Current status of structural studies of proteins originating from Arachnida

Safaa Kader[1,2], **Ricardo A. Hernandez**[1], **Kriti Khatri**[1], **Andrea O'Malley**[1], **Vojislava Grbic**[3], **Miodrag Grbic**[3,4], **Maksymilian Chruszcz**[1,*]

[1]Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA

[2]Department of Pathology and Forensic Medicine, University of Al-Nahrain, Baghdad, Iraq

[3]Department of Biology, Western University, London, Ontario, N6A 5B7, Canada

[4]University of La Rioja, Logrono, Spain

Abstract

Arthropods from class Arachnida constitute a large and diverse group with over 100,000 described species, and they are sources of many proteins that have a direct impact on human health. Despite the importance of Arachnida, few proteins originating from these organisms have been characterized in terms of their structure. Here we present a detailed analysis of Arachnida proteins that have their experimental structures determined and deposited to the Protein Data Bank (PDB). Our results indicate that proteins represented in the PDB are derived from a small number of Arachnida families, and two-thirds of Arachnida proteins with experimental structures determined are derived from organisms belonging to Buthidae, Ixodidae, and Theraphosidae families. Moreover, 90% of the deposits come from just a dozen of Arachnida families, and almost half of the deposits represent proteins originating from only fifteen different species. In summary, our analysis shows that the structural analysis of proteins originating from Arachnida is not only limited to a small number of the source species, but also proteins from this group of animals are not extensively studied. However, the interest in Arachnida proteins seems to be increasing, which is reflected by a significant increase in the related PDB deposits during the last ten years.

Keywords

Arachnida; proteins; structural biology; toxins; biomaterials; allergens

Introduction

Arthropods from class Arachnida constitute a large and very diverse group with over 100,000 described species (Chapman 2009). They include a large number of species that are important from a perspective of human health and economy. Arachnida are also sources of many proteins that have a direct impact on human health. For example, mites are an

*To whom correspondence should be addressed: Maksymilian Chruszcz, Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208 USA; chruszcz@mailbox.sc.edu; Tel: (803) 777-7399, FAX: (803) 777-9521.

important group of Arachnida that is relevant to human health, as they not only can transmit some diseases (Vogel *et al.* 2014), but can also be significant sources of allergens (Tehri & Gulati 2015). Over 10% of all registered allergens originate from eleven mite species, with *Dermatophagoides farinae* and *Dermatophagoides pteronyssinus* contributing to 69 allergens (Vogel *et al.* 2014; Azmiera et al. 2020).

Scorpions (order Scorpiones) are extensively studied due to their ability to produce toxins. There are over fifteen hundred scorpion species with a broad distribution throughout the world (Garcia et al. 2013; Chippaux & Goyffon 2008). Scorpion envenoming represents a public health problem, mostly in tropical countries. Yearly, more than one million cases of scorpion envenomation are recorded with a mortality risk of about 3% worldwide (Chapman 2009).

Spiders are another group of Arachnids that are studied not only for their toxins, but also for their fibroins, which are considered to be very promising biodegradable materials (Vidya & Rajagopal 2021). Therefore, a significant effort is made to study this group of proteins.

Lastly, ticks are also recognized as medical and economical burdens because of their ability to transfer diseases to humans and animals. Therefore, ticks are considered to be one of the biggest problems in public and veterinary health (Barker & Murrell 2004). These ecto-parasites can affect the production of the animals and their health, either directly by transmitting viruses, bacteria, rickettsiae, and protozoa, or by the effect of their bites. During the evolution of their saliva, ticks gained an enormous diversity of bioactive compounds such as anticoagulants, chemokine-binding proteins, platelet aggregation, and complement system inhibitors (Rajput et al. 2006; Chapman 2009). Many enzyme inhibitors, such as serine protease inhibitors (SPIs), have been described. SPIs' importance in several tick biological processes has been shown to be directly involved in the regulation of inflammation, wound healing, vasoconstriction, blood clotting, and the modulation of host defense mechanisms (Blisnick et al. 2017; Denisov et al. 2021).

Despite the importance of Arachnida, very few proteins originating from these organisms have been characterized structurally. Therefore, we performed a detailed analysis of Arachnida proteins that have their structures determined experimentally and deposited to the Protein Data Bank (PDB) (wwPDB consortium 2019). The PDB not only provides information on three dimensional structures of proteins, but also annotates all deposits, which allows for sophisticated analysis of the entries. For example, it is possible to analyze deposits taking into account protein sources, year of deposition or experimental methods used for structural studies. Our analysis provides a summary of Arachnida families and species that are sources of the most extensively studied proteins.

Materials and Methods

Selection of proteins originating from Arachnida and construction of the dataset used for analysis

The PDB deposits corresponding to proteins originating from Arachnida were selected using the "Advanced Search", "Polymer Molecular Features", and "Source Organism Taxonomy

Name (Full Lineage)" search options from the database interface. The PDB deposits selected for this analysis were chosen using the search query "Arachnida". This search in June 2021 resulted in 492 available PDB deposits. The PDB codes corresponding to the entries of interest were saved together with various information on the experimental procedures used to determine the protein structures. Moreover, we have collected information on the year of deposition, protein name, and protein family and function, as well as taxonomic information on the source organism. If the required information was not stored by the PDB, we have derived it from UniProt (UniProt consortium 2021) and/or relevant manuscripts associated with the deposit. All derived data were summarized in Supplementary Table S1.

Results

Overview of currently available experimental structures

Until June 2021, the PDB contained 492 entries of proteins originating from Arachnida (Fig. 1A), which corresponded to less than 0.5% of all deposited structures. Among the 492 entries, 44 structures represent proteins from mites, 192 structures belong to scorpions, 164 structures correspond to spiders, and 92 structures are from ticks (Fig. 1B). Almost two-thirds of the deposits correspond to proteins originating just from three families (Fig. 2A), namely Buthidae, Ixodidae, and Theraphosidae. At the same time, over 90% of the deposits can be attributed to twelve Arachnida families (Fig. 2A). Approximately half of the deposits (47%) represent only fifteen different species, and the top ten species are responsible for 180 of 492 deposits (Fig 2B). The experimental data of the structural models of these PDB entries are derived from the three major methods: X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR), and electron microscopy (EM). NMR was used to determine most of the deposited proteins (286), while X-ray crystallography and cryo-EM were used to determine 192 and 14 structures respectively (Fig. 1C). In terms of all PDB deposits, only 7% of structures were determined using NMR and 87% were determined using X-ray diffraction (Schiro *et al.* 2020). These differences in the preferred experimental approaches used between the whole PDB and entries originating from Arachnida clearly indicate that these proteins are relatively small and difficult to crystallize.

Mites

There are only 44 structures deposited to the PDB that represent proteins originating from mites. The majority of these structures correspond to the proteins from mites belonging to Sarcoptiformes (Fig. 3A). The remaining structures represent proteins originating from two orders, Trombidiformes and Mesostigmata. In fact, all of the proteins with experimental structures from Trombidiformes originate from *Tetranychus urticae*, while the two Mesostigmata proteins structures originate from *Varroa destructor* (Fig. 3B). These two mite species are important from an agricultural point of view, as *T. urticae* is a major plant pest and *V. destructor* is a mite parasite of bees. Sarcoptiformes proteins with structures determined originate from house dust mites (*Dermatophagoides farinae* and *Dermatophagoides pteronyssinus*), storage mite (*Blomia tropicalis*), and itch mite (*Sarcoptes scabiei*) (Fig. 3B). The listed Sarcoptiformes mites are detrimental to human health, as they are sources of allergens or parasites (*S. scabiei*).

Majority of mite proteins that have their structures determined (Table S1) are classified as allergens belonging to Groups 1 (cysteine proteases) and Group 2 (Niemann-Pick proteins type C2 (NPC2)) (Derewenda et al. 2002, Ichikawa et al. 2005, Chruszcz et al. 2009). However, recently there is a notable increase in number of proteins' structures that originate from agricultural pests, and that are responsible for xenobiotic detoxification (Schlachter et al. 2017; Schlachter et al. 2019; Daneshian et al. 2022) or Odorant-binding proteins (OBPs) (Amigues et al. 2021). Both NPC2s and OBPs may participate in chemical communication of arthropods.

Scorpions

There is a total of 192 protein structures belonging to the Scorpiones order deposited to the PDB between 1993–2021, most of which were characterized by NMR. Within the Scorpiones order, 182 structures correspond to proteins originating from Buthidae, while the other structures represent proteins stemming from two families, Scorpionidae (nine structures) and Hormuridae (one structure). Most of the structures are originating from *Mesobuthus* genus while the remaining structures originate from *Centruroides* (31 structures), *Leiurus* (22), *Androctonus* (18), *Tityus* (16), and other families (Fig. 4B).

Over 90% of the structure of proteins from scorpions deposited to the PDB (Table S1) are described as inhibitors of ion channels, with most (77%) being designated as sodium or potassium channels inhibitors (Norton & Chandy 2017).

Spiders

There are 164 protein structures deposited to the PDB originating from spiders. Among these, the majority belong to the family Theraphosidae (Fig. 5A). The rest of the structures correspond to proteins from families such as Araneidae (18 structures), Agelenidae (13), Pisauridae (10), Atracidae (10), Sicariidae (9), and Lycosidae (5). Looking at the category of these proteins, the number of structures related to toxins that spiders producing is the largest (130 deposits).

Similarly, as in the case of the PDB deposits of proteins originating from scorpions, the majority of proteins originating from spiders that have their structures determined (Table S1) are classified as neurotoxins and ion channel inhibitors.

Ticks

There were 92 PDB deposits representing proteins originating from ticks. Among the three families in Ixodida (Metastigmata) suborder of Acari, the majority of the deposits (72 structures) corresponded to the Ixodidae (hard ticks) family, and the remainder (20) belong to the Argasidae (soft ticks) family (Fig. 6A). No PDB deposits of proteins belonging to the Nuttalliellidae family are reported.

Among the Ixodidae family, most protein structures come from *Ixodidae scapularis* (black-legged tick or deer tick) followed by *Ixodidae ricinus* (castor bean tick). Both *I. scapularis* and *I. ricinus* are causative agents (vectors) of Lyme disease, which affects 476,000 people per year in the USA (Barker & Murrell 2004; Schwartz *et al.* 2021; Kugeler et al. 2021),

(Fig. 6). The remaining structures from the Ixodidae family belong to *Rhipicephalus appendiculatus* (14 structures), *R. bursa* (9), *R. sanguineus* (9). *R. microplus* (3), *R. pullchellus* (2), *Amblyomma variegatum* (2), *Haemaphysalis longicornis* (2), *Dermacentor andersoni* (1) and *A. maculatum* (1). Among the Argasidae tick family, most structures belong to *Ornithodoron moubta* (13), followed by *Argas monolakensis* (4) and *Argas reflexus* (2), (Fig. 6B).

Tick proteins that have their structures determined (Table S1) are more diverse in terms of their function in comparison with previously mentioned mite, scorpion or spider proteins. Significant fraction of the tick proteins can be classified as lipocalins, protease inhibitors (e.g. cystatins and serpins), proteases, as well as proteins like evasins, which play role in suppression of the host inflammatory response (Bhusal et al. 2020).

Discussion

Our analysis shows that proteins originating from Arachnida are not extensively studied in terms of their three-dimensional structures. However, in the last decade, the proteins derived from this group are more actively studied, with over 50% of all deposits being submitted to the PDB in the 2011–2021 period (Fig. 1A). Interestingly, 58% of all deposits originating from Arachnida correspond to protein structures that were determined using NMR (Fig. 1C), which is significantly more than for the whole PDB, where only 7% of protein structures are determined using this experimental technique. This extensive use of NMR can also be explained by the fact that these proteins are relatively small, due (generally less than 10 kDa) and have a high degree of conformational flexibility, which hinders crystallization. Another striking observation is related to a small number of Arachnida families from which proteins represented in the PDB are derived. As mentioned, two-thirds of Arachnida proteins with experimental structures originated from Buthidae, Ixodidae, and Theraphosidae families (Fig. 2A). Moreover, 90% of the deposits come from just a dozen Arachnida families, and almost half of the deposits represent proteins originating from only fifteen different species (Fig 2B). The small number of species from which the PDB deposits originate is contrasted with the size of class Arachnida, with approximately 100,000 species described (Laustsen et al. 2016). This clearly indicates that structural studies are at a very early stage and are extremely limited.

There are only 44 structures of mite proteins deposited to the PDB; in recent years, these Arachnida are attracting more attention for several reasons. First of all, mites are sources of very potent allergens that originate mainly from house dust mites (*D. farinae* and *D. pteronyssinus*) and storage mites (e.g. *B. tropicalis*) (Fig. 3B) (Thomas 2015). The experimental models of mite proteins deposited to the PDB are dominated by representatives of two protein families. For example, there are multiple deposits of Group 1 (cysteine proteases) and Group 2 (Niemann-Pick C2 protein family) house dust mite allergens, which constitute almost half of all mite-related PDB entries. Interestingly, several structures of proteins from Group 1 and Group 2 of house dust mite allergens correspond to complexes of the allergens with monoclonal antibodies (Chruszcz et al. 2012; Glesner et al. 2017; Glesner et al. 2019; Osinski et al. 2015). These structures provide insights into the antigenic structure of the allergens and are used to develop a new approach to immunotherapy. Together, the

PDB entries associated with mite allergens represent approximately three-quarters of all mite protein structures. Secondly, some of the mites (e.g. spider mites like *T. urticae*) are of interest from an agricultural point of view, as they show one of the highest incidences of pesticide resistance (Van Leeuwen et al. 2015). Moreover, *T. urticae* is an attractive experimental model system (Grbic et al. 2007). This is illustrated by the increase of PDB deposits for proteins originating from *T. urticae* (Schlachter et al. 2017; Schlachter et al. 2019; Daneshian et al. 2022). These deposits provide structural insights into enzymes that are used by spider mites to detoxify various plant metabolites. In this case, the structural studies aim to provide information on the physiological functions of the enzymes, to be used toward development of new generations of acaricides.

Scorpions are of medical relevance since they are the main species with venom that may cause harm to humans. Currently, it is clear that the toxins (ion channel inhibitors) derived from Buthidae species are of the biggest interest from the perspective of structural studies, as they correspond to 37% of all Arachnida-associated deposits. In this case, the biomedical importance of the venoms is driving the researchers' interest in these peptides and proteins. The main toxic effects of venoms are mainly derived from the inhibitory actions of neurotoxins, primarily sodium and potassium channel targeting neurotoxins. Proteins with sodium or potassium inhibitor channel activity account for most protein structures compared to proteins with enzyme inhibitor activity, short antimicrobial peptide activity, or defense response activity. In fact, almost 77% of all scorpion-related PDB deposits are derived from proteins from potassium and sodium channel inhibitor families. This is stemming from the interest in using toxins for drug applications and discovering crystal structures of receptor sites for toxins (Laustsen et al. 2016). This structural data is an excellent foundation for developing new antivenoms. On the other hand, these entries also provide information on the protein structure of substances in scorpion venoms that are important in the medical field (Laustsen et al. 2016; Abdel-Rahman et al. 2015; Lewis & Garcia 2003). For instance, chlorotoxins (PDB codes: 1CHL, 5L1C, and 6ATW) are of special importance due to their selective attachment to glioma cells, serving as a potential therapeutic option for brain gliomas (Cordeiro et al. 2015; Barker & Murrell 2004; Fletcher et al. 1997). Also, the cystine-dense peptide of (PDB code: 6AY7) could be applied to concentrate arthritis drugs in joints after conjugating them with a steroid (Cook Sangar et al. 2020). The experimental models of scorpion proteins deposited into the PDB represent the largest group among the other species of Arachnida.

Spiders are the second most important source of proteins that have their structures deposited in the PDB. Here proteins originating from Theraphosidae dominate; however, among the twelve most represented Arachnida families, there are five additional spider families (Araneidae, Agelenidae, Pisauridae, Atracidae, and Sicariidae). It clearly shows that the researchers are interested in many spider families, and they are driven by studies of venoms and spidroins. Similarly to scorpions, scientific interest in spider venom originates from the ability of these proteins to modulate ion channels and their therapeutic potentials (Chapman 2009; Cardoso et al. 2015). Spider venoms are composed of three main classes of components: inorganic and organic components, small molecular polypeptides, and high molecular proteins. Among these classes, polypeptides of a molecular weight between 3–8 kDa with neurotoxic properties are of particular interest to the scientific community

(Escoubas et al. 2000). These toxins belong to the group of small, disordered proteins that are difficult to crystallize. Therefore, it is not a surprise that NMR stands out as the major technique used to solve the structure of proteins originating from spiders. Furthermore, as detailed, toxins are of special interest due to their ability to modulate ion channels. In this sense, cryo-EM is the preferable technique to determine the structure of the complex between the toxin and the ion channel. In fact, seven out of the eight toxin-ion channel complexes in PDB, were solved using this novel experimental method. These proteins are of interest due to their remarkable properties and potential applications, such as insecticidal (Fletcher et al. 1997), antibacterial (Benli & Yigit 2008), and drug delivery and tissue regeneration, among many other biomedical applications (Bakhshandeh et al. 2021). While the first spider protein structure deposited to PDB dates from 1993, titled: NMR solution of toxin omega-Aga-IVB (PDB codes: 1OMA and 1OMB) (Yu et al. 1993), it is not until 2008 that the first structure of a protein unrelated to spider toxins was deposited. This deposit showed the repetitive domain of the egg case silk (Chapman 2009; Lin et al. 2009). Spidroin-related deposits (27 deposits) represent the second-most abundant group of spider proteins in the PDB. The remaining spider-related PDB deposits correspond to a miscellaneous group of polypeptides, such as arginine kinase and insulin-like growth factor-binding domain protein.

Ticks are the final group described here that contributes a significant fraction of the PDB deposits from Arachnida. In the case of this group, proteins derived from Ixodidae (hard ticks) are the most often studied, and the research generally concentrates on ticks' salivary proteins and their role in the transfer of various pathogens (Denisov & Dijkgraaf 2021). Given the significance of tick saliva in host immune defense modulation, these proteins exhibit crucial roles in binding interaction to host receptors (Kazimírová & Štibrániová 2013; Chmela et al. 2019). Among the tick proteins deposited in the PDB, transferases, endopeptidases, proteases, and lipocalins are the most common. Novel structures of lipocalins of tick-derived C5 inhibitors OmCI from *Ornithodoron moubta* (Argasidae family) and RaCI from *Rhipicephalus appendiculatus* (Ixodidae family) generate new insights on complement C5 activation in humans (PDB codes: 6RQJ and 6RPT). Interestingly, both free and complexed forms of these proteins with human complement C5 have been deposited, the comparison of which holds therapeutic significance. Structures of amine binding lipocalins monomine and monotonin from *Argasidae monolakensis* were determined to compare ligand binding sites for reconstructing evolutionary pathways of the amine binding tick lipocalins (PDB code: 3BRN). Tick saliva also causes significant allergic disease, as further interesting deposits from the Argasidae family come from the major allergen Arg r 1 from *Argas reflexus* (pigeon tick). A selenomethionine mutation-free Arg r 1 structure and Arg r 1 complexed with histamine structures are deposited in the PDB (PDB codes: 2X45 and 2X46). As interest in tick-transferred diseases is growing, it is expected that the salivary proteins will continue to be one of the major groups for Arachnida proteins represented in the PDB.

In summary, currently, the Arachnida proteins are of significant interest for biomedical, agricultural and material sciences, but are significantly understudied. Based on this review, the proteins or peptides studies are not very diverse and can be categorized mainly as toxins, spidroins, salivary proteins, and allergens (Fernández-Caldas et al. 2014; Wang et

al. 2014; Thomas 2015; Daly & Wilson 2018; Bhusal et al. 2020; Denisov & Dijkgraaf 2021). We conclude that structural studies of Arachnida-derived proteins are in their infancy, because of the importance of Arachnid proteins not only from the perspective of human health or biomaterials but also from potential application in agriculture. We also expect a rapid increase in resolution of various Arachnida protein structures, as these molecules have potential diverse applications.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abdel-Rahman MA, Harrison PL & Strong PN (2015) Snapshots of scorpion venomics. Journal of Arid Environments, 112, 170–176. 10.1016/j.jaridenv.2014.01.007

Amigues B, Zhu J, Gaubert A, Arena S, Renzone G, Leone P, Fischer IM, Paulsen H, Knoll W, Scaloni A, Roussel A, Cambillau C. & Pelosi P. (2021) A new non-classical fold of varroa odorant-binding proteins reveals a wide open internal cavity. Scientific Reports, 11, 13172. 10.1038/s41598-021-92604-2 [PubMed: 34162975]

Azmiera N, Mariana A, Pimsler ML & Heo CC (2020) Review of mites found on various animal hosts and at different localities in Malaysia. Journal of Medical Entomology, 57, 1354–1363. 10.1093/jme/tjaa086 [PubMed: 32440683]

Barker SC & Murrell A. (2004) Systematics and evolution of ticks with a list of valid genus and species names. Parasitology, 129 Suppl (7): S15–36. 10.1017/s0031182004005207 [PubMed: 15938503]

Bakhshandeh B, Nateghi S, Gazani M, Dehghani Z. & Mohammadzadeh F. (2021) A review on advances in the applications of spider silk in biomedical issues. International Journal of Biological Macromolecules, 192, 258–271. 10.1016/j.ijbiomac.2021.09.201 [PubMed: 34627845]

Benli M. & Yigit N. (2008) Antibacterial activity of venom from funnel web spider Agelena labyrinthica (Araneae: Agelenidae). Journal of Venomous Animals and Toxins including Tropical Diseases, 14, 641–650. 10.1590/S1678-91992008000400007

Bhusal RP, Eaton JRO, Chowdhury ST, Power CA, Proudfoot AEI, Stone MJ & Bhattacharya S. (2020) Evasins: tick salivary proteins that inhibit mammalian chemokines. Trends in Biochemical Sciences, 45, 108–122. 10.1016/j.tibs.2019.10.003 [PubMed: 31679840]

Blisnick AA, Foulon T. & Bonnet SI (2017) Serine protease inhibitors in ticks: an overview of their role in tick biology and tick-borne pathogen transmission. Frontiers in Cellular and Infection Microbiology, 7, 199. 10.3389/fcimb.2017.00199 [PubMed: 28589099]

Cardoso FC, Dekan Z, Rosengren KJ, Erickson A, Vetter I, Deuis JR, Herzig V, Alewood PF, King GF & Lewis RJ (2015) Identification and characterization of ProTx-III [mu-TRTX-Tp1a], a new voltage-gated sodium channel inhibitor from venom of the tarantula Thrixopelma pruriens. Molecular Pharmacology, 88, 291–303. 10.1124/mol.115.098178 [PubMed: 25979003]

Chapman AD (2009) Numbers of living species in Australia and the world. Australian Biodiversity Information Services.; Australian Biological Resources Study; Australia. Department of the Environment, Water, Heritage, and the Arts.

Chippaux JP & Goyffon M. (2008) Epidemiology of scorpionism: a global appraisal. Acta Tropica, 107, 71–79. 10.1016/j.actatropica.2008.05.021 [PubMed: 18579104]

Chmela J, Kotál J, Kovaíková A. & Kotsyfakis M. (2019) The use of tick salivary proteins as novel therapeutics. Frontiers in Physiology, 10, 812. 10.3389/fphys.2019.00812 [PubMed: 31297067]

Chruszcz M, Chapman MD, Vailes LD, Stura EA, Saint-Remy JM, Minor W. & Pomes A. (2009) Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. Journal of Molecular Biology, 386, 520–530. 10.1016/j.jmb.2008.12.049 [PubMed: 19136006]

Chruszcz M, Pomes A, Glesner J, Vailes LD, Osinski T, Porebski PJ, Majorek KA, Heymann PW, Platts-Mills TA, Minor W. & Chapman MD (2012) Molecular determinants for antibody binding on group 1 house dust mite allergens. Journal of Biological Chemistry, 287, 7388–7398. 10.1074/jbc.M111.311159 [PubMed: 22210776]

Cook Sangar ML, Girard EJ, Hopping G, Yin C, Pakiam F, Brusniak MY, Nguyen E, Ruff R, Gewe MM, Byrnes-Blake K, Nairn NW, Miller DM, Mehlin C, Strand AD, Mhyre AJ, Correnti CE, Strong RK, Simon JA & Olson JM (2020) A potent peptide-steroid conjugate accumulates in cartilage and reverses arthritis without evidence of systemic corticosteroid exposure. Science Translational Medicine, 12, 1041. 10.1126/scitranslmed.aay1041

Cordeiro FA, Amorim FG, Anjolette FA, Arantes EC, Anjolette FA & Arantes EC 2015. Arachnids of medical importance in Brazil: main active compounds present in scorpion and spider venoms and tick saliva. Journal of Venomous Animals and Toxins Including Tropical Diseases, 21, 24. 10.1186/s40409-015-0028-5 [PubMed: 26273285]

Daly NL & Wilson D. (2018) Structural diversity of arthropod venom toxins. Toxicon, 152, 46–56. 10.1016/j.toxicon.2018.07.018 [PubMed: 30040988]

Daneshian L, Renggli I, Hanaway R, Offermann LR, Schlachter CR, Hernandez Arriaza R, Henry S, Prakash R, Wybouw N, Dermauw W, Shimizu LS, Van Leeuwen T, Makris TM, Grbic V, Grbic M. & Chruszcz M. (2022) Structural and functional characterization of beta-cyanoalanine synthase from Tetranychus urticae. Insect Biochemistry and Molecular Biology, 142, 103722. 10.1016/j.ibmb.2022.103722

Denisov SS & Dijkgraaf I. (2021) Immunomodulatory proteins in tick saliva from a structural perspective. Frontiers in Cellular and Infection Microbiology, 11, 769574. 10.3389/fcimb.2021.769574

Denisov SS, Ippel JH, Castoldi E, Mans BJ, Hackeng TM & Dijkgraaf I. (2021) Molecular basis of anticoagulant and anticomplement activity of the tick salivary protein Salp14 and its homologs. Journal of Biological Chemistry, 297, 100865. 10.1016/j.jbc.2021.100865

Derewenda U, Li J, Derewenda Z, Dauter Z, Mueller GA, Rule GS & Benjamin DC (2002) The crystal structure of a major dust mite allergen Der p 2, and its biological implications. Journal of Molecular Biology, 318, 189–197. 10.1016/S0022-2836(02)00027-X [PubMed: 12054778]

Escoubas P, Diochot S. & Corzo G. (2000) Structure and pharmacology of spider venom neurotoxins. Biochimie, 82, 893–907. 10.1016/s0300-9084(00)01166-4 [PubMed: 11086219]

Fernández-Caldas E,· Puerta L. & Caraballo L. (2014) Mites and Allergy. Chemical Immunology and Allergy, 100, 234–242. 10.1159/000358860 [PubMed: 24925403]

Fletcher JI, Smith R, O'Donoghue SI, Nilges M, Connor M, Howden ME, Christie MJ & King GF (1997) The structure of a novel insecticidal neurotoxin, omega-atracotoxin-HV1, from the venom of an Australian funnel web spider. Nature Structural Biology, 4, 559–566. 10.1038/nsb0797-559 [PubMed: 9228949]

Garcia HH, Tanowitz HB & Del Brutto OH (2013) Neuroparasitology and tropical neurology. Handbook of Clinical Neurology, 114, 2–414.

Glesner J, Kapingidza AB, Godzwon M, Offermann LR, Mueller GA, DeRose EF, Wright P, Richardson CM, Woodfolk JA, Vailes LD, Wunschmann S, London RE, Chapman MD, Ohlin M, Chruszcz M. & Pomes A. (2019) A Human IgE antibody binding site on Der p 2 for the design of a recombinant allergen for immunotherapy. Journal of Immunology, 203, 2545–2556. 10.4049/jimmunol.1900580

Glesner J, Vailes LD, Schlachter C, Mank N, Minor W, Osinski T, Chruszcz M, Chapman MD, & Pomes A. (2017) Antigenic determinants of Der p 1: specificity and cross-reactivity associated with IgE antibody recognition. Journal Immunology, 198, 1334–1344. 10.4049/jimmunol.1600072

Grbic M, Khila A, Lee KZ, Bjelica A, Grbic V, Whistlecraft J, Verdon L, Navajas M, and Nagy L. (2007) Mity model: Tetranychus urticae, a candidate for chelicerate model organism. Bioessays, 29, 489–496. 10.1002/bies.20564 [PubMed: 17450600]

Ichikawa S, Takai T, Inoue T, Yuuki T, Okumura Y, Ogura K, Inagaki F. & Hatanaka H. (2005) NMR study on the major mite allergen Der f 2: its refined tertiary structure, epitopes for monoclonal antibodies and characteristics shared by ML protein group members. Journal of Biochemistry, 137, 255–263. 10.1093/jb/mvi039 [PubMed: 15809326]

Kazimírová M. & Štibrániová I. (2013) Tick salivary compounds: their role in modulation of host defences and pathogen transmission. Frontiers in Cellular and Infection Microbiology, 3, 43. 10.3389/fcimb.2013.00043 [PubMed: 23971008]

Kugeler KJ, Schwartz AM, Delorey MJ, Mead PS & Hinckley AF (2021) Estimating the frequency of lyme disease diagnoses, United States, 2010–2018. Emerging Infectious Diseases, 27, 616–619. 10.3201/eid2702.202731 [PubMed: 33496229]

Laustsen AH, Sola M, Jappe EC, Oscoz S, Lauridsen LP & Engmark M. (2016) Biotechnological trends in spider and scorpion antivenom development. Toxins, 8, 226. 10.3390/toxins8080226 [PubMed: 27455327]

Lewis RJ & Garcia ML (2003) Therapeutic potential of venom peptides. Nature Reviews Drug Discovery, 2, 790–802. 10.1038/nrd1197 [PubMed: 14526382]

Lin Z, Huang W, Zhang J, Fan JS & Yang D. (2009) Solution structure of eggcase silk protein and its implications for silk fiber formation. Proceedings of the National Academy of Sciences of the United States of America, 106, 8906–8911. 10.1073/pnas.0813255106 [PubMed: 19458259]

Norton RS & Chandy KG (2017) Venom-derived peptide inhibitors of voltage-gated potassium channels. Neuropharmacology, 127, 124–138. 10.1016/j.neuropharm.2017.07.002 [PubMed: 28689025]

Osinski T, Pomes A, Majorek KA, Glesner J, Offermann LR, Vailes LD, Chapman MD, Minor W. & Chruszcz M. (2015) Structural analysis of Der p 1-antibody complexes and comparison with complexes of proteins or peptides with monoclonal antibodies. Journal of Immunology, 195, 307–316. 10.4049/jimmunol.1402199

Rajput ZI, Hu SH, Chen WJ, Arijo AG & Xiao CW (2006) Importance of ticks and their chemical and immunological control in livestock. Journal of Zhejiang University Science-B, 7, 912–21. 10.1631/jzus.2006.B0912 [PubMed: 17048307]

Schiro A, Carlon A, Parigi G, Murshudov G, Calderone V, Revera E. & Luchinat C. On the complementarity of X-ray and NMR data. Journal of Structural Biology: X, 4, 100019. 10.1016/j.yjsbx.2020.100019

Schlachter CR, Daneshian L, Amaya J, Klapper V, Wybouw N, Borowski T, Van Leeuwen T, Grbic V, Grbic M, Makris TM & Chruszcz M. 2019. Structural and functional characterization of an intradiol ring-cleavage dioxygenase from the polyphagous spider mite herbivore Tetranychus urticae Koch. Insect Biochemistry and Molecular Biology, 107, 19–30. 10.1016/j.ibmb.2018.12.001 [PubMed: 30529144]

Schlachter CR, Klapper V, Wybouw N, Radford T, Van Leeuwen T, Grbic M. & Chruszcz M. (2017) Structural characterization of a eukaryotic cyanase from Tetranychus urticae. Journal of Agricultural Food and Chemistry, 65, 5453–5462. 10.1021/acs.jafc.7b01333

Schwartz AM, Kugeler KJ, Nelson CA, Marx GE & Hinckley AF (2021) Use of commercial claims data for evaluating trends in Lyme disease diagnoses, United States, 2010–2018. Emerging Infectious Diseases, 27, 499–507. 10.3201/eid2702.202728 [PubMed: 33496238]

Tehri K. & Gulati R. (2015) Might of the mite: a review. International Journal of Current Science, 15, 1–14.

Thomas WR (2015) Hierarchy and molecular properties of house dust mite allergens. Allergology International, 64, 304–311. 10.1016/j.alit.2015.05.004 [PubMed: 26433526]

UniProt Consortium. (2021) UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research, 49, D480–D489. 10.1093/nar/gkaa1100 [PubMed: 33237286]

Van Leeuwen T, Tirry L, Yamamoto A, Nauen R. & Dermauw W. (2015) The economic importance of acaricides in the control of phytophagous mites and an update on recent acaricide mode of

action research. Pesticide Biochemistry and Physiology, 121, 12–21. 10.1016/j.pestbp.2014.12.009 [PubMed: 26047107]

Vidya M. & Rajagopal S. (2021) Silk fibroin: a promising tool for wound healing and skin regeneration. International Journal of Polymer Science, 2021, 9069924. 10.1155/2021/9069924

Vogel P, Bosco SM & Ferla NJ (2015) Mites and the implications on human health. Nutricion Hospitalaria, 31, 944–951. 10.3305/nh.2015.31.2.7772

Wang S, Huang W. & Yang D. (2014) Structure and function of C-terminal domain of aciniform spidroin. Biomacromolecules, 15, 468–477. 10.1021/bm401709v [PubMed: 24422432]

wwPDB Consortium. (2019) Protein Data Bank: the single global archive for 3D macromolecular structure data. Nucleic Acids Research, 47, D520–D528. 10.1093/nar/gky949 [PubMed: 30357364]

Yu H, Rosen MK, Saccomano NA, Phillips D, Volkmann RA & Schreiber SL (1993) Sequential assignment and structure determination of spider toxin omega-Aga-IVB. Biochemistry, 32, 13123–13129. 10.1021/bi00211a022 [PubMed: 8241166]
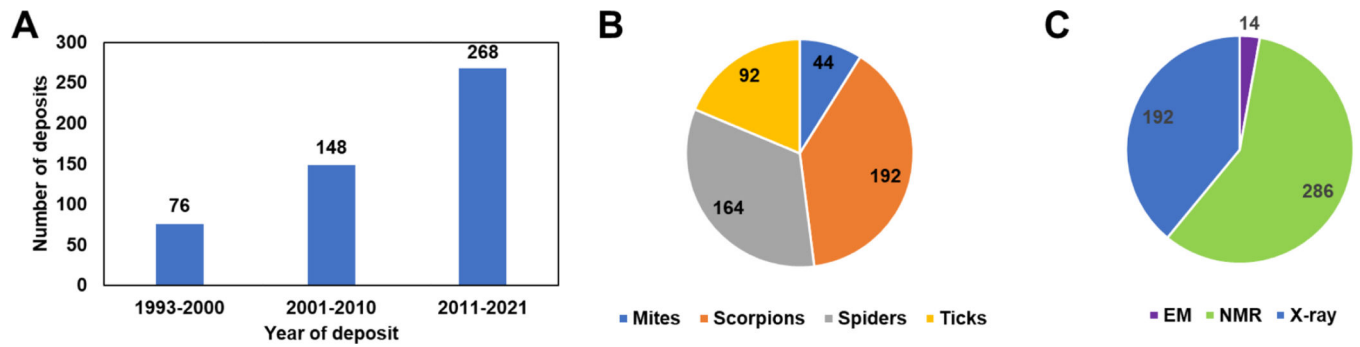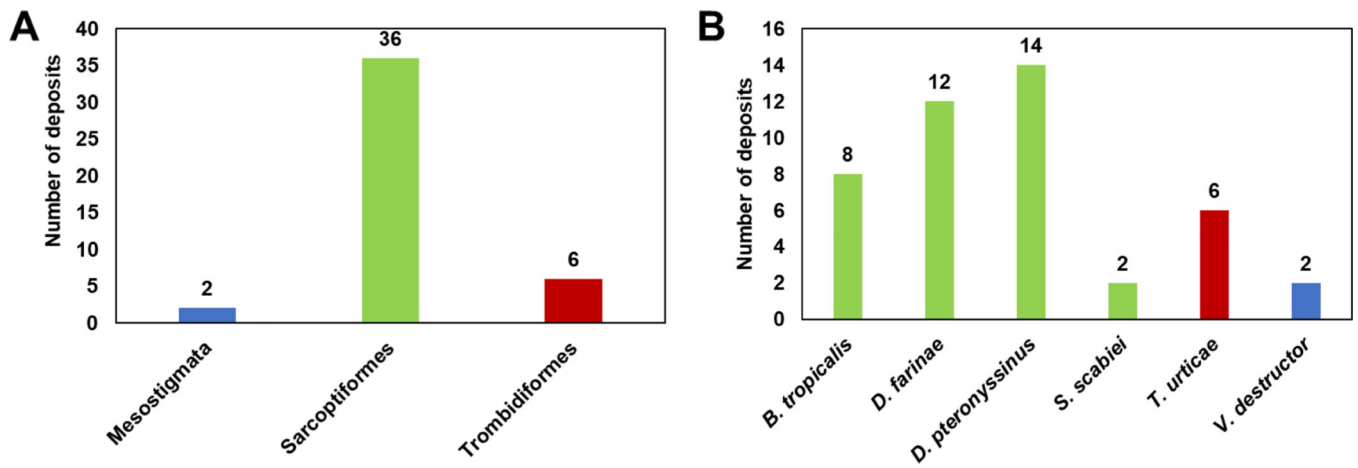
Figure 1.
(A) Number of the PDB deposits of proteins originating from Arachnida between 1992 and June 2021. (B) Sources of Arachnida proteins that have structures determined. (C) Experimental methods used to determine structures.
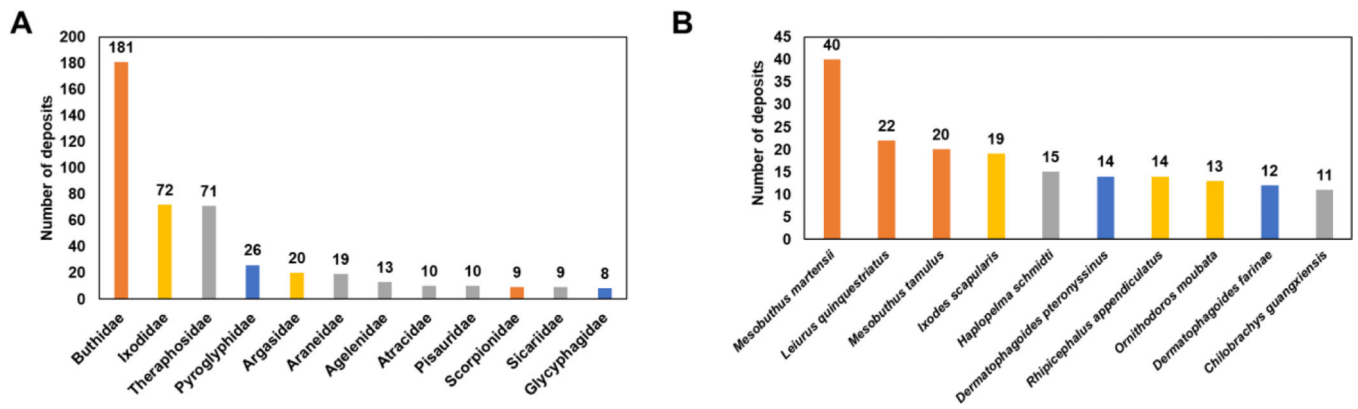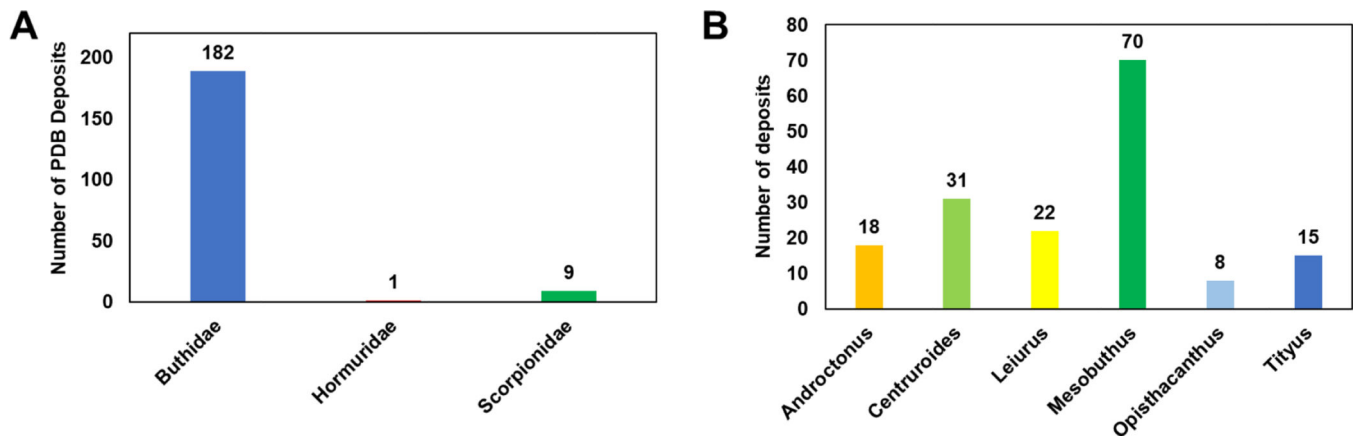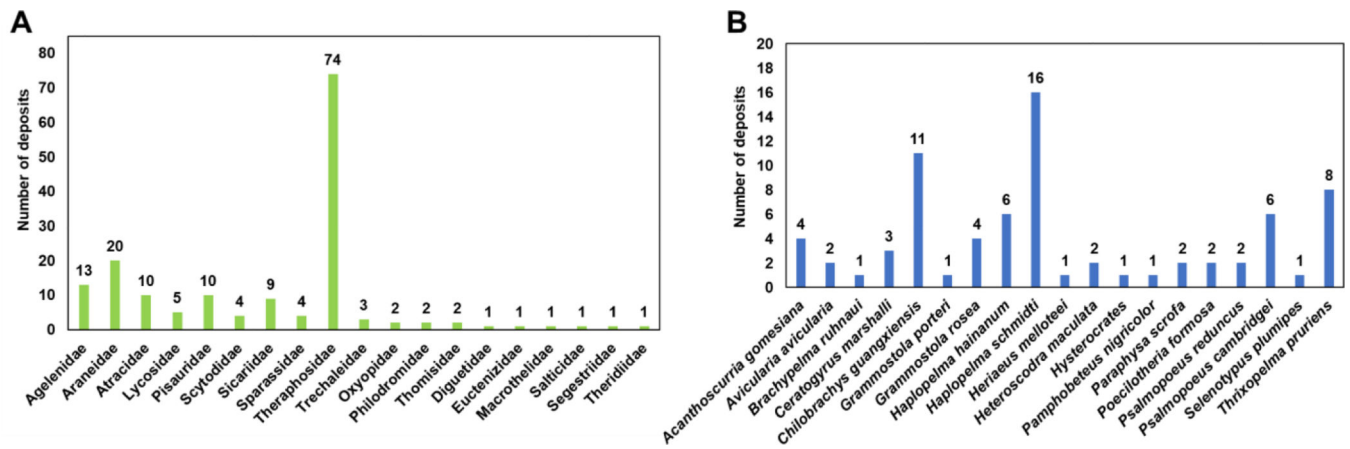
Figure 2.
(A) Number of the PDB deposits of proteins originating from Arachnida. Twelve families with the largest numbers of deposits are shown. (B) Arachnida species with at least eleven PDB deposits as of June 2021. Dark orange is used to mark scorpions, yellow is used for ticks, grey for spiders, and blue for mites.

Figure 3.
(A) Mite orders represented by the PDB deposits. (B) Numbers of the PDB deposits originating from different mite species.
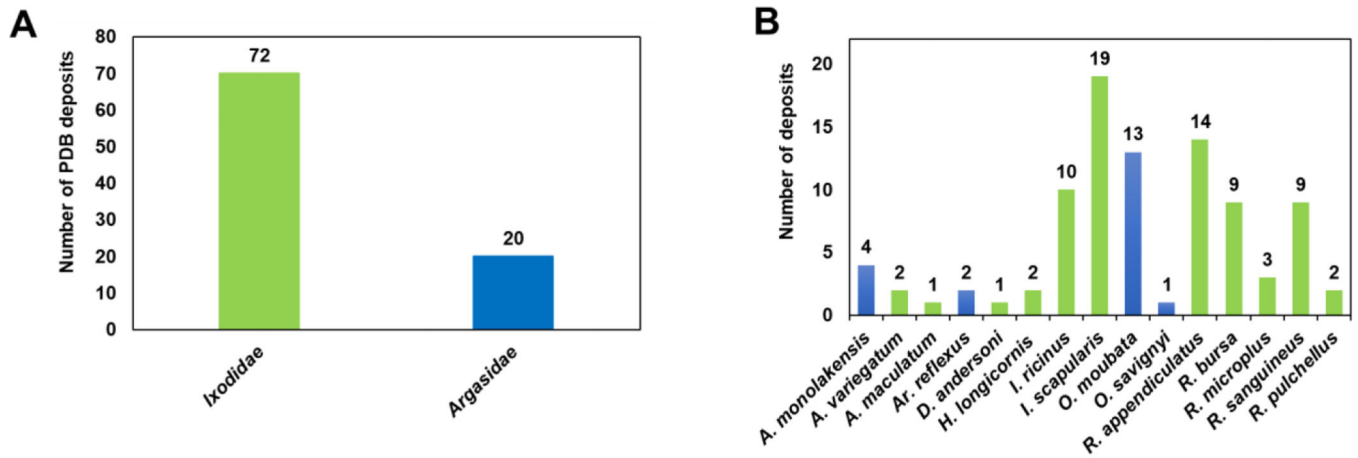
Figure 4.
Proteins originating from scorpions that have their structures determined. (A) The PDB deposits are grouped according to family. (B) The PDB deposits are counted according to their genus.

Figure 5.
(A) Number of PDB deposits originating from spider, arranged by families. (B) Deposits corresponding to proteins originating from spiders from Theraphosidae family, arranged by species.

Figure 6.
(A) The PDB deposits originating from Ixodidae (hard ticks; marked in green) and Argasidae (soft ticks; marked in blue). (B) Numbers of the PDB deposits for proteins originating from different tick species.