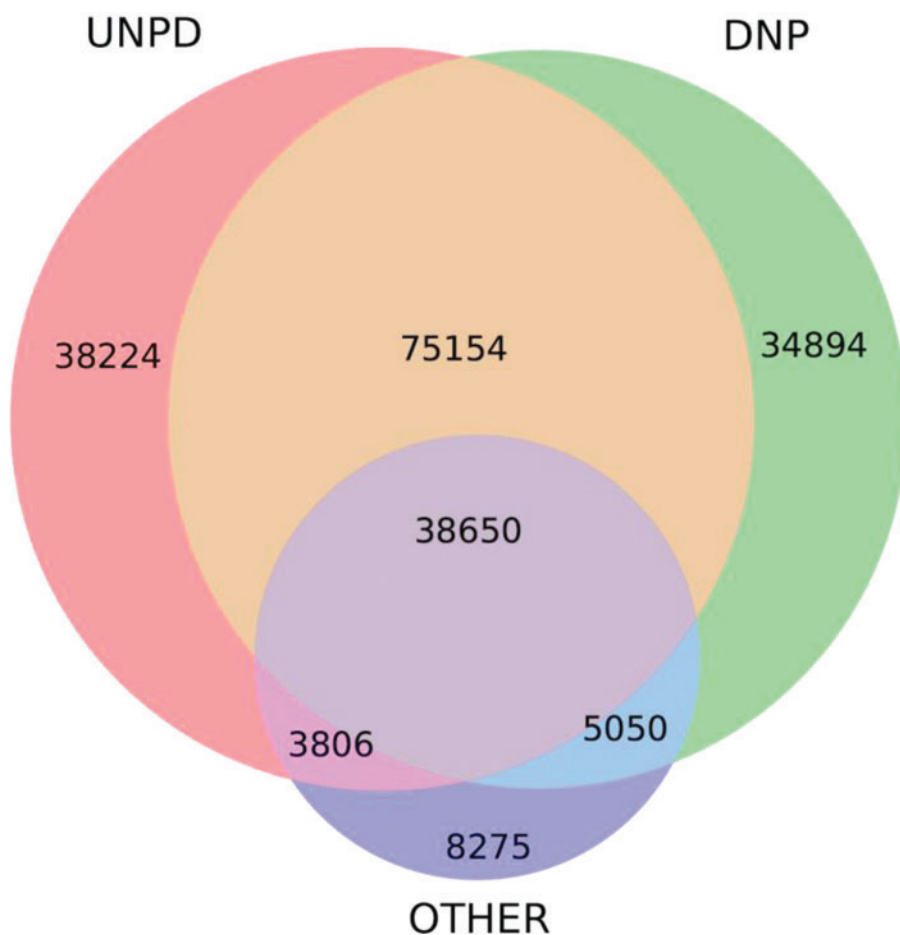


FIG. A2. Venn diagram illustrating the completeness of the Natural Product Database. Each circle describes major sources of natural products we took for our database. The left circle (red) shows the UNPD database (Gu *et al.*, 2013), the right circle is the DNP (DNP, 2016), and the bottom, “other,” represents a collection of natural products from both a manual literature search and from information in smaller databases. Details are provided in Appendix Table A2. The majority of natural products collected from these three different sources are found in at least two sources. This implies our search is moderately complete as to the chemicals known to be produced by life, and is unlikely to have omitted any major classes of chemicals. We note that the numbers are for unique natural products; databases often include more than one entry of the same natural product. We removed any duplicated entries before the completeness analysis and subsequent motif matching. Our final natural product database with only unique entries contains over 200,000 natural products.

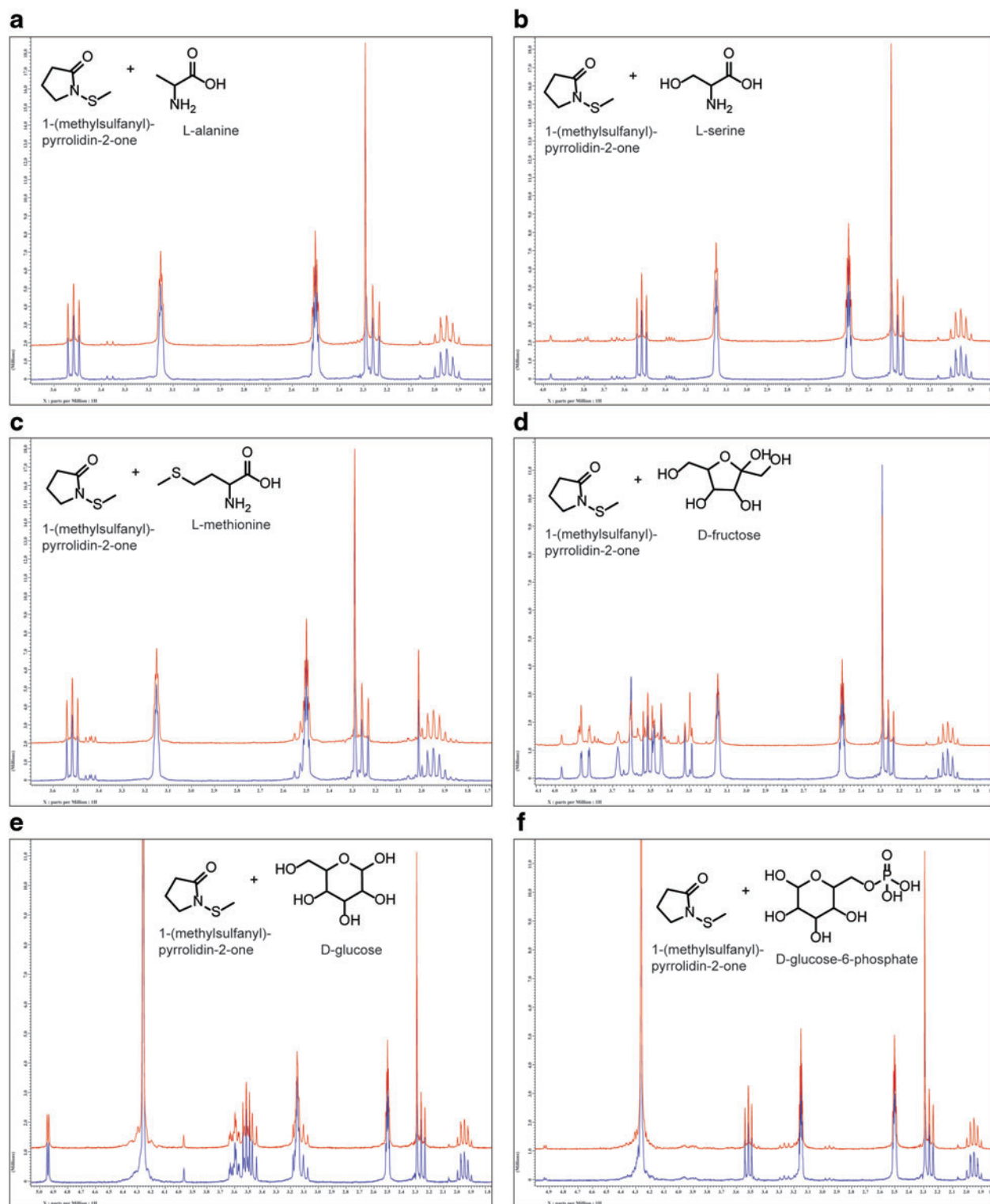


FIG. A3. The N-S(II) bond is stable in the presence of metabolites lacking the thiol functional group, as shown by NMR time-course reactivity assay of N-S compound 1-(methylsulfanyl)pyrrolidin-2-one with six major cellular metabolites containing major non-thiol functional groups. No changes, and hence no reactivity, were detected in the spectra of the mixtures of 1-(methylsulfanyl)pyrrolidin-2-one with (a) L-alanine, (b) L-serine, (c) L-methionine, (d) D-fructose*, (e) D-glucose, (f) D-glucose-6-phosphate (compare $t=0$ [blue] and $t=24$ h [red]). *The solution containing 1-(methylsulfanyl)pyrrolidin-2-one and D-fructose showed significant differences in the signals corresponding to the sugar, due to instability in the solvent system used. Signal from 1-(methylsulfanyl)pyrrolidin-2-one showed no change, demonstrating that the changes in the D-fructose signal were not due to reaction with compound 1-(methylsulfanyl)pyrrolidin-2-one.

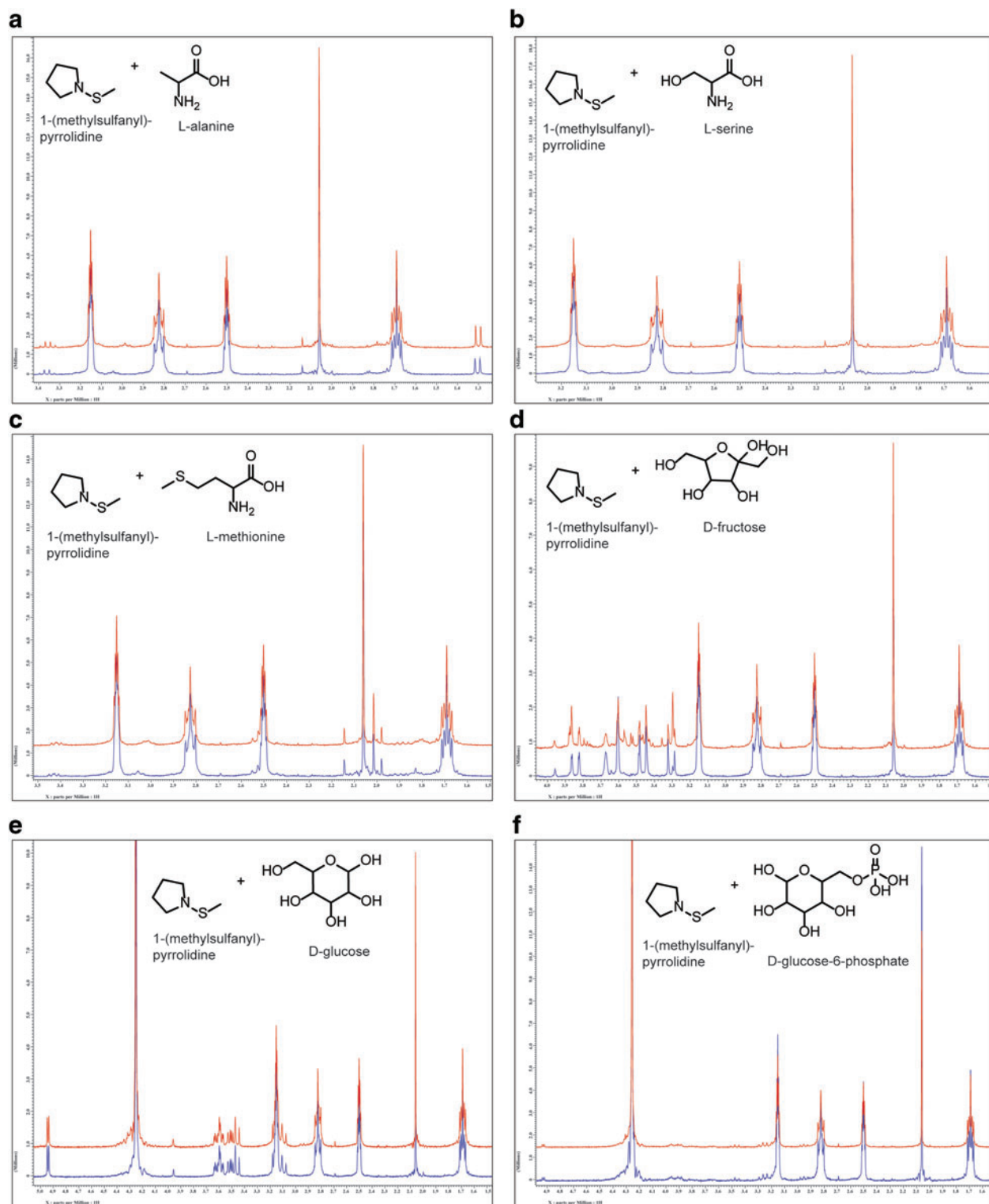


FIG. A4. The N-S(II) bond is stable in the presence of metabolites lacking the thiol functional group, as shown by NMR time-course reactivity assay of N-S compound 1-(methylsulfanyl)pyrrolidine with six major cellular metabolites containing major non-thiol functional groups. No changes, and hence no reactivity, were detected in the spectra of the mixtures of 1-(methylsulfanyl)pyrrolidine with (a) L-alanine, (b) L-serine, (c) L-methionine, (d) D-fructose*, (e) D-glucose, (f) D-glucose-6-phosphate (compare $t=0$ [blue] and $t=24$ h [red]). *The solution containing 1-(methylsulfanyl)pyrrolidine and D-fructose showed significant differences in the signals corresponding to the sugar, due to instability in the solvent system used. Signal from 1-(methylsulfanyl)pyrrolidine showed no change, demonstrating that the changes in the D-fructose signal were not due to reaction with 1-(methylsulfanyl)pyrrolidine.

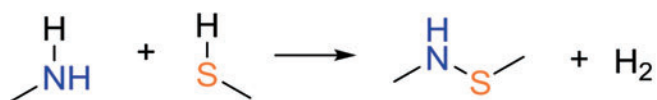
Category of Compound				Eukarya				Bacteria			
				Alveolata	Animalia	Fungi	Plants	Cyanobacteria	Bacteroidetes	Proteobacteria	Actinobacteria
sulfamate											
Number of natural products											
73											
secondary sulfamate		29	6	22	1	0	4	0	0	0	
tertiary sulfamate		3	0	3	0	0	0	0	0	0	
azetidinone sulfamate		15	0	0	0	0	0	5	10	0	
phospho sulfamate		4	0	0	0	0	0	0	3	1	
primary sulfamate ester		5	0	0	0	0	0	0	0	5	
pyrrole sulfamate		14	0	11	0	3	0	0	0	0	
secondary sulfamate ester		3	0	2	0	0	0	0	0	1	
sulfonamide											
Number of natural products											
8											
secondary sulfonamide		3	0	0	0	1	0	0	1	3	
tertiary sulfonamide		1	0	1	0	0	0	0	0	0	
primary sulfonamide		2	0	1	0	0	0	0	0	1	
aryl sulfonamides		2	0	0	1	0	0	0	0	1	
sulfoximines											
Number of natural products											
3											
primary sulfoximines		1	0	0	0	1	0	0	0	0	
phospho sulfoximines		2	0	0	0	0	0	0	0	2	

FIG. A5. N-S compounds are produced by a diverse range of organisms belonging to all major branches of life. Most classes of N-S bond are made by more than one kingdom of life, and most structural types of N-S bond were discovered throughout evolution by more than one kingdom of life, suggesting that evolutionary contingency is not a likely explanation for exclusion of N-S chemistry by life. The evolutionary barrier to containing N-S bonds in metabolism is evolving mechanisms to protect the cell from the toxic result of N-S chemistry, not that N-S chemistry itself is difficult to evolve.

aromatic N-S compounds											
	Isotiazoles		5	0	0	0	2	1	0	0	2
	Thiadiazoles		5	0	3	1	1	0	0	0	0
	Isothiazines		1	0	1	0	0	0	0	0	0
sulfenamide				Alveolata	Animalia	Fungi	Plants	Cyanobacteria	Bacteroidetes	Proteobacteria	Actinobacteria
	guanidine sulfenamide		1	0	0	0	1	0	0	0	0
	cyclic sulfenamides		1	0	1	0	0	0	0	1	0
	secondary linear sulfenamides		1	0	0	0	0	0	0	1	0
	sulfilimines		2	0	2	0	0	0	0	0	0
	aminopolysulfides		1	0	0	0	1	0	0	0	0
	dithiazines		1	0	0	0	1	0	0	0	0
	dithiadiazetidines		1	0	0	0	1	0	0	0	0
	nitrosothiols		3	1	3	1	1	1	1	1	1
	sulfenamides		3	0	3	0	0	0	0	0	0
	N-thiosulfoximides		1	0	1	0	0	0	0	0	0
TOTAL			110								

FIG. A5. (Continued).

a dehydrogenation approach



b rearrangement approach

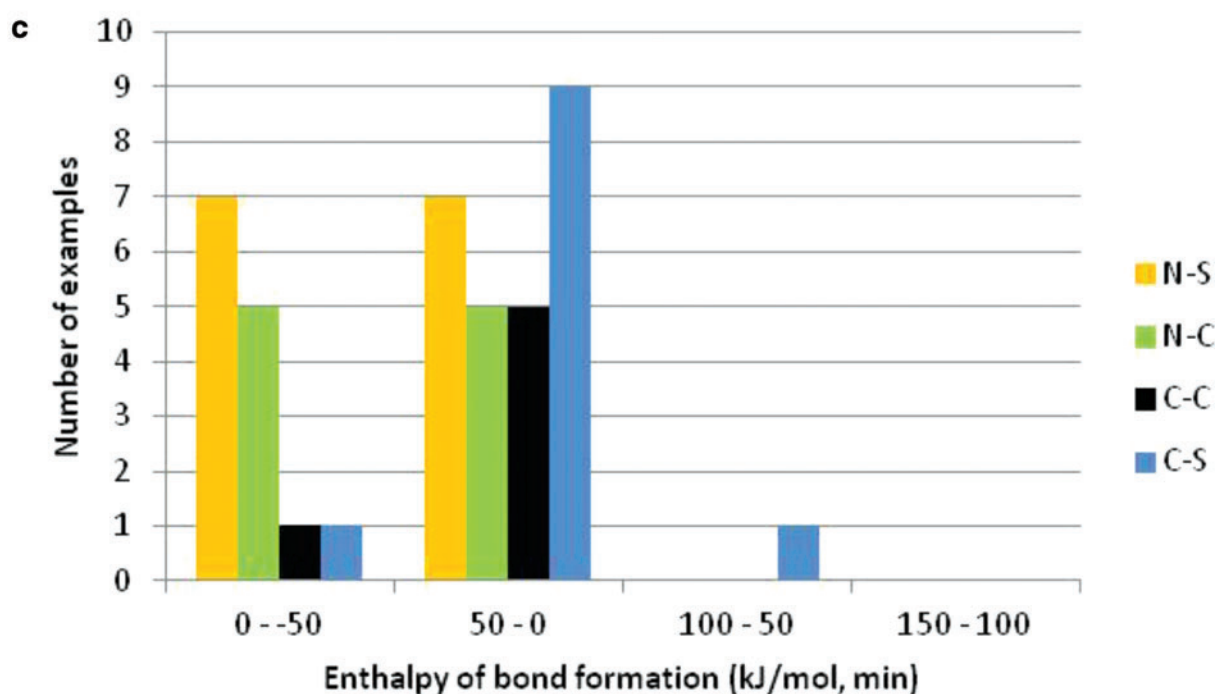
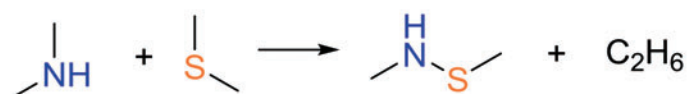


FIG. A6. Estimation of the energy needed to form a stable molecule containing an N-S bond from two stable substrate molecules that do not contain an N-S bond. The x axis shows energy needed to form a bond. The y axis shows number of bonds formed that had that energy. Yellow bars=N-S bonds. Three other bond types are shown for comparison: Black bars=C-C bonds, Green bars=C-N bonds, Blue bars=C-S bonds. N-S bonds are no more energy-requiring than other, biologically common bond types.

Energy calculations are done by using semi-empirical quantum mechanical methods (see main text). The enthalpy (“heat of formation”) of a bond is calculated from the energy difference between two molecules that do not contain the bond under consideration and two rearranged molecules that do contain that bond. Two transformations were considered: (a) dehydrogenation, where we assume that an N-S bond is formed by removal of two hydrogens from two substrate molecules, and (b) elimination of ethane, where an N-S bond is formed by removal of two methyl groups from preexisting N-CH₃ and S-CH₃ bonds. The difference between the energy of the two starting molecules and the two product molecules is the energy of formation of the N-S bond. (Note that these are comparison of the energy of four molecules, not realistic chemical reactions.) Comparisons were calculated for both transformations on 64 pairs of molecules and their transformed equivalents that represent all major chemical classes of N-S compounds, that is, N-S(II), N-S(IV), and N-S(VI).

Acknowledgments

We thank MIT and the MIT Amar G. Bose Research Grant for support. We thank Organix Inc. (<http://www.organixinc.com>) for performing experimental work. We also would like to extend special thanks to the reviewer whose comments and suggestions substantially improved the manuscript during the editing process.

References

- Adams, A., Freeman, W.A., Holland, A., Hossack, D., Inglis, J., Parkinson, J., Reading, H.W., Rivett, K., Slack, R., Sutherland, R., and Wien, R. (1960) Sulphasomizole (5-p-aminobenzenesulphonamido-3-methylisothiazole): a new antibacterial sulphonamide. *Nature* 186:221–222.
- Alnouti, Y. and Klaassen, C.D. (2006) Tissue distribution and ontogeny of sulfotransferase enzymes in mice. *Toxicol Sci* 93: 242–255.
- Arndt, K.E., Bland D.C., Podhorez D.E., and McConnell, J.R. (2009) Process for the oxidation of certain substituted sulfilimines to insecticidal sulfoximines. Dow AgroSciences LLC, Indianapolis, IN. Available at Google Patents.
- Bains, W. (2014) A trip through chemical space: why life has evolved the chemistry that it has. In *Evolutionary Biology: Genome Evolution, Speciation, Coevolution and Origin of Life*, edited by P. Pontarottis, Springer International Publishing, Cham, Switzerland, pp 371–394.
- Bains, W. and Schulze-Makuch, D. (2016) The cosmic zoo: the (near) inevitability of the evolution of complex, macroscopic life. *Life* 6, doi:10.3390/life6030025.
- Bains, W. and Seager S. (2012) A combinatorial approach to biochemical space: description and application to the redox distribution of metabolism. *Astrobiology* 12:271–281.
- Ball, P. (2008) Water as an active constituent in cell biology. *Chem Rev* 108:74–108.
- Beal, J.L., Foster, S.B., and Ashby, M.T. (2009) Hypochlorous acid reacts with the N-terminal methionines of proteins to give dehydromethionine, a potential biomarker for neutrophil-induced oxidative stress. *Biochemistry* 48:11142–11148.
- Beck, J.R. (1975) Substituted sulfanyl sulfilimine compounds as herbicides. Eli Lilly and Co Ltd, Basingstoke, UK. Available at Google Patents.
- Bektas, Y. and Eulgem, T. (2014) Synthetic plant defense elicitors. *Front Plant Sci* 5, doi:10.3389/fpls.2014.00804.
- Benhar, M., Forrester, M. T., and Stamler, J. S. (2009) Protein denitrosylation: enzymatic mechanisms and cellular functions. *Nat Rev Mol Cell Biol* 10:721–732.
- Benner, S.A., Ricardo, A., and Carrigan, M.A. (2004) Is there a common chemical model for life in the Universe? *Curr Opin Chem Biol* 8:672–689.
- Benner, S.A., Bains, W., and Seager, S. (2013) Models and standards of proof in cross-disciplinary science: the case of arsenic DNA. *Astrobiology* 13:510–513.
- Bergen, H.S. and Craver, J.K. (1947) Sulfonamide plasticizers and resins. *Industrial & Engineering Chemistry* 39:1082–1087.
- Bland, D.C., Ross, R., Johnson, P.L., and Johnson, T.C. (2014) Insecticidal n-substituted sulfilimine and sulfoximine pyridine n-oxides. Dow Agrosciences LLC, Indianapolis, IN. Available at Google Patents.
- Blankenship, R.E. (2010) Early evolution of photosynthesis. *Plant Physiol* 154:434–438.
- Blunt, C.E., Torcuk, C., Liu, Y., Lewis, W., Siegel, D., Ross, D., and Moody, C.J. (2015) Synthesis and intracellular redox cycling of natural quinones and their analogues and identification of indoleamine-2,3-dioxygenase (IDO) as potential target for anticancer activity. *Angew Chem Int Ed Engl* 54: 8740–8745.
- Bowden, N.B., Kuruvilla, D.J., Salem, A.K., and Yoo, J. (2014) Biodegradable polymers with sulfenamide bonds for drug delivery applications. University of Iowa Research Foundation (UIRF), Iowa City, IA. Available at Google Patents.
- Bradbury, R., Butters, A., Clive Moscrop, and Stark, A. (1996) Ink Composition—US Patent, Zeneca Ltd. USA.
- Braun, O., Knipp, M., Chesnov, S., and Vasak, M. (2007) Specific reactions of S-nitrosothiols with cysteine hydrolases: a comparative study between dimethylargininase-1 and CTP synthetase. *Protein Sci* 16:1522–1534.
- Broniowska, K.A. and Hogg, N. (2012) The chemical biology of S-nitrosothiols. *Antioxid Redox Signal* 17:969–980.
- Carlsson, P. and Kjellén, L. (2012) Heparin biosynthesis. In *Heparin—A Century of Progress*, edited by R. Lever, B. Mulloy, and C.P. Pages, Springer, Berlin, pp 23–41.
- Carta, F., Scozzafava, A., and Supuran, C.T. (2012) Sulfonamides: a patent review (2008–2012). *Expert Opin Ther Pat* 22:747–758.
- Casini, A., Scozzafava, A., and Supuran, C.T. (2002) Cysteine-modifying agents: a possible approach for effective anticancer and antiviral drugs. *Environ Health Perspect* 110:801–806.
- Chenier, P.J. (2002) *Survey of Industrial Chemistry*, Springer, Boston, MA.
- Chiu, Y.C., Black, A.L., and Fukuto, T.R. (1975) Thiolytic activation process in N-sulfenylated derivatives of methylcarbamate esters. *Pestic Biochem Physiol* 5:359–366.
- Choi, H., Cho, S.Y., Pak, H.J., Kim, Y., Choi, J.-y., Lee, Y.J., Gong, B.H., Kang, Y.S., Han, T., Choi, G., Lee, S., Ryhoo, D., and Park, H. (2017) NPCARE: database of natural products and fractional extracts for cancer regulation. *J Cheminform* 9, doi:10.1186/s13321-016-0188-5.
- Collier, P.J., Ramsey, A., Waigh, R.D., Douglas, K.T., Austin, P., and Gilbert, P. (1990) Chemical reactivity of some isothiazolone biocides. *J Appl Bacteriol* 69:578–584.
- Cortese-Krott, M.M., Butler, A.R., Woollins, J.D., and Feelisch, M. (2016) Inorganic sulfur-nitrogen compounds: from gunpowder chemistry to the forefront of biological signaling. *Dalton Trans* 45:5908–5919.
- Craine, L. and Raban, M. (1989) The chemistry of sulfenamides. *Chem Rev* 89:689–712.
- Cummings, C.F. (2013) *Formation and Function of Collagen IV Sulfilimine Bonds*, PhD dissertation, Vanderbilt University, Nashville, TN.
- Dahlin, J.L., Nissink, J.W.M., Strasser, J.M., Francis, S., Higgins, L., Zhou, H., Zhang, Z., and Walters, M.A. (2015) PAINS in the assay: chemical mechanisms of assay interference and promiscuous enzymatic inhibition observed during a sulfhydryl-scavenging HTS. *J Med Chem* 58:2091–2113.
- Dalvie, D.K., Kalgutkar, A.S., Khojasteh-Bakht, S.C., Obach, R.S., and O'Donnell, J.P. (2002) Biotransformation reactions of five-membered aromatic heterocyclic rings. *Chem Res Toxicol* 15:269–299.
- Davies, P.C.W. (2011) Searching for a shadow biosphere on Earth as a test of the 'cosmic imperative'. *Philos Trans A Math Phys Eng Sci* 369, doi:10.1098/rsta.2010.0235.
- Davis, F.A. (2006) Adventures in sulfur–nitrogen chemistry. *J Org Chem* 71:8993–9003.
- Defelipe, L.A., Lanzarotti, E., Gauto, D., Marti, M.A., and Turjanski, A.G. (2015) Protein topology determines cysteine

- oxidation fate: the case of sulfenyl amide formation among protein families. *PLoS Comput Biol* 11, doi:10.1371/journal.pcbi.1004051.
- Defty, C.L. and Marsden, J.R. (2012) Melphalan in regional chemotherapy for locally recurrent metastatic melanoma. *Curr Top Med Chem* 12:53–60.
- den Hertog, J., Groen, A., and van der Wijk, T. (2005) Redox regulation of protein-tyrosine phosphatases. *Arch Biochem Biophys* 434:11–15.
- de Paulis, T. (2002) Perospirone (Sumitomo Pharmaceuticals). *Curr Opin Investig Drugs* 3:121–129.
- Derese, S., Ndakala, A., Rogo, M., Oyim, J., and Manyim, S. (2015) Mitishamba database: a web based *in silico* database of natural products from Kenya plants [abstract SL 57]. In *The 16th Symposium of The Natural Products Research Network for Eastern and Central Africa (NAPRECA)*.
- Dewar, M.J.S. and Thiel, W. (1977) Ground states of molecules. 38. The MNDO method. Approximations and parameters. *J Am Chem Soc* 99:4899–4907.
- Dewar, M.J.S., Zoebisch, E.G., Healy, E.F., and Stewart, J.J.P. (1985) Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *J Am Chem Soc* 107:3902–3909.
- Dicks, A.P., Li, E., Munro, A.P., Swift, H.R., and Williams, D.L.H. (1998) The reaction of S-nitrosothiols with thiols at high thiol concentration. *Canadian Journal of Chemistry* 76: 789–794.
- DNP. (2016) *Dictionary of Natural Products Online*, Taylor and Francis Group.
- Drahl, C., Cravatt, B.F., and Sorensen, E.J. (2005) Protein-reactive natural products. *Angew Chem Int Ed Engl* 44:5788–5809.
- Dubbs, J.M. and Mongkolsuk, S. (2012) Peroxide-sensing transcriptional regulators in bacteria. *J Bacteriol* 194:5495–5503.
- El-Sayed, A.M. (2016) *The Pherobase: Database of Pheromones and Semiochemicals*. Available online at <http://www.pherobase.com>
- Farnsworth, N.R. (2016) *Natural Products Alert*. Available online at <https://www.napralert.org>
- Fleischer, J.C., Clark, G.T., and Weaver, M.A. (1983) Isothiazole type azo dyes containing imidazo type couplers. Eastman Kodak Co, Rochester, NY. Available at Google Patents.
- Frey, D., Braun, O., Briand, C., Vašák, M., and Grütter, M.G. (2006) Structure of the mammalian NOS regulator dimethylarginine dimethylaminohydrolase: a basis for the design of specific inhibitors. *Structure* 14:901–911.
- Frings, M., Bolm, C., Blum, A., and Gnam, C. (2017) Sulfoximines from a medicinal Chemist's perspective: physicochemical and *in vitro* parameters relevant for drug discovery. *Eur J Med Chem* 126:225–245.
- Fukuto, T.R., March, R.B., and Miller, T.A. (1983) *Chemistry and Mode of Action of Insecticides: Phase II*, United States Environmental Protection Agency, Washington, DC.
- García, I., Vior Natalia, M., Braña Alfredo, F., González-Sabin, J., Rohr, J., Moris, F., Méndez, C., and Salas, J.A. (2012) Elucidating the biosynthetic pathway for the polyketide-nonribosomal peptide collismycin A: mechanism for formation of the 2,2'-bipyridyl ring. *Chem Biol* 19:399–413.
- Gersch, M., Kreuzer, J., and Sieber, S.A. (2012) Electrophilic natural products and their biological targets. *Nat Prod Rep* 29:659–682.
- Go, Y.-M., Chandler, J.D., and Jones, D.P. (2015) The cysteine proteome. *Free Radic Biol Med* 84:227–245.
- Gould, N., Doulias, P.-T., Tenopoulou, M., Raju, K., and Ischiropoulos, H. (2013) Regulation of protein function and signaling by reversible cysteine S-nitrosylation. *J Biol Chem* 288:26473–26479.
- Gray, M.J., Wholey, W.-Y., Parker, B.W., Kim, M., and Jakob, U. (2013) NemR Is a bleach-sensing transcription factor. *J Biol Chem* 288:13789–13798.
- Gray, M.J., Li, Y., Leichert, L.I.-O., Xu, Z., and Jakob, U. (2015) Does the transcription factor NemR use a regulatory sulfenamide bond to sense bleach? *Antioxid Redox Signal* 23: 747–754.
- Gu, J., Gui, Y., Chen, L., Yuan, G., Lu, H.-Z., and Xu, X. (2013) Use of natural products as chemical library for drug discovery and network pharmacology. *PLoS One* 8, doi: 10.1371/journal.pone.0062839.
- Hatherley, R., Brown, D.K., Musyoka, T.M., Penkler, D.L., Faya, N., Lobb, K.A., and Tastan Bishop, Ö. (2015) SANCDDB: a South African natural compound database. *J Cheminform* 7:1–9.
- Heinrich, J.N., Butera, J.A., Carrick, T., Kramer, A., Kowal, D., Lock, T., Marquis, K.L., Pausch, M.H., Popiolek, M., Sun, S.-C., Tseng, E., Uveges, A.J., and Mayer, S.C. (2009) Pharmacological comparison of muscarinic ligands: historical versus more recent muscarinic M1-preferring receptor agonists. *Eur J Pharmacol* 605:53–56.
- Heldreth, B., Long, T.E., Jang, S., Reddy, G.S.K., Turos, E., Dickey, S., and Lim, D.V. (2006) N-Thiolated β -lactam antibacterials: effects of the N-organothio substituent on anti-MRSA activity. *Bioorg Med Chem* 14:3775–3784.
- Hemenway, J.N. (2006) *Preparation, Physicochemical Properties and Animal Studies of New Water-Soluble Prodrugs of Carbamazepine and Oxcarbazepine*, PhD dissertation, University of Kansas, Lawrence, KS.
- Hemenway, J.N., Nti-Addae, K., Guarino, V.R., and Stella, V.J. (2007) Preparation, characterization and *in vivo* conversion of new water-soluble sulfenamide prodrugs of carbamazepine. *Bioorg Med Chem Lett* 17:6629–6632.
- Hillemann, C.L. (1986) Herbicidal sulphonimidamide compounds. E I du Pont de Nemours and Co, Wilmington, DE. Available at Google Patents.
- Hitoshi, S., Yanase Y., Sekino T., Ishikawa K., Kuwatsuka T., Tanikawa H., Kawashima H., Tomura N., and Kanemoto Y. (1993) Isotianil—US Patent. USA.
- Hong, L. and Fast, W. (2007) Inhibition of human dimethylarginine dimethylaminohydrolase-1 by S-nitroso-L-homocysteine and hydrogen peroxide: analysis, quantification, and implications for hyperhomocysteinemia. *J Biol Chem* 282:34684–34692.
- Hooks, H. and Ottmann, G.F. (1966) Halogenated aromatic polysulfinylamines and their preparation. Olin Corporation, Clayton, MO.
- Hörlein, H. (1909) Prontalbin. In *Deutsches Reich Patentschrift*.
- Hu, T.-M. and Chou, T.-C. (2006) The kinetics of thiol-mediated decomposition of S-nitrosothiols. *AAPS J* 8:E485–E492.
- Huttunen, K.M., Leppanen, J., Vepsäläinen, J., Sirvio, J., Laine, K., and Rautio, J. (2012) *In vitro* and *in vivo* evaluation of a sulfenamide prodrug of basic metformin. *J Pharm Sci* 101: 2854–2860.
- Kalgotkar, A.S., Driscoll, J., Zhao, S.X., Walker, G.S., Shepard, R.M., Soglia, J.R., Atherton, J., Yu, L., Mutlib, A.E., Munchhof, M.J., Reiter, L.A., Jones, C.S., Doty, J.L., Trevena, K.A., Shaffer, C.L., and Ripp, S.L. (2007) A rational chemical intervention strategy to circumvent bioactivation liabilities associated with a nonpeptidyl thrombopoietin re-

- ceptor agonist containing a 2-amino-4-arylthiazole motif. *Chem Res Toxicol* 20:1954–1965.
- Kalgutkar, A.S., Jones, R., and Sawant, A. (2010) Sulfonamide as an essential functional group in drug design. In *Metabolism, Pharmacokinetics and Toxicity of Functional Groups: Impact of Chemical Building Blocks on ADMET*, The Royal Society of Chemistry, Cambridge, UK, pp 210–274.
- Kalogirou, A.S. and Koutentis, P.A. (2014) Reactions of selected 3-bromoisothiazole-5-carbonitriles with the secondary dialkylamines pyrrolidine and morpholine. *Tetrahedron* 70: 7092–7099.
- Katritzky, A.R. and Rees, C.W. (1984) *Comprehensive Heterocyclic Chemistry*, Pergamon Press, Oxford, UK.
- Katritzky, A.R., Rees, C.W., and Scriven, E.F.V. (1996) *Comprehensive Heterocyclic Chemistry II*, Pergamon Press, Oxford, UK.
- Katritzky, A.R., Ramsden, C.A., Scriven, E.F.V., and Taylor, R.J.K. (2008) *Comprehensive Heterocyclic Chemistry III*, Elsevier, Amsterdam.
- Keceli, G. and Toscano, J.P. (2015) Reactivity of HNO-induced modifications and its relevance to phospholamban function. *FASEB J* 29:LB207.
- Keceli, G., Moore, C.D., Labonte, J.W., and Toscano, J.P. (2013) NMR detection and study of hydrolysis of HNO-derived sulfinamides. *Biochemistry* 52:7387–7396.
- Kirkpatrick, P. and Ellis, C. (2004) Chemical space. *Nature* 432:823.
- Klementz, D., Döring, K., Lucas, X., Telukunta, K.K., Erleben, A., Deubel, D., Erber, A., Santillana, I., Thomas, O.S., Bechthold, A., and Günther, S. (2016) StreptomeDB 2.0—an extended resource of natural products produced by streptomycetes. *Nucleic Acids Res* 44:D509–D514.
- Knipp, M., Braun, O., and Vašák, M. (2005) Searching for DDAH Inhibitors: S-nitroso-l-homocysteine is a chemical lead. *J Am Chem Soc* 127:2372–2373.
- Koval, I.V. (1990) Progress in the chemistry of sulphilimines. *Russian Chemical Reviews* 59:819–831.
- Koval, I.V. (1996a) Chemistry of sulfenamides. *Russian Journal of Organic Chemistry* 32:1239–1270.
- Koval, I.V. (1996b) Synthesis and application of sulfenamides. *Russian Chemical Reviews* 65:421–440.
- Kresze, G. and Wucherpfennig, W. (1967) New methods of preparative organic chemistry V: organic syntheses with imides of sulfur dioxide. *Angew Chem Int Ed Engl* 6:149–167.
- Lemfack, M.C., Nickel, J., Dunkel, M., Preissner, R., and Piechulla, B. (2014) mVOC: a database of microbial volatiles. *Nucleic Acids Res* 42:D744–D748.
- Licke, G.C. (1971) Gasoline containing N-sulfinyl amine. Ethyl Corporation, Richmond, VA.
- Lieberman, J.A., Javitch, J.A., and Moore, H. (2008) Cholinergic agonists as novel treatments for schizophrenia: the promise of rational drug development for psychiatry. *Am J Psychiatry* 165:931–936.
- Loso, M.R., Nugent, B.M., Zhu, Y., Rogers, R.B., Huang, J.X., Renga, J.M., Benko, Z.L., Whiteker, G.T., and Daeuble, J.F. (2012) Insecticidal N-substituted (heteroaryl)alkyl sulfilimines. Dow AgroSciences LLC, Indianapolis, IN.
- Lücking, U. (2013) Sulfoximines: a neglected opportunity in medicinal chemistry. *Angew Chem Int Ed Engl* 52:9399–9408.
- Mazzei, T. (2010) The pharmacokinetics and pharmacodynamics of the carbapenems: focus on doripenem. *J Chemother* 22:219–225.
- McCall, A.S., Cummings, C.F., Bhavé, G., Vanacore, R., Page-McCaw, A., and Hudson, B.G. (2014) Bromine is an essential trace element for assembly of collagen IV scaffolds in tissue development and architecture. *Cell* 157:1380–1392.
- Morton, D. and Stockman, R.A. (2006) Chiral non-racemic sulfinimines: versatile reagents for asymmetric synthesis. *Tetrahedron* 62:8869–8905.
- Mueller, J.W. and Shafqat, N. (2013) Adenosine-5'-phosphosulfate—a multifaceted modulator of bifunctional 3'-phospho-adenosine-5'-phosphosulfate synthases and related enzymes. *FEBS J* 280:3050–3057.
- Nagy, P. (2013) Kinetics and mechanisms of thiol–disulfide exchange covering direct substitution and thiol oxidation-mediated pathways. *Antioxid Redox Signal* 18:1623–1641.
- Nagy, P. and Winterbourn, C.C. (2010) Redox chemistry of biological thiols. In *Advances in Molecular Toxicology*, edited by J.C. Fishbeins, Elsevier, Amsterdam, pp 183–222.
- Nakamura, Y., Asahi, H., Altaf-Ul-Amin, M., Kurokawa, K., and Kanaya, S. (2015) *KNAPSAcK: A Comprehensive Species-Metabolite Relationship Database*, NAIST Comparative Genomics Laboratory, Nara, Japan.
- Nelsen, S.F., Steffek D.J., Cunkle G.T., and Gannett P.M. (1982) One-electron oxidation of trialkylsulfenamides. *J Am Chem Soc* 104:6641–6646.
- Nti-Addae, K.W. (2008) *Synthesis and Physicochemical Characterization of Sulfenamide Prodrugs of Antimicrobial Oxazolidinones*, PhD dissertation, University of Kansas, Lawrence, KS.
- Nti-Addae, K.W. and Stella, V.J. (2011) *In vitro* conversion of model sulfenamide prodrugs in the presence of small molecule thiols. *J Pharm Sci* 100:1001–1008.
- Nti-Addae, K.W., Laurence, J.S., Skinner, A.L., and Stella, V.J. (2011) Reversion of sulfenamide prodrugs in the presence of free thiol containing proteins. *J Pharm Sci* 100:3023–3027.
- Ntie-Kang, F., Telukunta, K.K., Döring, K., Simoben, C.V., Moubock, A.F., Malange, Y.I., Njume, L.E., Yong, J.N., Sippl, W., and Günther, S. (2017) NANPDB: a resource for natural products from Northern African sources. *J Nat Prod* 80:2067–2076.
- Olbe, L., Carlsson, E., and Lindberg, P. (2003) A proton-pump inhibitor expedition: the case histories of omeprazole and esomeprazole. *Nat Rev Drug Discov* 2:132–139.
- Pace, N.R. (2001) The universal nature of biochemistry. *Proc Natl Acad Sci USA* 98:805–808.
- Paulsen, C.E. and Carroll, K.S. (2013) Cysteine-mediated redox signaling: chemistry, biology, and tools for discovery. *Chem Rev* 113:4633–4679.
- Peskin, A.V., Turner, R., Maghzal, G.J., Winterbourn, C.C., and Kettle, A.J. (2009) Oxidation of methionine to dehydromethionine by reactive halogen species generated by neutrophils. *Biochemistry* 48:10175–10182.
- Petkowski, J.J., Bains, W., and Seager, S. (2018) Natural products containing a nitrogen-sulfur bond. *Journal of Natural Products* 81:423–446.
- Petrov, K.A., Rudnev, G.V., and Sorokin, V.D. (1990) Sulphenamides and their derivatives. *Russian Chemical Reviews* 59:832–843.
- Poole, L.B. (2015) The basics of thiols and cysteines in redox biology and chemistry. *Free Radic Biol Med* 80: 148–157.
- Proença, C., Serralheiro, M.L., Araújo, M.E., Pamplona, T., Santos, S., Santos, M.S., and Frazão, F. (2011) Novel sulfenamides as promising acetylcholinesterase inhibitors. *J Heterocycl Chem* 48:1287–1294.

- Prosen, K.R., Carroll, R.K., Burda, W.N., Krute, C.N., Bhat-tacharya, B., Dao, M.L., Turos, E., and Shaw, L.N. (2011) The impact of fatty acids on the antibacterial properties of N-thiolated β -lactams. *Bioorg Med Chem Lett* 21:5293–5295.
- Raftery, M.J., Yang, Z., Valenzuela, S.M., and Geczy, C.L. (2001) Novel intra- and inter-molecular sulfenamides in S100A8 produced by hypochlorite oxidation. *J Biol Chem* 276:33393–33401.
- Ramaraju, P., Gergeres, D., Turos, E., Dickey, S., Lim, D.V., Thomas, J., and Anderson, B. (2012) Synthesis and antimicrobial activities of structurally novel S,S'-bis(heterosubstituted) disulfides. *Bioorg Med Chem Lett* 22:3623–3631.
- Raymond, J., Siefert, J.L., Staples, C.R., and Blankenship, R.E. (2004) The natural history of nitrogen fixation. *Mol Biol Evol* 21:541–554.
- Reich, H.J. and Hondal, R.J. (2016) Why nature chose selenium. *ACS Chemical Biology* 11:821–841.
- Reitz, A.B., Smith, G.R., and Parker, M.H. (2009) The role of sulfamide derivatives in medicinal chemistry: a patent review (2006–2008). *Expert Opin Ther Pat* 19:1449–1453.
- Revell, K.D., Heldreth, B., Long, T.E., Jang, S., and Turos, E. (2007) N-thiolated β -lactams: studies on the mode of action and identification of a primary cellular target in *Staphylococcus aureus*. *Bioorg Med Chem* 15:2453–2467.
- Reymond, J.-L., Ruddigkeit, L., Blum, L., and van Deursen, R. (2012) The enumeration of chemical space. *Wiley Interdiscip Rev Comput Mol Sci* 2:717–733.
- Robertson, W.E., Rose, K.L., Hudson, B.G., and Vanacore, R.M. (2014) Supramolecular organization of the α 121- α 565 collagen IV network. *J Biol Chem* 289:25601–25610.
- Ruddigkeit, L., van Deursen, R., Blum, L.C., and Reymond, J.-L. (2012) Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. *J Chem Inf Model* 52:2864–2875.
- Sanchez, R.I. and Kauffman, F.C. (2010) 9.05—Regulation of xenobiotic metabolism in the liver A2. In *Comprehensive Toxicology*, 2nd ed., edited by C.A. McQueen, Elsevier, Oxford, UK, pp 109–128.
- Sannino, A., Bolzoni, L., and Bandini, M. (2004) Application of liquid chromatography with electrospray tandem mass spectrometry to the determination of a new generation of pesticides in processed fruits and vegetables. *J Chromatogr A* 1036:161–169.
- Sarma, B.K. and Mughes, G. (2007) Redox regulation of protein tyrosine phosphatase 1B (PTP1B): a biomimetic study on the unexpected formation of a sulfonyl amide intermediate. *J Am Chem Soc* 129:8872–8881.
- Schmidt, M.W., Baldrige, K.K., Boatz, J.A., Elbert, S.T., Gordon, M.S., Jensen, J.H., Koseki, S., Matsunaga, N., Nguyen, K.A., Su, S., Windus, T.L., Dupuis, M., and Montgomery, J.A., Jr. (1993) General atomic and molecular electronic structure system. *J Comput Chem* 14:1347–1363.
- Scozzafava, A., Carta, F., and Supuran, C.T. (2013) Secondary and tertiary sulfonamides: a patent review (2008–2012). *Expert Opin Ther Pat* 23:203–213.
- Seager, S., Bains, W., and Petkowski, J.J. (2016) Toward a list of molecules as potential biosignature gases for the search for life on exoplanets and applications to terrestrial biochemistry. *Astrobiology* 16:465–485.
- Shang, J.-L., Guo, H., Li, Z.-S., Ren, B., Li, Z.-M., Dai, H.-q., Zhang, L.-X., and Wang, J.-G. (2013) Synthesis and evaluation of novel sulfenamides as novel anti methicillin-resistant *Staphylococcus aureus* agents. *Bioorg Med Chem Lett* 23:724–727.
- Shiau, T.P., Erlanson, D.A., and Gordon, E.M. (2006) Selective reduction of peptide isothiazolidin-3-ones. *Org Lett* 8:5697–5699.
- Sikorski, J.A. and Hoobler, M.A. (1984) Herbicidal amino-sulfenamide derivatives of N-phosphonomethylglycine triesters. Monsanto, St. Louis, MO. Available at Google Patents.
- Singh, R. and Whitesides, G.M. (2010) Thiol—disulfide interchange. In *Sulphur-Containing Functional Groups (1993)*, John Wiley & Sons, Hoboken, NJ, pp 633–658.
- Sivaramakrishnan, S., Cummings, A.H., and Gates, K.S. (2010) Protection of a single-cysteine redox switch from oxidative destruction: on the functional role of sulfonyl amide formation in the redox-regulated enzyme PTP1B. *Bioorg Med Chem Lett* 20:444–447.
- Slack, R. and Wooldridge, K.R.H. (1965) Isothiazoles. In *Advances in Heterocyclic Chemistry*, edited by A.R. Katritzky, Academic Press, New York, pp 107–120.
- Smith, B.C. and Marletta, M.A. (2012) Mechanisms of S-nitrosothiol formation and selectivity in nitric oxide signaling. *Curr Opin Chem Biol* 16:498–506.
- Steinhagen, H. (2011) The evolution of drug discovery: from traditional medicines to modern drugs. By Enrique Raviña. *ChemMedChem* 6:1746–1747.
- Stewart, J.J.P. (1989) Optimization of parameters for semiempirical methods I. Method. *J Comput Chem* 10:209–220.
- Stewart, J.J.P. (2007) Semiempirical molecular orbital methods. In *Reviews in Computational Chemistry*, edited by K.B. Lipkowitz and D.B. Boyd, John Wiley & Sons, Hoboken, NJ, pp 45–81.
- Tawfik, D.S. and Viola, R.E. (2011) Arsenate replacing phosphate—alternative life chemistries and ion promiscuity. *Biochemistry* 50:1128–1134.
- Teffera, Y., Choquette, D., Liu, J., Colletti, A.E., Hollis, L.S., Lin, M.-H.J., and Zhao, Z. (2010) Bioactivation of isothiazoles: minimizing the risk of potential toxicity in drug discovery. *Chem Res Toxicol* 23:1743–1752.
- Tomizawa, M. and Casida, J.E. (2003) Selective toxicity of neonicotinoids attributable to specificity of insect and mammalian nicotinic receptors. *Annu Rev Entomol* 48: 339–364.
- Turos, E. (2005) N-thiolated beta-lactams: novel antibacterial agents for methicillin-resistant *Staphylococcus aureus*. University of South Florida, Tampa, FL. Available at Google Patents.
- Turos, E., Carpenter, E.T., Long, T., Lim, D.V., and Dickey, S.S. (2002) N-thiolated β -lactam antibiotics. University of South Florida, Tampa, FL. Available at Google Patents.
- Umetsu, N., Kuwano, E., and Fukuto, T.R. (1980) Nature of N-S bond cleavage of 2,3-dihydro-2,2-dimethyl-7-benzofuranyl (di-n-butylaminosulfonyl) (methyl)carbamate. *J Environ Sci Health B* 15:1–23.
- Valli, M., dos Santos, R.N., Figueira, L.D., Nakajima, C.H., Castro-Gamboa, I., Andricopulo, A.D., and Bolzani, V.S. (2013) Development of a natural products database from the biodiversity of Brazil. *J Nat Prod* 76:439–444.
- Vanacore, R., Ham, A.-J.L., Voehler, M., Sanders, C.R., Conrads, T.P., Veenstra, T.D., Sharpless, K.B., Dawson, P.E., and Hudson, B.G. (2009) A sulfilimine bond identified in collagen IV. *Science* 325:1230–1234.
- Vanelslander, B., Paul, C., Grueneberg, J., Prince, E.K., Gillard, J., Sabbe, K., Pohnert, G., and Vyverman, W. (2012) Daily bursts of biogenic cyanogen bromide (BrCN) control biofilm formation around a marine benthic diatom. *Proc Natl Acad Sci USA* 109:2412–2417.

- Waldman, A.J., Ng, T.L., Wang, P., and Balskus, E.P. (2017) Heteroatom-heteroatom bond formation in natural product biosynthesis. *Chem Rev* 117:5784–5863.
- Wallace, G.C. and Zerba, E.N. (1989) *In vitro* evidence for activative thiolysis and selfsynergism of sulfenyl dicarbamate derivatives of 3,4 methylenedioxyphenyl N-methylcarbamate. *Pesticide Science* 27:233–241.
- Westheimer, F.H. (1987) Why nature chose phosphates. *Science* 235:1173–1178.
- Williams, D.L.H. (1996) The mechanism of nitric oxide formation from S-nitrosothiols (thionitrites). *Chem Commun* 1996:1085–1091.
- Winum, J.-Y., Scozzafava, A., Montero, J.-L., and Supuran, C.T. (2005) Sulfamates and their therapeutic potential. *Med Res Rev* 25:186–228.
- Wong, P.S.Y., Hyun, J., Fukuto, J.M., Shirota, F.N., DeMaster, E.G., Shoeman, D.W., and Nagasawa, H.T. (1998) Reaction between S-nitrosothiols and thiols: generation of nitroxyl (HNO) and subsequent chemistry. *Biochemistry* 37:5362–5371.
- Ye, H., Ye, L., Kang, H., Zhang, D., Tao, L., Tang, K., Liu, X., Zhu, R., Liu, Q., Chen, Y.Z., Li, Y., and Cao, Z. (2011) HIT: linking herbal active ingredients to targets. *Nucleic Acids Res* 39:D1055–D1059.
- Yi, J.H., Perumalsamy, H., Sankarapandian, K., Choi, B.-R., and Ahn, Y.-J. (2015) Fumigant toxicity of phenylpropanoids identified in *Asarum sieboldii* aerial parts to *Lycoriella ingenua* (Diptera: Sciaridae) and *Coboldia fuscipes* (Diptera: Scatopsidae). *J Econ Entomol* 108:1208–1214.
- Young, P.R. and Briedis, A.V. (1988) Kinetics and mechanism of the glutathione-dependent reduction of dehydromethionine. *Biochim Biophys Acta* 967:318–321.
- Zeng, X., Zhang, P., He, W., Qin, C., Chen, S., Tao, L., Wang, Y., Tan, Y., Gao, D., Wang, B., Chen, Z., Jang, Y.Y., and Chen, Y.Z. (2018) NPASS: natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Res* 46:D1217–D1222.
- Zhou, S., Gu, Y., Liu, M., Wu, C., Zhou, S., Zhao, Y., Jia, Z., Wang, B., Xiong, L., Yang, N., and Li, Z. (2014a) Insecticidal activities of chiral N-trifluoroacetyl sulfilimines as potential ryanodine receptor modulators. *J Agric Food Chem* 62:11054–11061.
- Zhou, S., Jia, Z., Xiong, L., Yan, T., Yang, N., Wu, G., Song, H., and Li, Z. (2014b) Chiral dicarboxamide scaffolds containing a sulfiliminyl moiety as potential ryanodine receptor activators. *J Agric Food Chem* 62:6269–6277.
- Zhou, S., Yan, T., Li, Y., Jia, Z., Wang, B., Zhao, Y., Qiao, Y., Xiong, L., Li, Y., and Li, Z. (2014c) Novel phthalamides containing sulfiliminyl moieties and derivatives as potential ryanodine receptor modulators. *Org Biomol Chem* 12:6643–6652.

Address correspondence to:
Janusz J. Petkowski
Dept. of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
77 Mass. Ave.
Cambridge, MA 02139
E-mail: jjpetkow@mit.edu

Submitted 6 February 2018
Accepted 30 August 2018