The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework

Nielsen, Lau D; Foged, Mads M.; Albert, Anastasia; Bertelsen, Andreas B.; Søltoft, Cecilie L; Robinson, Samuel D.; Petersen, Steen Vang; Purcell, Anthony W.; Olivera, Baldomero M.; Norton, Raymond S.; Vasskog, Terje; Safavi-Hemami, Helena; Teilum, Kaare; Ellgaard, Lars

Published in:
The Journal of Biological Chemistry

DOI:
10.1074/jbc.RA119.007491

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Venomous marine cone snails produce peptide toxins (cono-toxins) that bind ion channels and receptors with high specificity and therefore are important pharmacological tools. Conotoxins contain conserved cysteine residues that form disulfide bonds that stabilize their structures. To gain structural insight into the large, yet poorly characterized conotoxin H-superfamily, we used NMR and CD spectroscopy along with MS-based analyses to investigate H-Vc7.2 from *Conus victoriae*, a peptide with a VI/VII cysteine framework. This framework has Cys¹–Cys⁴/Cys¹¹–Cys¹⁸/Cys³¹–Cys⁵¹ connectivities, which have invariably been associated with the inhibitor cystine knot (ICK) fold. However, the solution structure of recombinantly expressed and purified H-Vc7.2 revealed that although it displays the expected cysteine connectivities, H-Vc7.2 adopts a different fold consisting of two stacked β-hairpins with opposing β-strands connected by two parallel disulfide bonds, a structure homologous to the N-terminal region of the human granulin protein. Using structural comparisons, we subsequently identified several toxins and nontoxin proteins with this “mini-granulin” fold. These findings raise fundamental questions concerning sequence–structure relationships within peptides and proteins and the key determinants that specify a given fold.

Venoms from a variety of animals such as snakes, spiders, scorpions, and marine snails contain diverse peptide toxins that bind with high specificity and affinity to their molecular targets. These characteristic features make venom peptide toxins potentially valuable pharmacological tools and drug leads (1). One toxin from a predatory marine cone snail, ω-MVIIIA from *Conus magus*, is a Food and Drug Administration–approved drug (with the commercial name Prialt) used to treat severe and chronic pain (2), and several other conotoxins are in development for various pathologies, including pain, epilepsy, stroke, and diabetes (3–6). In addition, many venom peptides are used as affinity reagents and constitute valuable research tools. Each of the ~800 species of cone snails of the genus *Conus* produces a unique set of 100–400 venom peptides that are typically referred to as conotoxins or conopeptides. Conotoxins are ribosomally synthesized as preproteins, containing a conserved signal sequence for endoplasmic reticulum (ER) localization, a propeptide region, and the C-terminal mature toxin that is released upon proteolytic removal of the propeptide. Conotoxins can be divided into gene superfamilies based on their conserved signal sequence (7, 8). To date, more than 53 gene superfamilies have been described in this genus (9). The peptide investigated in this study, H-Vc7.2, was first identified by transcriptional sequencing of the venom gland of *Conus victoriae* (10) and belongs to the H-superfamily of conotoxins.

Mature conotoxins display highly variable amino acid sequences at nearly all positions apart from the conserved cysteine residues that define their cysteine framework (11). Presently, the reported conotoxin sequences are divided into 26 cysteine frameworks (8). H-Vc7.2 adopts one of the most...
Granulin fold arising from a common ICK cysteine framework

Results

Native H-Vc7.2 is devoid of post-translational modifications

Previous analyses of the venom gland transcriptome of C. victoriae revealed the uncharacterized conotoxin H-superfamily, from which three transcripts were identified (10). By searching high-resolution MS/MS spectra of reduced and alkylated C. victoriae venom against the translated venom gland transcriptome, the mature form of the H-Vc7.2 conotoxin was identified. H-Vc7.2 exists in the venom as a 25-amino acid–residue peptide with its six cysteine residues arranged in the type VI/VII framework (C–C–CC–C–C) (Fig. 1). No variants of H-Vc7.2, e.g. post-translationally-modified, or peptides of different chain length were detected.

H-Vc7.2 expressed in E. coli purifies as a single major species

Disulfide-containing proteins can be expressed in the oxidized state in the cytosol of E. coli using the CyDisCo system (16, 17), which can produce highly complex disulfide-bonded proteins (16, 19). In this system, the protein of interest is co-expressed with two redox enzymes, the mitochondrial oxidase Erv1p from Saccharomyces cerevisiae and human PDI. Erv1p provides the oxidizing equivalents to generate disulfide bonds de novo, whereas PDI isomerizes non-native disulfides. In contrast to the commonly used Shuffle and Origami cells, where the reducing pathways present in the cytoplasm are disrupted, the CyDisCo system thus uses an active enzyme system to make disulfide bonds. Published comparisons of CyDisCo versus Origami show the former system to be more efficient for the proteins investigated (16). Moreover, the CyDisCo system is highly versatile in the sense that it works in any strain and in any media tested to date.5

We modified the original CyDisCo system for targeted expression of conotoxins by co-expressing, in addition to human PDI and Erv1p, a conotoxin-specific PDI (Fig. 2A) (18). In previous work (18), we have shown that in vitro this csPDI

5 L. Ruddock, University of Oulu, Finland, personal communication.
enzyme greatly accelerates the generation of correctly folded conotoxins during oxidative folding. Therefore, we reasoned that co-expressing the csPDI may help improve expression levels of some conotoxins.

Recombinant H-Vc7.2 was expressed as a fusion protein with an N-terminal ubiquitin (Ub)-tag containing 10 consecutive histidines (Ub–His_{10}) followed by a TEV protease recognition site (Fig. 2B) (20). Upon TEV cleavage, this recombinant H-Vc7.2 peptide contains four additional N-terminal amino acid residues, Gly–Ala–Met–Gly (GAMG), as compared with the native peptide because of the cloning procedure used. This N-terminally extended 29-residue recombinant peptide is hereafter referred to as N_{ext}-H-Vc7.2, and was used for determination of the cysteine connectivity and structural analysis by NMR spectroscopy. For CD spectroscopy, comparison of reversed-phase HPLC (RP-HPLC) retention time with C. victoriae venom and bioactivity studies during this work, we also generated recombinant H-Vc7.2 (named rH-Vc7.2) devoid of the four N-terminal nonnative residues. All numbering throughout this study refers to the extended peptide, except for Fig. 1 where the numbering refers to the native sequence.

Overnight induction with IPTG resulted in clearly visible bands for hPDI and csPDI (that co-migrated by SDS-PAGE), Erv1p, and Ub–His_{10}–N_{ext}-H-Vc7.2 (Fig. 2C). Moreover, the majority of Ub–His_{10}–N_{ext}-H-Vc7.2 was found in the soluble fraction. N_{ext}-H-Vc7.2 was purified as outlined in Fig. S1A. The His_{10}-tagged fusion protein was purified from crude lysate by metal-affinity chromatography (Fig. S1B). Upon TEV protease cleavage, uncleaved fusion protein, His_{6}-TEV protease, and the liberated Ub–His_{10}-tag were removed by batch purification on cobalt resin, and the peptide was further purified by RP-HPLC. The obtained chromatogram (Fig. 2D) showed a major peak as well as smaller peaks with increased retention time. The fraction representing the major peak was collected as indicated in Fig. 2D as a single major species.

To determine the cysteine connectivity in N_{ext}-H-Vc7.2, we used sequential reduction and alkylation followed by mass....
Granulin fold arising from a common ICK cysteine framework

Figure 3. rH-Vc7.2 elutes with the same retention time as the native venom peptide, and disulfide bonds are critical for stability.

To investigate whether rH-Vc7.2 (devoid of the four N-terminal nonnative residues) shared the same structural characteristics as the native peptide, we compared retention times of rH-Vc7.2 and the native peptide from *C. victoriae* venom by RP-HPLC. Moreover, the importance of the disulfide bonds for the stability of the rH-Vc7.2 peptide was investigated by CD spectroscopy.

The venom fraction that eluted at the same percent of solvent B as rH-Vc7.2 (24.3–24.7% solvent B) contained a peptide with the same mass as rH-Vc7.2 (observed [M + H]⁺ = 2785.15; calculated [M + H]⁺ = 2785.12) demonstrating that rH-Vc7.2 exhibits the same physicochemical properties as the native venom peptide (Fig. 3, A–C). Adjacent fractions were also subjected to MALDI-TOF MS. A corresponding mass was also observed in the fraction eluting immediately prior to rH-Vc7.2 (23.9–24.3% solvent B), but not in any other adjacent fractions. Future studies may determine whether the peptide with the same mass as H-Vc7.2 could represent the same peptide with an alternative fold (e.g. ICK or other) or a proline cis/trans isomer.

We then used CD spectroscopy to characterize further structural properties of the peptide. CD spectra of rH-Vc7.2 were spectrometric analyses as described recently (22). To suppress the risk of disulfide scrambling, reduction and alkylation steps were performed at low pH using a large molar excess of alkylation agent and with the peptide adsorbed to a solid phase, which further minimizes disulfide reconnection because of the restrained conformational freedom of the peptide. The sequential reduction and alkylation were performed with different maleimides resulting in samples where the cysteines of the individual disulfide bridges have distinct alkylation patterns. This approach allows unambiguous identification of the cysteine connectivity. In total, four different alkylation patterns with *N*-methylmaleimide (NMM), *N*-ethylmaleimide (NEM), and *N*-cyclohexylmaleimide (NCM) were obtained for *N*exH-Vc7.2. All alkylation patterns were compatible with a single cysteine connectivity in the peptide. The MS/MS spectra of *N*exH-Vc7.2 alkylated with two NMM, two NEM, and two NCM are shown in Fig. S3.

The *N*exH-Vc7.2 peptide produced mainly y-type ions with additional (y-NH3) and (y-H2O) fragments. Based on the MS/MS spectra in Fig. S3, the alkylation pattern was assigned as C8-NEM/C13-NCM/C17-NMM/C18-NEM/C23-NCM/C28-NMM. The resulting cysteine connectivity in the peptide consequently is C8–C18/C13–C23/C17–C28. Detailed information about the identified fragments and three additional alkylation patterns with identified fragments all supporting the same cysteine connectivity can be found in Fig. S4 and Tables S1–S4.

Recombinant H-Vc7.2 elutes with the same retention time as the native venom peptide, and disulfide bonds are critical for stability

...
recorded at 25 and 90 °C and (after cooling) again at 25 °C, under both reducing and nonreducing conditions (Fig. 3D). The magnitude of the molar residual ellipticity in all spectra was low, indicating the presence of few regular secondary structure elements in the peptide. The spectrum of the oxidized peptide at 25 °C showed two minima, one at 205 nm and the other at 222 nm. A negative ellipticity minimum around 206 nm has been observed for small β-sheet and disulfide-rich proteins (23, 24). The combination of weak negative bands at around 208 and 222 nm is characteristic of both α-helical peptides and peptides with type I β-turns (25). Deconvolution of the CD spectrum predicted a protein devoid of α-helix but containing anti-parallel β-sheets (42.9%), turns (15.8%), and “others” (41.2%) (Fig. S5). Together, these CD data are consistent with the presence of β-turns and some β-sheet secondary structure in rH-Vc7.2 (see below). Upon reduction of the peptide, the CD spectrum changed considerably, displaying an ellipticity minimum at 198 nm and a shape of the curve suggesting an unstructured peptide.

To investigate the thermal stability of oxidized rH-Vc7.2, the change in ellipticity was followed at 224 nm from 10 to 90 °C. The experiment revealed a linear temperature dependence (data not shown), demonstrating that rH-Vc7.2 does not unfold cooperatively over this temperature range. The spectrum recorded at 90 °C displayed the same overall features as seen at 25 °C with a slightly lower ellipticity in the 205–250-nm range, indicating that the fold was preserved but that the conformational ensemble had changed. This change, however, was reversible upon cooling of the sample to 25 °C. We concluded that the folded peptide shows strong resistance to thermal denaturation and that the disulfide bonds are critical to maintain the structure of rH-Vc7.2.

**Nexo-H-Vc7.2 structure is characterized by two short, stacked β-hairpins**

Having determined the cysteine connectivity and the importance of disulfide bonds for structural integrity, we used NMR spectroscopy to investigate the structure of Nexo-H-Vc7.2. Using homonuclear 2D total correlation spectroscopy (TOCSY) and 2D NOE spectroscopy (NOESY) spectra in combination with triple resonance backbone spectra, 40% of backbone heavy atoms, and 60% of the nonexchangeable protons were assigned. Resonances from 20 out of 27 possible backbone amide groups were assigned. No peaks in the 15N heteronuclear single quantum coherence spectroscopy (HSQC) spectrum could be assigned to residues 6 and 7 and 14–18. The 15N HSQC (Fig. 4B) revealed a large variation in the intensity of observed peaks, with peaks from unassigned residues likely broadened beyond detection. Such peak broadening is the result of exchange between two or more conformations of the molecule on the millisecond time scale and indicates a dynamic nature of these bonds connecting Cys-8 and Cys-18 (1), Cys-13 and Cys-23 (2), and Cys-17 and Cys-28 (3), and the two β-hairpins. Structural elements are colored with β-strands in red, sulfur atoms of the six cysteine residues in yellow, and loops in teal. Positions of the N and C termini are indicated. D, superimposition of the 20 lowest energy structures. The structures are shown in the same overall orientation as the conformer in C and with the same color-coding (see also Table 1).
regions of the structure. We note that residues 6 and 7 and residues 17 and 18, which are located in β-strands, are connected by a disulfide bond. If there are dynamics in the structure on one side of the disulfide bond, the other side likely also experiences these dynamics.

One process that could cause the proposed peak broadening would be transient dimer formation (26). We therefore measured the diffusion coefficient for NextH-Vc7.2 by NMR to 1.96 ± 0.08 \times 10^{-10} \text{m}^2/\text{s} (data not shown). This is only slightly faster than the theoretical value of 1.54 \times 10^{-10} \text{m}^2/\text{s} calculated using HydroPro (27). This result therefore indicates that the peak broadening is not caused by the formation of dimers or higher-order oligomers. We also screened pH (in the range of pH 4.0–6.1), temperature (5–45 °C), and salt concentrations (up to 150 mM NaCl) to determine whether a single well-ordered conformation could be stabilized. However, we did not observe any large effect on the line shape of the different conformations.

To gain further support for the presence of the two β-hairpins, we determined temperature coefficients for the amide proton chemical shifts (Fig. 5, A and B). Hydrogen-bonded protons typically have temperature coefficients closer to 0 than −4.6 ppb/K, whereas nonhydrogen-bonded protons typically have temperature coefficients lower than −4.6 ppb/K (28, 29). The temperature coefficients for rH-Vc7.2 were determined by linear regression using the amide hydrogen chemical shifts from a series of $^{15}$N HSQC spectra measured in the interval between 5 and 45 °C (Fig. 4B and 5A). Plotting the temperature coefficients onto the structure revealed that the amide protons predicted from the calculated structure to be involved in hydrogen bonding indeed had temperature coefficients of −4.6 ppb/K or higher (Fig. 5B). We note that although Gly-10 and Ile-26 are not involved in β-sheet formation, their temperature coefficients indicate hydrogen bonding. Based on analysis of the structure, the amide proton of Gly-10 could possibly form a hydrogen bond with the sulfur atom of Cys-8, whereas the most likely candidate for hydrogen bonding with the Ile-26 amide proton is the carbonyl oxygen of Arg-24.

Importantly, the raw NMR data were inconsistent with an ICK fold, which displays a two-stranded anti-parallel β-sheet that is often extended with an additional anti-parallel β-strand (12, 30). Instead, the data substantiated the presence of two short β-hairpins, the first comprising a type I and the second a type I’ β-turn as determined from the backbone dihedral angles.

**Structural topology of NextH-Vc7.2 is unlike the ICK fold**

On the basis of the disulfide framework VI/VII alone, it was predicted that H-Vc7.2 would have an ICK fold. However, in NextH-Vc7.2, the connections between β-strands β1, β2, and β3 and between β2, β3, and β4 are both left-handed in contrast to the ICK fold, where these connections are left-handed and right-handed, respectively. Furthermore, the crossing disulfides that are a hallmark of the ICK fold (Fig. 6A) are not present in NextH-Vc7.2 (Fig. 6B). The first disulfide in NextH-Vc7.2 connects β1 with the opposing β3, whereas the second disulfide connects the opposing β2 and β4. As a result, the first two disulfides in NextH-Vc7.2 are parallel. The third disulfide connects the beginning of β3 with the unstructured C-terminal region without threading the loop formed by the first two disulfide bonds, another hallmark of the ICK fold. The NextH-Vc7.2 structure therefore unexpectedly displays a fold that is clearly distinct from an ICK.

### Table 1

NMR structural restraints and structure statistics

None of the structures exhibited distance violations >0.3 Å or dihedral angle violations >5°.

<table>
<thead>
<tr>
<th>Restraints</th>
<th>(N_{ex} \text{H-Vc7.2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOE-based restraints</td>
<td>60</td>
</tr>
<tr>
<td>Intraspatial ((i - j) = 0)</td>
<td>60</td>
</tr>
<tr>
<td>Sequential ((i - j) = 1)</td>
<td>49</td>
</tr>
<tr>
<td>Medium range (2 \leq</td>
<td>i - j</td>
</tr>
<tr>
<td>Long (</td>
<td>i - j</td>
</tr>
<tr>
<td>Total</td>
<td>152</td>
</tr>
<tr>
<td>Dihedral angle restraints</td>
<td>34</td>
</tr>
</tbody>
</table>

**Restraint statistics**

Mean RMSD from experimental restraints:

- Backbone dihedral angles: 0.012 ± 0.002 Å
- Backbone heavy atoms: 0.18 ± 0.04 Å

**Structure statistics**

- Most favored regions: 84.1%
- Allowed regions: 15.2%
- Generously allowed regions: 0.5%
- Disallowed regions: 0.2%

**Coordinate precision, pairwise RMSD (Å)**

| H-Vc7.2 (6–25) | 0.43 ± 0.23 |
| Backbone heavy atoms (N, Cα, and C′) | 1.37 ± 0.29 |

*ProCHECK was used.

* MOLMOL was used.
Granulin fold arising from a common ICK cysteine framework

Considering the results described above, we searched the Protein Data Bank (PDB) for structural homologs of NextH-Vc7.2. The results showed the NextH-Vc7.2 core structure to have the same fold as the N-terminal domain of the human progranulin A module (Fig. 7A). Although the second β-hairpin in the N-terminal domain of the human progranulin A module is longer than the corresponding β-hairpin in NextH-Vc7.2, the structures align well and display the same arrangement of disulfide bonds. Another conotoxin, Φ-MiXXVIIA, with structural homology to granulin was recently discovered (31). The two toxins have highly similar structures with the inter-β-hairpin disulfides placed similarly, whereas the remaining disulfides in the two toxins are at opposite ends of the peptides (Fig. 7B). Notably, H-Vc7.2, granulins, and Φ-MiXXVIIA are not the only peptides displaying the observed structure, which we identified in a number of other proteins. Thus, we found that domains from other toxins, such as the D-GeXXa conotoxin and leech antistasins (serine protease inhibitors), as well as from the nontoxin proteins follistatin-related protein 3 (and other proteins with EGF-like domains, e.g. fibrillin) and the zinc-binding lobe of the E3 ubiquitin ligase, Pirh2, where cysteine residues coordinate Zn$^{2+}$ ions rather than forming disulfide bonds, have similar folds (Fig. 7, C–F; see under “Discussion”). RMSD values for the structures aligned in Fig. 7, as well as an alignment of their sequences, are shown in Table S6.

In conclusion, we find that the fold comprising two short, stacked β-hairpins stabilized by two parallel disulfide bonds across a small β-sandwich structure is present in a variety of toxins and other proteins. Given the wide distribution of this structure, which apparently constitutes an autonomous folding unit, and its homology with the N-terminal domain of granulins, we propose to name this structure the “mini-granulin” fold.
Granulin fold arising from a common ICK cysteine framework
Investigation of bioactivity

rH-Vc7.2 was tested for bioactivity by intracranial injection in mice. In this assay, a dose of 3.5 nmol of peptide did not produce any strong reproducible behavioral changes compared with control animals injected with saline (data not shown). Similarly, rH-Vc7.2 (10 μM) did not produce any observable changes to normal or depolarization-induced intracellular Ca²⁺ levels in mouse dorsal root ganglion cells (data not shown).

Discussion

Whereas the same cysteine framework can give rise to different cysteine connectivities (32), the finding that the same cysteine connectivities can dictate different folds is unusual. This work therefore emphasizes that caution must be taken in assuming a specific structure based on a certain cysteine framework, or even a known disulfide pattern. Specifically, we demonstrate that, despite having the disulfide pattern expected of a peptide with a VI/VII cysteine framework, N₆₋H-Vc7.2 did not adopt an ICK fold. Rather, the peptide adopted a mini-granulin fold, a structure composed of two short, stacked β-hairpins connected by two parallel disulfide bonds.

Progranulins are composed of separate disulfide-rich granulin modules of ~55 residues each (33). Structurally, each module is constituted by stacks (four in most cases) of β-hairpins interconnected by stabilizing disulfide bonds (Fig. 7A) (34, 35). The N-terminal region comprising the first ~30 residues of the human progranulin A module is structurally much better defined than the C-terminal region (33). Indeed, a fragment comprising residues 1–30 of carp granulin displays essentially the same structure as the corresponding region of the intact protein (36). Similarly, an N-terminal 35-residue peptide of a granulin-like module from a rice protease inhibitor constitutes a well-defined structure with the expected fold (37), and designed N-terminal peptides of human (38) and live fluke (39) granulins also adopt the same overall structure. Thus, it is well-established that the N-terminal region of several granulins constitutes an autonomous folding unit.

Here, we corroborate and extend the findings for granulin by identifying a variety of structural homologs of Nₓ₋H-Vc7.2 where the mini-granulin fold constitutes a separate domain (Fig. 7). Our analyses show that the recent finding that the Φ-MiXXVIIIA conotoxin adopts the mini-granulin fold (Fig. 7B) (31) is not a unique case. For instance, the C-terminal domain of the αD-GeXXa conotoxin, which has not previously been assigned to a particular fold, also constitutes a mini-granulin structure (Fig. 7C) (40). A domain with the mini-granulin fold is also found in the antistasins, i.e., leech serine protease inhibitors such as antistasin, hirustasin, and bdellastasin (Fig. 7D) (41). Another structural homolog of Nₓ₋H-Vc7.2 is the EGF-like subdomain of follistatin-related protein 3 (42), a transforming growth factor β family antagonist (Fig. 7E). Structural homology between the N-terminal domain of granulin and the N-terminal subdomain of EGF-like molecules has been noted previously (36, 43). We also detected homology with the C-lobe of the N-terminal domain of the E3 ubiquitin ligase Pirh2, a zinc-binding region of the protein (Fig. 7F) (44), in which cysteines coordinate zinc atoms located between the β-hairpins rather than forming disulfide bonds. Overall, the mini-granulin structure is widely present as a separate domain in toxins as well as nontoxin proteins.

Although the presence of the mini-granulin fold in nature may well have been underappreciated to date, ICK peptides are still much more common (14). Their evolutionary success is likely a consequence of the high thermodynamic stability and resistance to proteolytic cleavage afforded by the ICK fold, and the fact that this fold can function as a scaffold for a vast array of functionalities (13, 45). In comparison, the core mini-granulin fold contains only two disulfides and is probably dependent on the acquisition of additional disulfides, as seen in e.g. H-Vc7.2 and Φ-MiXXVIIIA, for added structural stability. Although Φ-MiXXVIIIA has been demonstrated to display (weak) anti-apoptotic and cell-proliferative activities (31), as is the case for certain granulin modules, these are unlikely to constitute the native activities of this conotoxin, and the ability of granulin-like toxins to harbor different functionalities warrants further investigation.

Conotoxins and other disulfide-rich toxins have previously been produced successfully in E. coli using various expression systems (46, 47). Nₓ₋H-Vc7.2 and rH-Vc7.2 were produced using the CyDisCo expression system that supports the formation of disulfide bonds in the reducing cytosol of E. coli. Based on our previous finding that the csPDI enzyme accelerates oxidative folding of selected conotoxins in vitro (18), we surmised that co-expression with csPDI could increase the folding efficiency of at least some conotoxins in this expression system. Further studies are required to address whether the csPDI–CyDisCo expression system is widely applicable to the expression of small disulfide-rich peptides and to investigate a potential synergism between hPDI and csPDI in disulfide-bond formation of conotoxins as we recently demonstrated in vitro for Conus PDI and csPDI (48).

Here, our results showed that when produced in the absence of the csPDI–CyDisCo system, a large fraction of Nₓ₋H-Vc7.2 molecules ended up in the insoluble fraction (Fig. S6). A systematic investigation of the importance of this system for the production of oxidized Nₓ₋H-Vc7.2, e.g., the specific effect of the csPDI as compared with human PDI or a comparison with other E. coli expression systems such as Shuffle and Origami cells, was beyond the scope of the current investigation. Ongoing work in our laboratory is aimed at understanding in more detail the effect of the csPDI on the folding of Nₓ₋H-Vc7.2 peptides.

Figure 7. Nₓ₋H-Vc7.2 adopts a fold similar to the N-terminal domain of granulin. Nₓ₋H-Vc7.2 (dark green) is aligned to different structural homologs (displayed in gray outside the region of structural homology). A, N-terminal domain of the human progranulin A module (dark blue, PDB code 2JYE) (33). Φ-MiXXVIIIA conotoxin (violet) (31). C, C-terminal domain of the αD-GeXXa conotoxin (red, PDB code 4X92) (40). The second molecule of the αD-GeXXa dimer is depicted in light blue. D, N-terminal domain of the leech serine protease inhibitor bdellastasin (orange, PDB code 1C9T) (41). E, N-terminal domain of the follistatin-related protein 3 (FSTL-3) (yellow, PDB code 2KCX). F, zinc-binding β-hairpin repeat from the C-lobe of the N-terminal domain of Pirh2 (light green, PDB code 2K2C) (44). Top, alignment with the second and third β-hairpins; bottom, alignment with the first two β-hairpins. Zinc atoms are shown as purple spheres.
Granulin fold arising from a common ICK cysteine framework

detail the effects of the csPDI–CyDisCo system on conotoxin production for a larger and diverse set of toxins. The function of H-Vc7.2, and indeed all H-superfamily conotoxins, remains an open question. In the assays used here, no obvious bioactivity was detected. Potentially, H-Vc7.2 interacts with a receptor target in such a way as to not produce observable responses in these assays, or perhaps, the peptide is selective for one or more receptors not present in the assay systems used. Future in vitro activity assays using gastropods, the prey of C. victoriae, may shed light on the peptide’s bioactivity.

From the many available structures of ICK peptides with large sequence variations in the loops between cysteine residues, it is clear that the disulfide bonds are crucial for directing the ICK fold (30). Apart from the side chain of Val-11 that packs against the first two disulfide bonds, the structure of N-Vc7.2 shows no evidence of a hydrophobic core that could help direct the folding of the peptide. Instead, as proposed for the Φ-MiXXVIIIA conotoxin (31), the short loop lengths in H-Vc7.2 could well restrict the polypeptide conformation to preclude the formation of an ICK fold. An analysis of the more than 3300 sequences of presumed ICK proteins listed in the KNOTTIN database (http://www.dsimb.insERM.fr/KNOTTIN/) supports this suggestion. Among these many sequences, only 10 contain four (or fewer) residues in the loop between CysI and CysII, three (or fewer) residues between CysII and CysIII, and no residues between CysIII and CysIV. No structures have been solved for any of these 10 proteins. Only when the loop length between CysIII and CysIV reaches three residues (while retaining ≤4 and ≤3 residues in the first two loops, respectively) have ICK structures been solved. However, a more thorough analysis of the consequences of varying loop lengths in cysteine framework VI/VII proteins is needed to determine the underlying features of the H-Vc7.2 sequence that govern its structure and lead to the formation of the mini-granulin fold rather than the ICK fold. Such work will be fundamental to allow the design of better structure prediction algorithms and help further our understanding of sequence–structure relationships in peptides and proteins.

Experimental procedures

Mass spectrometry (MS) on venom extract

C. victoriae snails were collected from Broome, Western Australia, under a commercial fishing license of the Western Australian Specimen Shell Managed Fishery (license number 2577). The venom was obtained by manual extrusion from freshly dissected venom glands. An aliquot (0.5 μg) of venom was reduced in 20 mM tris(2-carboxyethyl)phosphine (TCEP), pH 8, for 30 min at 60 °C, then alkylated by incubating in 40 mM iodoacetamide for 30 min. This reduced and alkylated venom was loaded onto a microfluidic trap column packed with ChromXP C18-CL 3-μm particles (300 Å nominal pore size). An analytical (15-cm × 75-μm ChromXP C18-CL 3-μm) microfluidic column was then switched in line, and venom was separated using a linear gradient of 0–80% acetonitrile, 0.1% formic acid over 90 min at a flow rate of 300 nl/min. Separated venom peptides were analyzed using an AB SCIEX 5600 TripleTOF mass spectrometer equipped with a Nanospray III ion source and accumulating 30 tandem MS (MS/MS) spectra/s. MS/MS data were analyzed with ProteinPilot software (version Beta 4.1.46) using the Paragon algorithm. The search databases comprised a six-frame translation of the C. victoriae venom gland transcriptome, as described previously (10), including the translated prepropeptide sequence of H-Vc7.2. MS/MS spectra identified by ProteinPilot as matching peptides were manually validated by comparison against a theoretical peak list (Protein Prospector MS-Product, University of California, San Francisco).

Plasmid generation

Based on the nucleotide sequence of H-Vc7.2 from C. victoriae (10), an E. coli codon-optimized H-Vc7.2 gene containing Ncol and BamHI restriction sites was synthesized (Eurofins Genomics GmbH) and cloned into the pET39 Ub19 plasmid (20) (a kind gift from V. Rogov and V. Dötsch, Goethe University, Frankfurt, Germany). The resulting plasmid expressing ubiquitin (Ub)–His10-tagged, N-terminally extended H-Vc7.2 (Ub–His10–N-Vc7.2) was referred to as pLE566. A plasmid (pLE879) differing only from pLE566 by encoding the native sequence devoid of the four non-native residues in N-Vc7.2 was purchased from TWIST Bioscience. The fusion protein produced from pLE879 was named Ub–His10–H-Vc7.2.

A CyDisCo plasmid (pMJS205 (17)) encoding Erv1p and hPDI was a gift from L. W. Ruddock (Dept. of Biochemistry, University of Oulu, Finland). A codon-optimized sequence for E. coli expression of csPDI,GH,HGH from C. geographus (GenBank accession no. KT874567) (18) was synthesized with XbaI and Xhol restriction sites (Eurofins Genomics GmbH) and cloned into pMJS205, generating the plasmid pLE577 that encodes Erv1p, hPDI, and csPDI,GH,HGH (see Fig. 2).

Protein expression

Chemically competent E. coli BL21 Tuner(DE3) cells (Novagen) were co-transformed with pLE577 and pLE566 or pLE879 and plated on an LB agar plate containing kanamycin (50 μg/ml) and chloramphenicol (30 μg/ml). One colony was inoculated in LB medium containing the same antibiotics and incubated overnight at 37 °C at 200 rpm in an orbital shaker. Overnight cultures were diluted in LB or M9 media (3 g/liter KH2PO4, 15.1 g/liter Na2HPO4, 12H2O, 5 g/liter NaCl, 1 mM MgSO4, 1 ml/liter M2 Trace element solution (203 g/liter MgCl2, 6H2O, 2.1 g/liter CaCl2, 2H2O, 2.7 g/liter FeSO4 7H2O, 20 mg/liter AlCl3 6H2O, 10 mg/liter CuSO4 7H2O, 2 mg/liter KCr(SO4)2 12H2O, 2 mg/liter CuCl2 2H2O, 1 mg/liter H3BO4, 20 mg/liter KI, 20 mg/litter MnSO4 H2O, 1 mg/liter NiSO4 6H2O, 4 mg/liter Na2MoO4 2H2O, 4 mg/liter ZnSO4 7H2O, 21 g/liter citric acid monohydrate), 1 g/liter [15N]NH4Cl, and 4 g/liter [13C]glucose) to an optical density at 600 nm (OD600) of 0.1 and incubated at 37 °C until OD600 reached ~0.8. Protein expression was induced by 1 mM IPTG, and cultures were grown overnight at 30 °C. Expression of the 13C/15N-labeled peptide in 580 ml of M9 minimal medium was performed as described for the unlabeled peptide.

Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.
Granulin fold arising from a common ICK cysteine framework

Cell harvest

Induced overnight cultures were centrifuged for 15 min at 6,000 × g; supernatants were discarded, and pellets were resuspended in 10 ml of lysis buffer (300 mM NaCl, 50 mM Na2HPO4/NaH2PO4, 10 mM imidazole, pH 8.0)/liter of culture medium. Sonication was performed on a Bandelin Sonopuls HD2200 equipped with a Boosterhorn SH 213 G and a 3 mm probe. Cells were sonicated 10× for 10 s at 30% power with 30 s on ice between pulses. After centrifugation for 50 min at 15,000 × g, the cleared lysates were transferred to fresh tubes, while pellets were dissolved in lysis buffer containing 8 M urea for SDS-PAGE analysis.

Protein purification

Ub–His10–NexH-Vc7.2 was purified from the cleared lysate on an ÄKTA purifier 900 chromatography system equipped with a GE Healthcare Tricorn 10/50 column packed with 5 ml of Qiagen Superflow Ni-NTA resin equilibrated in lysis buffer. Upon sample application, the column was washed with lysis buffer containing 20 mM imidazole until a stable baseline at A280 was achieved, and the protein was eluted with lysis buffer containing 400 mM imidazole. The pooled material (7 ml) was dialyzed twice against 2 liters of lysis buffer at 4 °C using Spectrum Laboratories Inc.). The dialyzed sample was concentrated to ~500 µl using Amicon Ultra 15-ml 3K centrifugal filters (Merck Millipore) according to the manufacturer’s instructions.

Ni-NTA–purified Ub–His10–NexH-Vc7.2 was cleaved using His6–TEV protease (obtained as described below), which was activated by incubation with 10 mM dithiothreitol (DTT) for 30 min. To avoid reduction of disulfide bonds in the target fusion protein, DTT was diluted to 0.2 mM in the TEV solution for 30 min. To avoid reduction of disulfide bonds in the target Ub–His10–NextH-Vc7.2, liberated Ub–His10 and His6–TEV, molar ratio of 100:1 was used. To remove any uncleaved sample, and the mixture was incubated overnight at 4 °C on a rotating mixer. Talon beads were removed by centrifugation, and the supernatant was subjected to RP-HPLC. Trifluoroacetic acid (TFA) was added to 0.1%; pH was adjusted to ~2 with 1 M HCl, and samples were spun at 16,100 × g for 10 min. RP-HPLC purification was performed on an ÄKTA purifier 900 chromatography system equipped with a Grace Vydac 218 TP C18 column (4.6 × 150 mm, 5 µm) using solvent A: 5% ethanol, 0.1% TFA, and solvent B: 90% ethanol, 0.085% TFA. The 0–60% solvent B gradient was developed over 60 min with a flow rate of 1 ml/min. Following purification, peptide-containing fractions from the major peak were pooled and lyophilized. RP-HPLCs (and plots of molar ellipticity; see below) were visualized in MATLAB (MathWorks®). Ub–His10–rH-Vc7.2 was purified using the same procedure.

His6–TEV protease expression and purification

An expression plasmid (pLE478) encoding maltose-binding protein (MBP) and His6-tagged TEV protease (MBP–His6–TEV) (20) was a gift from V. Rogov and V. Dötsch (Goethe University, Frankfurt, Germany). Upon expression, the fusion protein undergoes autocleavage to release TEV protease with an N-terminal His-tag (His6–TEV). Chemically competent BL21(DE3) cells (Invitrogen) were transformed with pLE478 and plated on an LB agar plate containing 100 µg/ml ampicillin. One colony was inoculated in 10 ml of LB medium containing 100 µg/ml ampicillin and grown overnight at 37 °C in an orbital shaker (200 rpm). The overnight culture was diluted to OD600 of 0.1 and incubated at 37 °C in an orbital shaker until the OD600 reached ~1.0. Subsequently, IPTG was added to the culture to a concentration of 1 mM and incubated at 20 °C overnight. The His6–TEV protease was purified from overnight cultures essentially as the Ub–His10–NexH-Vc7.2 construct, but using different lysis buffer (50 mM Tris, 200 mM NaCl, 1% glycerol, 10 mM imidazole, pH 7.8), Ni-NTA–binding buffer (50 mM Tris, 200 mM NaCl, 20 mM imidazole, pH 7.8), elution buffer (50 mM Tris, 200 mM NaCl, 300 mM imidazole, pH 7.8), and dialysis buffer (50 mM Tris–HCl, 100 mM NaCl, pH 7.8). The concentration of dialyzed His6–TEV protease was calculated based on the A280 value and an extinction coefficient, ε280 of 32,290 M⁻¹ cm⁻¹. Purified His6–TEV protease was stored in dialysis buffer containing 5 mM DTT and 50% glycerol at ~80 °C.

Concentration determination

As the H-Vc7.2 sequence contains no Trp or Tyr residues, the concentrations of NexH-Vc7.2 and rH-Vc7.2 were determined by measuring absorbance at 214 nm and using the extinction coefficient 34,621 M⁻¹ cm⁻¹ calculated as reported previously (49).

Confirmation of peptide oxidation state

The oxidation status of purified NexH-Vc7.2 was evaluated by MALDI-TOF MS using an Autoflex Smartbeam III instrument (Bruker) calibrated by external calibration (Peptide calibration standard I; Bruker Daltonics). Lyophilized rH-Vc7.2 was dissolved in 20 µl of 20 mM Tris, pH 8.0, containing 6 mM guanidinium hydrochloride and incubated 2 h at 37 °C in the presence or absence of 50 mM TCEP. Subsequently, 100 mM iodoacetamide was added and samples were incubated for 2 h at room temperature in the dark to allow for S-carbamidomethylation of free thiols. For MS analysis, the samples were acidified by the addition of TFA and desalted by reversed-phase chromatography using Poros R1 microcolumns as described previously (50). The recovered samples were mixed with α-cyano-4-hydroxycinnamic acid prepared in 70% acetonitrile, 0.1% TFA, spotted on a stainless-steel target plate, and analyzed in positive reflectron mode. The generated data were evaluated using the GPMAW software and monoisotopic masses.

MS confirmation of csPDI and hPDI expression

Lysates from IPTG-induced E. coli cells were separated by reducing SDS-PAGE, and in-gel digestion was performed using...
Granulin fold arising from a common ICK cysteine framework

porcine trypsin (Promega). Generated peptides were recovered by reversed-phase chromatography (StageTips, C18, Thermo Fisher Scientific) and eluted directly onto the target plate using α-cyano-4-hydroxyquinolinic acid as described above. Peptides were analyzed in positive reflector mode, annotated, and interrogated by peptide mass fingerprinting using the MASCOT search engine (51). Cysteine propionamide was set as a fixed modification, due to the reaction of protein cysteine residues with acrylamide present in the gel support during electrophoresis. Oxidation of methionine residues was set as a variable modification. One missed cleavage was allowed, and the peptide mass tolerance was set to 50 ppm. Identifications were validated by LIFT analysis on selected peptides.

Determination of cysteine connectivity

The cysteine connectivity in N_{ex}-H-Vc7.2 was determined following the method of Albert et al. (22). In brief, 5 nmol of the peptide were sequentially reduced with 500 mM TCEP and alkylated at pH 3.0 on an Empore C18 solid-phase extraction cartridge (3M, St. Paul, MN). For sequential alkylation, the maleimides NMM, NEM, and NCM were used (at 20 mM). The peptide was run on an ACQUITY BEH C18 UPLC column (1.0 × 150 mm, 1.7 μm) installed in an ACQUITY UPLC I-Class system (Waters) and coupled to a Xevo G2 Q-TOF mass spectrometer (Waters). Different alkylation patterns and subsequently the disulfide connectivity of the peptide were identified from characteristic combinations of y-type ions from obtained mass spectra.

RP-HPLC elution experiments with rH-Vc7.2 and C. victoriae venom

Approximately 1 nmol of rH-Vc7.2 was analyzed by RP-HPLC on an analytical C18 column (5-μm particle size, 4.6 × 250 mm, Vydac) using a gradient of 15–45% solvent B (90% acetonitrile, 0.1% TFA) over 30 min. The HPLC column was repeatedly washed with 5–100% solvent B to ensure that no more recombinant peptide was bound to the column prior to venom analysis. Venom was extruded from a single frozen venom duct of C. victoriae and extracted in 400 μl of 40% acetonitrile, 0.1% TFA with a disposable plastic pestle. The mixture was centrifuged for 5 min at 10,000 g to remove insoluble material. A 10-μl aliquot of the supernatant was diluted in 0.1% TFA, water and analyzed by RP-HPLC using the same conditions described above. Fractions eluting at similar % of solvent B as rH-Vc7.2 were collected in 0.25% B intervals for subsequent mass spectrometric analysis.

Fractions were lyophilized and submitted to the Proteomics Core Facility at the University of Utah. MALDI-TOF MS analysis was performed on rH-Vc7.2 and venom fractions eluting at a similar percent of solvent B. MALDI-MS measurements were performed on a Bruker Daltonics UltraflxeXtreme ToF mass spectrometer in positive ion reflectron mode using pulsed ion extraction set at 100 ns. The instrument was calibrated using polyalanine, and samples were analyzed by averaging ~500 single-shot spectra. Data were analyzed using flexAnalysis 1.4 (Bruker).

CD spectroscopy

CD spectroscopy of oxidized and reduced rH-Vc7.2 (40 μM in 10 mM KH₂PO₄/K₂HPO₄, pH 6.1, in the absence of TCEP or 25 mM NaH₂PO₄/Na₂HPO₄, pH 6.1, in the presence of 5 mM TCEP) was carried out using a JASCO J-810 spectropolarimeter. The fully reduced sample was prepared by incubating rH-Vc7.2 overnight with 10 mM DTT and 10 mM TCEP before purification by RP-HPLC. The lyophilized sample was dissolved in H₂O for concentration determination before dilution into phosphate buffer. For thermostability assays, ellipticity at 224 nm was recorded from 10 to 90 °C at a rate of 1 °C/min followed by cooling to 5 °C. CD spectra were recorded at 25 °C in the wavelength range of 190–250 nm before and after heating. The final spectra were obtained after averaging over 15 spectra recorded at a scan rate of 10 nm/min, and baseline (buffer only, same conditions) was subtracted. High-frequency noise was removed using a Fourier transformation filter. The region from 190 to 200 nm was omitted due to a high-tension voltage value exceeding 600 Hz in the presence of TCEP.

NMR spectroscopy

Samples of ~1 mM of both unlabeled and 13C,15N-labeled N_{ex}-H-Vc7.2 were prepared in 25 mM sodium phosphate, 10% D₂O, pH 6.1. Chemical shifts were assigned by standard sequential assignment based on the following recorded spectra: 2D 1H-1H-TOCSY and 2D 1H-1H-NOEY (on the unlabeled sample), and 13N HSQC, 13C HSQC, HNCA, HNCO, HNCOCA, HNCACO, and CBCACONH (on the double-labeled sample). All spectra were recorded at 25 °C on Bruker Avance III-HD NMR spectrometers operating at 750 and 600 MHz, equipped with TCI and QCI cryoprobes, respectively. The data were processed with nmrPipe (52) and analyzed in CCPNMR (53).

NMR structure calculations

CYANA was used for automated NOE assignment of the 2D 1H-1H NOESY spectrum and initial structure calculation, also including dihedral angles calculated with TALOS+ (54) and the disulfide pattern (C8–C18/C13–C23/C17–C28, determined as described above) as restraints (55). The final structure refinement was done using XPLOR-NIH with the NOE-derived distance restraints, dihedral angle restraints, and disulfide bond restraints (56). The 20 lowest energy structures out of 100 were selected and contained no NOE violations (>0.3 Å) or dihedral angle violations > 5°. Ramachandran-plot statistics were calculated using PROCHECK (57). Structure alignment was done using MOLMOL (58), and structure visualization was performed with PyMOL (DeLano Scientific). The structure has been deposited in the PDB (59) with PDB code 6Q5Z and chemical shifts and chemical shifts have been deposited in the Biological Magnetic Resonance Bank (http://www.bmrb.wisc.edu/ (60)) with ID 34335.

Identification of structural homologs

To identify structural homologs of N_{ex}-H-Vc7.2, we first downloaded all structures in the PDB, and the sequences from
the structure files were extracted using an in-house automated Matlab script. The extracted sequences were then searched with a $\text{CX}_{1-5}\text{CX}_{1-3}\text{CX}_{0-2}\text{C}$ sequence motif (where $X$ denotes any residue), representing a fragment of the H-Vc7.2 sequence covering the first four cysteine residues. All structures containing a sequence matching this motif were inspected in PyMOL. Structures containing a fold homologous to that of N$_\text{exo}$H-Vc7.2 were aligned to the N$_\text{exo}$H-Vc7.2 structure using the CLICK alignment server (61, 62).

**Mouse bioassay and calcium imaging**

All experiments involving the use of animals were approved by the Institutional Animal Care and Use Committee of the University of Utah. Swiss Webster mice (14–21 days old) were injected intracranially with 3.5 nmol of rH-Vc7.2 peptide (dissolved in 25 mM NaHPO$_4$ buffer, pH 6.1, containing 10% D$_2$O, 0.25 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid, 0.02% NaN$_3$). Following intracranial injection, mouse behavior was observed for 1 h to determine differences between treated and control animals. Calcium imaging on mouse lumbar dorsal root ganglion neurons, using 10 $\mu$M peptide, was carried out as described previously (63).

**Acknowledgments—**We thank Nick Williamson, University of Melbourne, and the Mass Spectrometry and Proteomics Core Facility at the University of Utah (spectrometry equipment obtained through a Shared Instrumentation Grant 1 S10 OD018210 01A1) for technical assistance with MS. Norut Northern Research Institute was supported by Tromsø County Council Grant TFK2013-252.

**References**

Granulin fold arising from a common ICK cysteine framework

meant compounds into cells. J. Biol. Chem. 283, 27048–27056 CrossRef Medline
Granulin fold arising from a common ICK cysteine framework


The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework


doi: 10.1074/jbc.RA119.007491 originally published online April 11, 2019

Access the most updated version of this article at doi: 10.1074/jbc.RA119.007491

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 61 references, 5 of which can be accessed free at
http://www.jbc.org/content/294/22/8745.full.html#ref-list-1