

# NMR-spectroscopic screening of spider venom reveals sulfated nucleosides as major components for the brown recluse and related species

Frank C. Schroeder<sup>\*†‡</sup>, Andrew E. Taggi<sup>†§</sup>, Matthew Gronquist<sup>¶</sup>, Rabia U. Malik<sup>†</sup>, Jacqueline B. Grant<sup>||</sup>, Thomas Eisner<sup>\*\*</sup>, and Jerrold Meinwald<sup>†‡</sup>

<sup>\*</sup>Boyce Thompson Institute, Cornell University, Tower Road, Ithaca, NY 14853; Departments of <sup>†</sup>Chemistry and Chemical Biology and <sup>\*\*</sup>Neurobiology and Behavior, Cornell University, Ithaca, NY 14853; <sup>¶</sup>Department of Chemistry, SUNY Fredonia, Fredonia, NY 14063; and <sup>||</sup>School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, MI 49931

Contributed by Jerrold Meinwald, July 17, 2008 (sent for review May 12, 2008)

**Extensive chemical analyses of spider venoms from many species have revealed complex mixtures of biologically active compounds, of which several have provided important leads for drug development. We have recently shown that NMR spectroscopy can be used advantageously for a direct structural characterization of the small-molecule content of such complex mixtures. Here, we report the application of this strategy to a larger-scale analysis of a collection of spider venoms representing >70 species, which, in combination with mass spectrometric analyses, allowed the identification of a wide range of known, and several previously undescribed, small molecules. These include polyamines, common neurotransmitters, and amino acid derivatives as well as two additional members of a recently discovered family of natural products, the sulfated nucleosides. In the case of the well studied brown recluse spider, *Loxosceles reclusa*, sulfated guanosine derivatives were found to comprise the major small-molecule components of the venom.**

chemical prospecting | *Loxosceles* | metabolomics | natural products | neurotoxins

**B**iological samples commonly contain complex mixtures of small molecules with diverse properties and functions. Their characterization remains a challenge, because many analytical approaches are primarily aimed at identifying specific classes of known compounds. Several recent studies have demonstrated the feasibility of NMR-spectroscopic analyses of complex small-molecule mixtures, including the use of diffusion-ordered spectroscopy (DOSY) (1) or principal component analysis (PCA) in metabolomics (2) as well as the characterization of crude unfractionated natural products extracts using routine 2D NMR spectra (3–5). Compared with MS analyses, 2D NMR spectroscopic investigations of small-molecule mixtures offer the benefit of more detailed structural information, which is of particular relevance for the detection of unanticipated chemotypes. We have recently reported on the utility of direct NMR spectroscopic analysis of unpurified biological extracts as a powerful method for the discovery of natural products (5). Such analyses provide a means for the rapid evaluation of biological samples in an unbiased and nondestructive manner and are particularly useful for the characterization of glandular secretions with specific biological functions. A standard set of spectra [<sup>1</sup>H, dqfCOSY, heteronuclear multiple quantum correlation (HMQC), heteronuclear multiple bond correlation (HMBC), and NOESY] collected for a crude biological sample will usually permit partial or complete structural characterization of its major components as well as of many minor components. If these “direct” NMR-spectroscopic analyses suggest the presence of compounds of interest, final identification of individual components can be accomplished in a subsequent step, for example by using additional MS analyses, through synthesis of proposed structures, or via NMR-monitored fractionation of the sample

aimed at enrichment or isolation of the compounds of interest. In the case of the hobo spider, *Tegenaria agrestis*, direct NMR spectroscopic analysis facilitated detection and characterization of an entire family of natural products, the sulfated nucleosides, including the sulfated guanosine derivatives **1** and **2** (Fig. 1) (5). Except for a single example (6), sulfated nucleosides had completely escaped detection by conventional isolation and characterization protocols, although spider venoms had been subject to intensive chemical scrutiny that had led to identification of hundreds of proteins, peptides, acylated polyamines, and various small-molecule neurotransmitters (7, 8).

The promising results with *T. agrestis* prompted us to initiate NMR-spectroscopic screening of venom from a large selection of spider species, both to determine the extent to which sulfated nucleosides might occur among the various spider taxa and to obtain an expanded and unbiased view of spider venom composition in general.

## Results

Spider venom samples representing >70 species from 27 families (9, 10) were subjected to direct NMR spectroscopic analysis. For each species, one or more lyophilized samples corresponding to 5–10 μl of crude venom were dissolved, without purification,<sup>††</sup> in D<sub>2</sub>O, and <sup>1</sup>H NMR spectra were recorded. These <sup>1</sup>H spectra were used to assess the overall composition of nonvolatiles in each venom sample. In a number of instances, the <sup>1</sup>H NMR spectrum immediately suggested that simple amino acids and amino acid derivatives typically found in spider venom represented the major components. For example, in the <sup>1</sup>H spectrum of *Pisaura mirabilis* venom, signals representing >90% of the small-molecule content of the venom could be assigned to glutamic acid, aspartic acid, 4-aminobutyric acid, and spermidine (Fig. 2). For venoms of this type, structural assignments were confirmed through NMR-spectroscopic mixing experiments using authentic samples of the identified compounds. The

Author contributions: F.C.S., A.E.T., M.G., T.E., and J.M. designed research; F.C.S., A.E.T., M.G., R.U.M., and J.B.G. performed research; F.C.S., A.E.T., M.G., R.U.M., and J.B.G. analyzed data; and F.C.S., M.G., and J.M. wrote the paper.

The authors declare no conflict of interest.

<sup>†</sup>To whom correspondence may be addressed. E-mail: circe@cornell.edu or fs31@cornell.edu.

<sup>§</sup>Present address: DuPont Crop Protection, Stine–Haskell Research Center, Newark, DE 19711.

<sup>††</sup>In some instances, centrifugation was necessary to remove insoluble material. In addition, it cannot be excluded that in some cases lyophilization caused loss of volatile venom components; however, for the present study, limited availability of nonlyophilized venom samples precluded additional analyses aimed at volatile components, for example, via GC-MS.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0806840105/DCSupplemental](http://www.pnas.org/cgi/content/full/0806840105/DCSupplemental).

© 2008 by The National Academy of Sciences of the USA

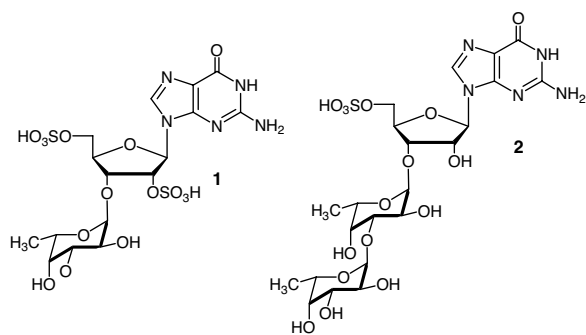


Fig. 1. Sulfated nucleosides from the hobo spider, *Tegenaria agrestis*.

$^1\text{H}$  NMR spectra of several other spider venom samples, for example, *Dysdera* sp., *Phyxioshema* sp., and *Scytodes fusca*, showed primarily broad, unresolved peaks characteristic of large peptides and proteins but were virtually devoid of any significant signals representing small molecules. A summary of the small molecules we identified in our venom library is given in Table 1.

Venom samples of more interesting composition that could not be characterized based solely on their  $^1\text{H}$  NMR spectra were analyzed further by using 2D NMR spectroscopy, including dqfCOSY, NOESY, HMQC, and HMBC spectra. For the venom samples analyzed in this way, the resulting sets of spectra provided sufficient signal dispersion and connectivity information to assign all major signals to discrete structural fragments, which, in many cases, allowed immediate identification of complete structures. For example, the highly convoluted  $^1\text{H}$  NMR spectrum of *Coelotes pastoralis* venom (11) clearly indicates that this venom represents a complex mixture (Fig. 3). However, based on the corresponding dqfCOSY spectrum, structures could be easily proposed for the venom's major components as well as for many minor components. Among the minor components, we identified a sulfated nucleoside, whose presence was suggested by cross-peaks at 5.35 ppm/6.13 ppm, 5.35 ppm/4.74 ppm, and 4.36 ppm/4.42 ppm that appeared to belong to the 2' and 5' protons of a ribose spin system, which are characteristically shifted downfield in sulfated derivatives (5). Further analysis using HMQC and HMBC spectra, as well as negative-ion electrospray-ionization mass spectrometry (ESI-MS) and UV spectroscopy, confirmed the presence of the previously undescribed disulfated guanosine derivative 3 [Fig. 4 and supporting information (SI) Fig. S1 and Table S1]. In addition, several compounds commonly found in spider venom were identified, including citric acid, isocitric acid, cytosine,  $\gamma$ -amino butyric acid, and octopamine.

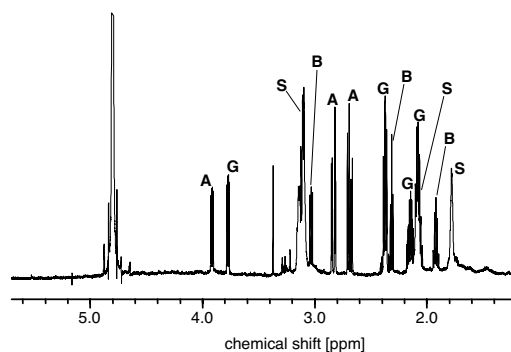


Fig. 2. The  $^1\text{H}$  NMR spectrum of *P. mirabilis* venom. Signals corresponding to glutamic acid, aspartic acid, 4-aminobutyric acid, and spermidine are labeled G, A, B, and S, respectively.

Like many of the venoms investigated, *C. pastoralis* venom was found to contain significant amounts of acylated polyamines, a large family of natural products that has been extensively studied (7). The presence of acyl polyamines in our unpurified venoms was easily recognized based on the characteristic spin systems of their aromatic head groups, whose connections to the polyamine chains were evident from the HMBC spectra. Polyamines frequently occurred as mixtures of several structurally very similar compounds, and as a result, overlap of the corresponding NMR-spectroscopic signals in both 1D and 2D spectra did not always permit complete characterization of individual compounds. Nonetheless, NMR-spectroscopic analyses enabled elucidation of the general structural features of the polyamines present in a sample, which, in cases where the spectra suggested the presence of unusual structural features, were investigated further. In the case of *C. pastoralis*, two major as well as several minor acylated polyamines were detected in the  $^1\text{H}$  NMR spectrum. Analysis of the 2D NMR spectra yielded partial structures for the two major polyamines (Tables S2 and S3). Both compounds appeared to possess a 2,5-dihydroxybenzoylamino-propylamine head group and terminate with a 4-guanidinobutylamine (Agmatine) moiety. Complete characterization of the intervening diaminoalkane segments was not possible, however, because of multiple signal overlap. Subsequent analysis of the sample by ESI-MS-MS nevertheless allowed the remainder of both structures, 5 and 6, to be determined (Figs. S2 and S3). For these analyses, the molecular  $[\text{M} + 1]^+$  peaks corresponding to polyamines 5 and 6 were identified based on the structural characteristics previously identified in the NMR spectra. For example, collisionally induced fragmentation of a terminal agmatine moiety produces a characteristic fragment ion having a mass-to-charge ( $m/z$ ) ratio of 114 (12). A tandem experiment scanning for those precursor ions giving rise to a fragment of  $m/z = 114$  revealed two major components in the venom sample, whose abundance, elemental composition, and molecular mass indicated that they correspond to the two major polyamines observed in the NMR spectra. Complete structures could then be confirmed based on the predictable fragmentation patterns (12) observed in daughter scans acquired for the selected molecular  $[\text{M} + 1]^+$  peaks. Aside from these two compounds from *C. pastoralis*, the major acylated polyamines present in our venom samples appeared to belong to known types and were not characterized in detail.

Among the most interesting spiders investigated in the course of this survey were three species from the genus *Loxosceles*: *Loxosceles arizonica* (13), *Loxosceles deserta* (14), and the well known brown recluse, *Loxosceles reclusa* (13). *Loxosceles* spiders are notable for the severe necrotic lesions that may result from their bites to humans (15). Surprisingly, we found sulfated nucleosides to comprise by far the most abundant components in the venoms of all three species. Our NMR-spectroscopic analyses revealed that all three *Loxosceles* species contained 2,5-disulfated guanosine (3) as the major component, in addition to small amounts of what appeared to be 2-sulfated guanosine (4) (Fig. 5 and Table S4). Because the concentrations of the monosulfated derivative 4 in *Loxosceles* spp. venom were low, and the amounts of venom we had available were limited, we were unable to characterize 4 fully via HSQC and HMBC spectra. Therefore, a sample of 4 was synthesized via nonselective, partial sulfation of a large excess of unprotected guanosine. NMR-spectroscopic analysis of the resulting mixture of monosulfated guanosine derivatives confirmed our structural assignment of 4 as the 2-sulfated derivative. The presence of the guanine moiety in 3 and 4 was further confirmed by UV spectroscopy (Fig. S1) and via negative-ion ESI-MS (Tables S1 and S4), which showed peaks corresponding to  $[\text{M} - 1]^-$  at  $m/z$  442.0 and 362.0, respectively.

Table 1. Small molecules detected in spider venoms organized by family

Family	Species	Small molecules	Family	Species	Small molecules
Agelenidae	<i>Agelena gracilens</i>		Philodromidae	<i>Thanatus striatus</i>	e
	<i>Agelena labyrinthica</i>	i	Pisauridae	<i>Dolomedes tenebrosus</i>	b,d,e
	<i>Agelena orientalis</i>	i		<i>Dolomedes gertschi</i>	b,d,e
	<i>Agelenopsis aperta</i> (21, 22)	d,i		<i>Pisaura mirabilis</i>	b,e
	<i>Agelenopsis pennsylvania</i>		Plectreuridae	<i>Plectreurys tristis</i> (23)	e
	<i>Benoitia tadhika</i>			Salticidae	<i>Phidippus ardenus</i>
	<i>Calilena</i> sp.		<i>Phidippus johnsoni</i>		
	<i>Hololena curta</i> (6, 21)	d,h		<i>Phidippus</i> sp.	d,f
	<i>Loxosceles arizonica</i>	h	Scytodidae	<i>Scytodes fusca</i>	
	<i>Loxosceles deserta</i>	h		Segestriidae	<i>Ariadna</i> sp.
<i>Loxosceles reclusa</i>	h	Sparassidae	<i>Delena cancerides</i>		d,e,g,i
<i>Tegenaria agrestis</i> (5)	d,h		<i>Eusparassus oculatus</i>	a,b,e,f,g,i	
<i>Tegenaria domestica</i>	b,d,g,i		<i>Holconia flindersi</i>	b,d,e	
Amaurobiidae	<i>Coelotes pastoralis</i>	d, h		<i>Isopedella canberrensis</i>	d,e
	<i>Paracoelotes birulai</i> (21)	d,i		<i>Isopedella leai</i>	b,d
Araneidae	<i>Aculepeira</i> sp.	d,i		<i>Micrommata virescens virvirescens</i>	a,ed
	<i>Araneus diadematus</i>	c,d,e		<i>Neosparassus diana</i>	b,d,e
	<i>Araneus</i> sp.	e		<i>Olios</i> sp.	e
	<i>Araneus tartaricus</i>	d,e	Theraphosidae	<i>Aphonopelma seemani</i>	a,b,d,i
	<i>Argiope lobata</i>			<i>Aphonopelma</i> sp.	b,d,i
	<i>Eriophora edax</i>	d,e		<i>Avicularia avicularis</i>	b,e,i
	<i>Larinioides cornutus</i>	d,e,i		<i>Ceratogyrus marshalli (cornuatus)</i>	b,d,g
	<i>Larinioides</i> sp.	e,i		<i>Chilobrachys</i> sp.	a,d,i
		<i>Citharischius crawshayi</i>		d,e	
Clubionidae	<i>Cheiracanthium punctorium punctorium</i>				
Cyrtachenidae	<i>Anemesia incana</i>	e		<i>Haplopelma albostratum</i>	a,d,e,i
Dipluridae	<i>Phyxioshema</i> sp.			<i>Haplopelma lividum</i>	a,d,e,f,i
Dysderidae	<i>Dysdera</i> sp.			<i>Heteroscodra maculata</i>	B
Eresidae	<i>Eresus</i> sp.	e		<i>Ornithoctonus huwena</i>	a,b,d,e
	<i>Stegodyphus</i> sp.	e		<i>Phormictopus cancerides</i>	b,d
Filistatidae	<i>Kukulcania</i> sp.	b		<i>Psalmopoeus cambridgei</i>	b,d,g
Gnaphosidae	<i>Drassodes</i> sp.	e		<i>Selenocosmia hainana</i>	a,b,d,e,i
Hexathelidae	<i>Macrothele</i> sp. (24)	b,i	Theridiidae	<i>Latrodectus tadhicus</i>	i
Lycosidae	<i>Rabidosa</i> sp.	d,e			<i>Steatoda giorra</i>
	<i>Hogna</i> sp.	d,e,i		<i>Steatoda paykulliana</i>	b,e,i
	<i>Lycosa proegrandis</i>	d,e,f	Thomisidae	<i>Misumena vatia</i>	b,d,i
Nemesiidae	<i>Raveniola micrura</i>	b,d,e,g,i		<i>Thomisus onustus</i>	b,i
Oecobiidae	<i>Urotea</i> sp.		Titanocidae	<i>Titanoeca</i> sp.	d,e
Oxyopidae	<i>Peucetia viridans</i>	e,i		Zodariidae	<i>Lashesana</i> sp.
Palpimanidae	<i>Palpimanus</i> sp.	d			

a, acetyl choline; b,  $\gamma$ -aminobutyric acid; c, choline; d, citric acid; e, glutamic acid; f, histamine; g, octopamine; h, sulfated nucleosides; i, other nucleosides.

## Discussion

A well known limitation of NMR spectroscopy is its relatively poor sensitivity. Despite continuing advances in probe design and available magnet field strength, NMR spectroscopic sensitivity is often orders of magnitude lower than that of MS (16). Nonetheless, our analyses, which in most cases had to rely on samples of only 5–25  $\mu$ l of crude venom, allowed detailed NMR-spectroscopic characterization of most components accounting for more than  $\approx$ 1–5% of the small-molecule content of the venom samples. For example, as shown in Fig. 3, most of the signals visible in the dqfCOSY spectrum of *C. pastoralis* could be assigned to distinct spin systems. The threshold for detection necessarily varied for different compound classes and depended on the size and complexity of the venom samples. For example, whereas 1% of a sulfated ribonucleoside derivative could have been detected even in the most dilute venom samples investigated, the presence of as much as 5% of an unusually functionalized polyamine might not have been detected in venoms containing large amounts of other polyamines.

Our results show that direct NMR-spectroscopic analysis can be used to screen complex biological samples such as spider venom for the presence of small molecules. In combination with mass-spectrometric analyses and synthesis, this methodology enabled us to detect and identify compounds without purification. In the case of the sulfated guanosine derivatives that we identified in venoms obtained from *C. pastoralis* and *Loxosceles* spp., complete structures could be proposed based on the NMR data obtained for the entire venom, with subsequent confirmation via independent synthesis. A different situation was presented by the acylated polyamines detected in *C. pastoralis*, which occur as part of a larger family of structurally similar compounds. Here, signal overlap in the NMR spectra prevented full NMR-spectroscopic characterization of the individual components. However, the NMR spectroscopic analyses revealed important structural features which provided guidance in the rational pursuit of further analyses via MS leading to full identification.

Complementing MS analyses, NMR-spectroscopic screening seems particularly well suited for the survey of arrays of related samples, because the efficiency of analysis inevitably increases as



