



Article

Extent of N-Terminus Folding of Semenogelin 1 Cleavage Product Determines Tendency to Amyloid Formation

Daria A. Osetrina ¹, Aleksandra M. Kusova ^{1,2} , Aydar G. Bikmullin ^{1,3}, Evelina A. Klochkova ^{1,3},
Aydar R. Yulmetov ¹ , Evgenia A. Semenova ¹, Timur A. Mukhametzyanov ¹ , Konstantin S. Usachev ^{3,4} ,
Vladimir V. Klochkov ^{1,*} and Dmitriy S. Blokhin ¹

- ¹ NMR Laboratory, Medical Physics Department, Institute of Physics, Kazan Federal University, Kremlevskaya Str., 18, Kazan 420008, Russia; dmitr.blokhin@gmail.com (D.S.B.)
² Kazan Institute of Biochemistry and Biophysics, FRC Kazan Scientific Center, Russian Academy of Sciences, Kazan 420111, Russia
³ Laboratory of Structural Biology, Institute of Fundamental Medicine and Biology, Kazan Federal University, Kazan 420021, Russia
⁴ Laboratory for Structural Analysis of Biomacromolecules, Federal Research Center "Kazan Scientific Center of Russian Academy of Sciences", Kazan 420111, Russia
* Correspondence: vladimir.klochkov@kpfu.ru; Tel.: +7-(843)-233-7634

Abstract: It is known that four peptide fragments of predominant protein in human semen Semenogelin 1 (SEM1) (SEM1(86–107), SEM1(68–107), SEM1(49–107) and SEM1(45–107)) are involved in fertilization and amyloid formation processes. In this work, the structure and dynamic behavior of SEM1(45–107) and SEM1(49–107) peptides and their N-domains were described. According to ThT fluorescence spectroscopy data, it was shown that the amyloid formation of SEM1(45–107) starts immediately after purification, which is not observed for SEM1(49–107). Seeing that the peptide amino acid sequence of SEM1(45–107) differs from SEM1(49–107) only by the presence of four additional amino acid residues in the N domain, these domains of both peptides were obtained via solid-phase synthesis and the difference in their dynamics and structure was investigated. SEM1(45–67) and SEM1(49–67) showed no principal difference in dynamic behavior in water solution. Furthermore, we obtained mostly disordered structures of SEM1(45–67) and SEM1(49–67). However, SEM1(45–67) contains a helix (E58–K60) and helix-like (S49–Q51) fragments. These helical fragments may rearrange into β -strands during amyloid formation process. Thus, the difference in full-length peptides' (SEM1(45–107) and SEM1(49–107)) amyloid-forming behavior may be explained by the presence of a structured helix at the SEM1(45–107) N-terminus, which contributes to an increased rate of amyloid formation.

Keywords: SEM1(49–107); SEM1(45–107); Semenogelin 1; amyloid; HIV; NMR spectroscopy; CD spectroscopy; DLS spectroscopy; ThT fluorescence; spatial structure



Citation: Osetrina, D.A.; Kusova, A.M.; Bikmullin, A.G.; Klochkova, E.A.; Yulmetov, A.R.; Semenova, E.A.; Mukhametzyanov, T.A.; Usachev, K.S.; Klochkov, V.V.; Blokhin, D.S. Extent of N-Terminus Folding of Semenogelin 1 Cleavage Product Determines Tendency to Amyloid Formation. *Int. J. Mol. Sci.* **2023**, *24*, 8949. <https://doi.org/10.3390/ijms24108949>

Academic Editors: Koichi Kato and Stanislaw Oldziej

Received: 23 March 2023

Revised: 15 May 2023

Accepted: 16 May 2023

Published: 18 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Human immunodeficiency virus (HIV) was first identified in 1981 [1]. This is one of the most dangerous diseases due to its influence on cells of the human immune system. Over time, HIV causes acquired immunodeficiency syndrome (AIDS) which may lead to susceptibility to various infections and tumors and eventually death [2]. Since the main route of transmission of the virus is unprotected intercourse, seminal fluid is considered as the main factor in increasing HIV activity [3].

Human semen forms a coagulum immediately after ejaculation. Semenogelin 1 (SEM1) and Semenogelin 2 (SEM2) are predominant proteins in seminal vesicles occupying approximately 60% of the ejaculate volume [4–8]. Both proteins originate from the glandular epithelium of the seminal vesicles, which secrete them in high concentrations (60 g/L) [9]. Semenogelins 1 and 2 are the predominant structural proteins of the loose gel formed in

freshly ejaculated human semen. The concentration of SEM1 is five to ten times higher compared with SEM2 in semen. Semenogelins perform many biochemical functions, such as semen liquefaction, antibacterial activity, protection and other functions of semen [9]. The Semenogelins participate in noncovalently linked gel-like structure formation that ensnares the spermatozoa. Within the first 20 min, the gel-like structure is liquefied by serine proteases, primarily prostate-specific antigen (PSA), which cleaves the full-sized proteins into small fragments [10].

In previous works, it was demonstrated that certain cationic polypeptides within the semen and seminal plasma could form amyloid fibrils amplifying HIV infection [11]. Semenogelin cleavage products found in semen in high concentration could be responsible for an increase in the activity of HIV virions in seminal fluid. The analysis of semen amyloid compounds showed the presence of SEM1 fragments (SEM1(45–107), SEM1(49–107), SEM1(68–107) and SEM1(86–107)) (Figure 1) [12,13].

Peptide	Amino acid sequence
SEM1 (86–107)	DLNALHKTTKSQRHLGGSQQLL
SEM1 (68–107)	TYHVDANDHDQSRKSQQYDLNALHKTTKSQRHLGGSQQLL
SEM1 (49–107)	SGQKKGKQQTESKGSFSIQYTYHVDANDHDQSRKSQQYDLNALHKTTKSQRHLGGSQQLL
SEM1 (45–107)	GQHY SGQKKGKQQTESKGSFSIQYTYHVDANDHDQSRKSQQYDLNALHKTTKSQRHLGGSQQLL

Figure 1. Amino acid sequences of amyloidogenic peptides from Semenogelin 1.

It is believed that semen amyloid fibrils enhance the adhesion of HIV virions to cell membranes by decreasing electrostatic repulsion between membranes of the virus and target cell [14,15]. Additionally, studies have shown that amyloid fibrils from Semenogelin 1 play a physiological role in the process of fertilization [16]. The most obvious similarity between these processes is their same physiological circumstances [17]. The fusion of semen and oocyte membranes, as well as HIV and target cell membranes, is an energetically unfavorable process that requires many different cooperative protein–protein interactions [18–20].

Herein, we report the discovery of amyloidogenic peptides SEM1(45–107) and SEM1(49–107), which are derived from human SEM1 (Figure 1). A clear difference was observed in their amyloid-forming process rate via Thioflavin T (ThT) fluorescence spectroscopy. We observed a more intensive aggregation process of SEM1(45–107) in comparison with SEM1(49–107) via ThT fluorescence and TEM data. Interestingly, the primary structure of SEM1(45–107) differs from SEM1(49–107) only by the presence of four amino acid residues (G45Q46H47Y48) at N-domains. In this regard, the N-domains of SEM1(45–107) and SEM1(49–107) (SEM1(45–67) and SEM1(49–67), respectively) have been characterized using structural and dynamics NMR methods, circular dichroism (CD) and dynamic light scattering (DLS) spectroscopies.

Nuclear magnetic resonance spectroscopy (NMR) is an informative method for determining the spatial structure of proteins and peptides in solution [21–26]. Pulsed-field gradient NMR spectroscopy (PFG-NMR) and dynamic light scattering spectroscopy (DLS) are effective methods for studying the translational mobility of a protein in solution [27–32]. CD was applied for the verification of the protein secondary structure [33]. In our discussion, we will describe the difference in the amyloid behavior of SEM1 cleavage products based on an individual study of the N-domains (SEM1(49–67) and SEM1(45–67)) of SEM1(49–107) and SEM1(45–107). Based on structural studies of N-domains and the Molecular dynamics (MD) simulation of full-sized peptides, we suppose the responsibility of helical fragments in N-terminuses of SEM1(45–107) in fast amyloid formation. In a previous article, Lui et al. (2010) [34] reported monomer helices' rearrangement into amyloid β -strand structures during the formation of amyloid fibrils. Thus, the presence of helix-like structures of SEM1(45–107) may lead to an increase in the amyloid-forming rate in comparison with SEM1(49–107).

2. Results

2.1. Thioflavin T Fluorescence Measurements of SEM1(49–107) and SEM1(45–107)

Figure 2A depicts the ThT fluorescence spectra of SEM1(45–107) and SEM1(49–107) peptides and a control pure ThT solution (black line). The red and green lines show the ThT fluorescence spectra of the SEM1(49–107) and SEM1(45–107) solutions, respectively. The red and black lines are almost identical, pointing to the absence of the amyloid fibril formation of SEM1(49–107) in water solution. In the case of SEM1(45–107), ThT fluorescence intensity (green line) increases by 4–5 times, indicating the presence of amyloid fibrils in the peptide solution.

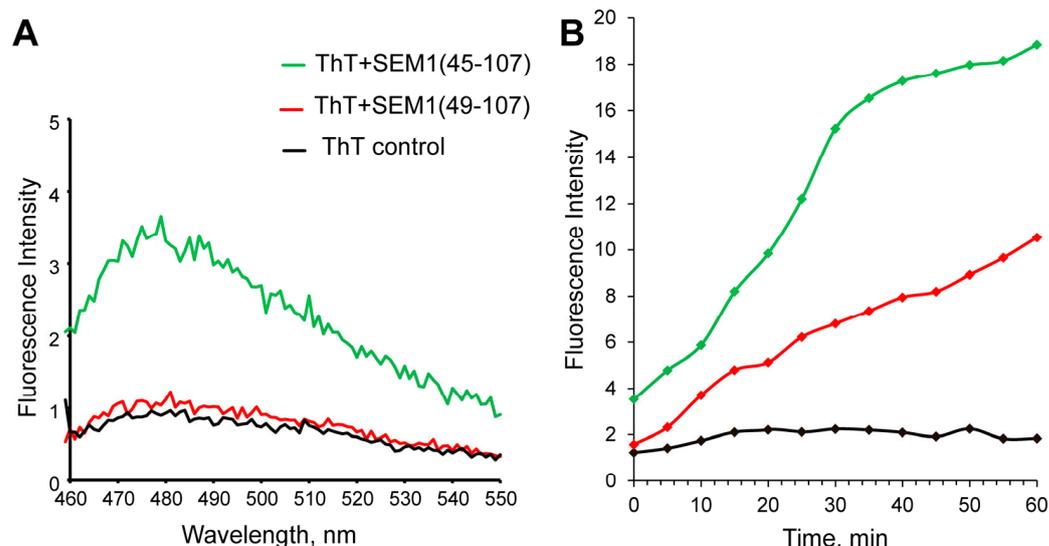


Figure 2. (A) Thioflavin T (ThT) fluorescence spectra of SEM1(49–107) (red line), SEM1(45–107) (green line) and control (black line). (B) Real-time ThT fluorescence assays of SEM1(49–107) (red line), SEM1(45–107) (green line) and control (black line) solutions.

As mentioned above, only fresh peptide solutions (up to 10 min after synthesis) were used for ThT fluorescence measurements. Therefore, the fibril formation of SEM1(45–107) starts immediately after purification in comparison with SEM1(49–107). However, the increased ThT fluorescence intensity measured 30 min after purification showed the amyloid formation of both SEM1(49–107) and SEM1(45–107) peptides (see Supplementary Material Figure S6). The time dependence of the fluorescence intensity of ThT showed less time for the amyloid formation of SEM1(45–107) in comparison with SEM1(49–107) (Figure 2B). We used transmission electron microscopy (TEM) data to detail the fibril formation process. TEM images of SEM1(45–107) and SEM1(49–107) fibrils were recorded an hour after preparation (Figure 3). TEM images show that the length of SEM1(45–107) fibrils is larger than that of SEM1(49–107) ones. Thus, TEM data are in good agreement with ThT data, where we observed the more intensive fibril formation of SEM1(45–107) compared with SEM1(49–107).

2.2. Nuclear Magnetic Resonance Spectroscopy

According to ThT fluorescence spectroscopy data, an unstable monomeric state of SEM1(49–107) and SEM1(45–107) peptides in water solution was shown, which made it impossible to study their structure via NMR spectroscopy. Seeing that the peptide amino acid sequence of SEM1(45–107) differs from SEM1(49–107) only by the presence of four additional amino acid residues in the N-domain, these domains of both peptides were obtained via solid-phase synthesis and studied via NMR spectroscopy. The ^1H , ^{15}N and ^{13}C chemical shifts of SEM1(45–67) (BMRB-51673) and SEM1(49–67) (BMRB-51691) assignment were determined with the help of NMR spectra (2D ^1H - ^1H TOCSY, ^1H - ^1H

NOESY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, ^1H - ^{13}C HSQC-TOCSY and ^1H - ^{15}N HSQC (Figure 4 and Figure S1)). The experimental restraints for the structural calculation were determined from NMR data (Table 1 for SEM1(45–67); Table 2 for SEM1(49–67)). The $^3J_{\text{H}\alpha\text{-HN}}$ scalar coupling constants were determined via the 2D JRES spectrum analysis. The values of the $^3J_{\text{H}\alpha\text{-HN}}$ scalar coupling constant were used to determine ψ dihedral angle restraints via Karplus curves [35]. The H, N, $\text{H}\alpha$, $\text{C}\alpha$ and C chemical shifts were applied for the determination of dihedral angles using the TALOS+ service (<https://spin.niddk.nih.gov/bax/nmrserver/talos>, accessed on 25 November 2022). Distance constraints (Table S1 for SEM1(45–67), Table S2 for SEM1(49–67)) were determined with the help of ^1H - ^1H NOESY spectra analysis, where high-intensity signals corresponded to short distances between nucleuses, and low-intensity signals corresponded to long distances.

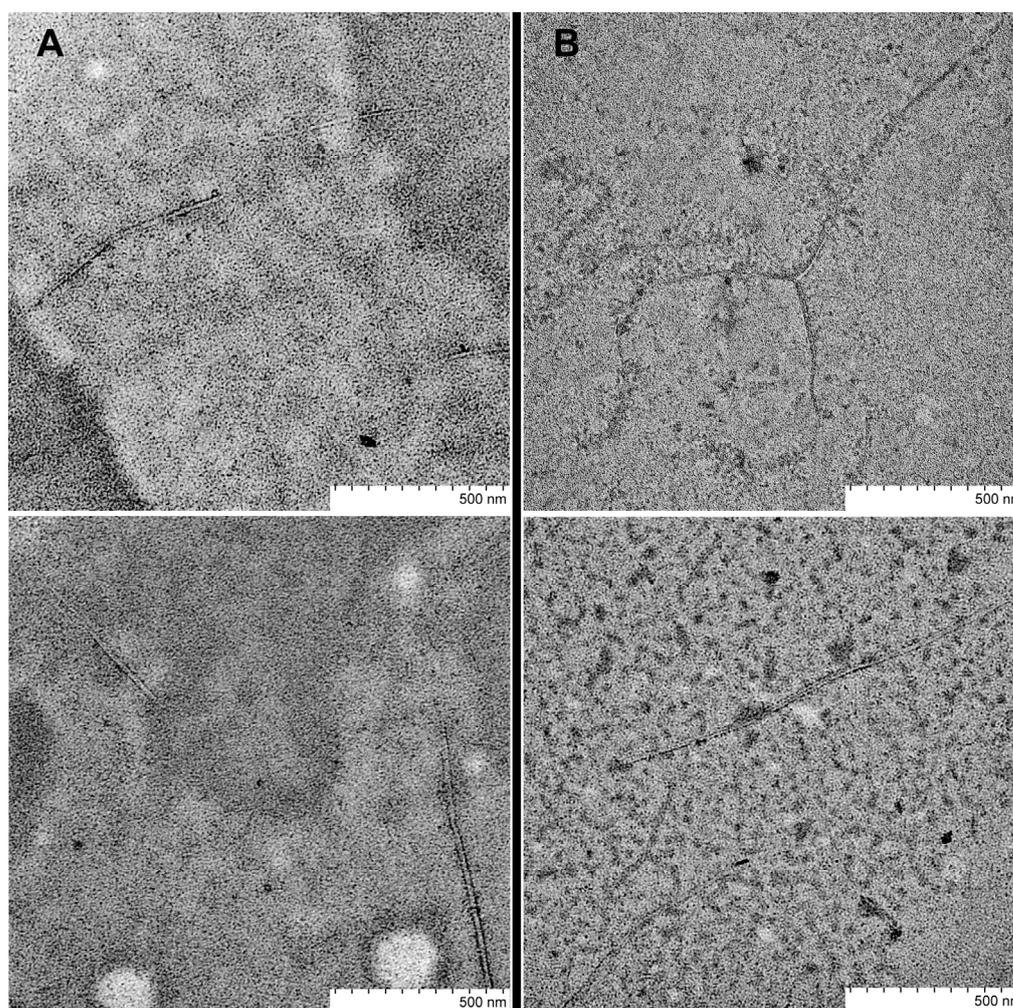


Figure 3. TEM images of SEM1(49–107) (A) and SEM1(45–107) (B) fibrils.

The NOE data of SEM1(45–67) provide 143 distance restraints, including 88 intra-residual distances and 55 inter-residual distances (24 distances between sequential amino acid residues ($i-j = 1$), 16 distances between medium-distance amino acid residues ($i-j = 2-6$) and 15 distances between long-range ones ($i-j > 6$)) (Table S1). The NOE data of SEM1(49–67) provide 67 distance restraints including 54 distances within amino acid residues and 13 distances between spatially close amino acid residues (8 distances between sequential amino acid residues ($i-j = 1$), 2 distances between medium-distance amino acid residues ($i-j = 3$) and 3 distances between long-range ones ($i-j > 4$)) (Table S2).

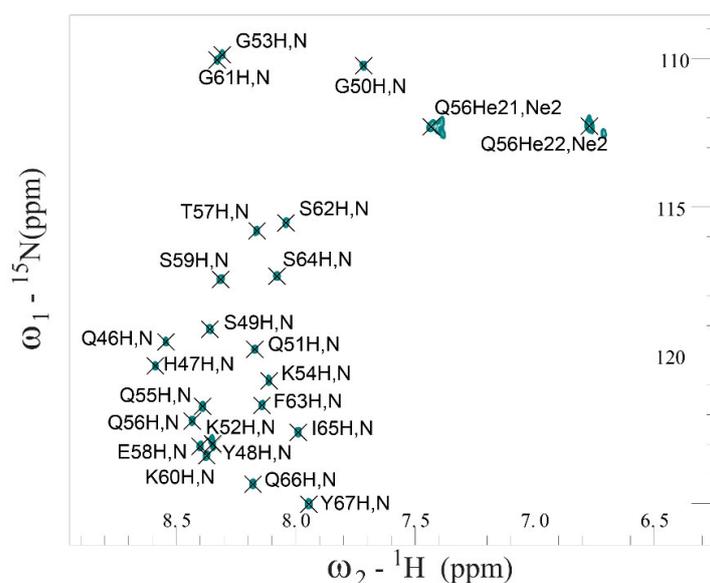


Figure 4. Fragment of 2D ^1H - ^{15}N HSQC NMR spectrum of SEM1(45–67) in water solution at 298 K.

Table 1. Statistical information for SEM1(45–67) structural ensemble.

Distance Constraints	
total	143
intra-residual	88
inter-residual	55
sequential ($i-j = 1$)	24
medium ($i-j = 2-6$)	16
long-range ($i-j > 6$)	15
$^3\text{J}_{\text{H}\alpha\text{-HN}}$ alpha coupling constants (jna)	
Restrains	22
Structural Statistics	
number of NOE violations	0
number of jna restraint violations	0
RMSD for bond deviations (\AA)	0.005
RMSD for angle deviations (deg)	0.598
RMSD of all backbone atoms (\AA)	
E58-K60	0.13 ± 0.04
S49-Y67	0.39 ± 0.11
Y48-K52	0.14 ± 0.06
Ramachandran Plot	
residues in the most favored region (%)	80.1
residues in the additionally allowed region (%)	19.9
residues in the disallowed region (%)	0.0

The distances constraints, dihedral angles and $^3\text{J}_{\text{H}\alpha\text{-HN}}$ scalar coupling constants were used as input parameters for the spatial structure calculation with the help of XPLOR-NIH (v. 3.6) software. Figure 5 shows structural ensembles of SEM1(45–67) (PDB ID: 8BOO) (Figure 5A) and SEM1(49–67) (PDB ID: 8BVZ) (Figure 5B). Ramachandran plots of SEM1(45–67) (Figure S2) and SEM1(49–67) (Figure S3) validated the obtained peptides' structures [36]. Both peptides have mostly unordered structures and have a horseshoe shape, which is consistent with a large number of long-range distances ($i-j > 6$) of 15 and 3 for SEM1(45–67) and SEM1(49–67), respectively. DSSP analysis (<http://bioinformatica.isa.cnr.it/SUSAN/NAR2/dsspweb.html>, accessed on 11 January 2023) [37] of the calculated structures (Table S3) showed that SEM1(45–67) contains 3_{10} -helix (E58-K60) and ordered helix-like fragments (S49-Q51), while any ordered regions in the case of SEM1(49–67) were

not observed (Table S4). The presence of 3_{10} -helix (E58-K60) was confirmed via the NOEs characteristics [38] ($d_{\alpha,N}(i, i + 3)$: E58H $_{\alpha}$ -G61H $_N$, $d_{\alpha,\beta}(i, i + 3)$: T57H $_{\alpha}$ -K60H $_{\beta}$).

Table 2. Statistical information for SEM1(49–67) structural ensemble.

Distance Constraints	
total	67
intra-residual	54
inter-residual	13
sequential (i-j = 1)	8
medium (i-j = 3)	2
long-range (i-j > 13, 14)	3
Dihedral Angle Constraints	
Phi restraints	17
Psi restraints	17
$^3J_{H_{\alpha}-HN}$ alpha coupling constants (jna)	
Restraints	16
Structural Statistics	
number of NOE violations	0
number of dihedral angle restraint violations	0
number of jna restraint violations	0
RMSD for bond deviations (Å)	0.004
RMSD for angle deviations (deg)	0.794
RMSD of all backbone atoms (Å)	
E58-K60	0.24 ± 0.10
S49-Y67	1.12 ± 0.30
Y49-K52	0.69 ± 0.26
T57-S59	0.17 ± 0.08
G50-T57	0.70 ± 0.25
Ramachandran Plot	
residues in the most favored region (%)	94.1
residues in the additionally allowed region (%)	5.9
residues in the disallowed region (%)	0.0

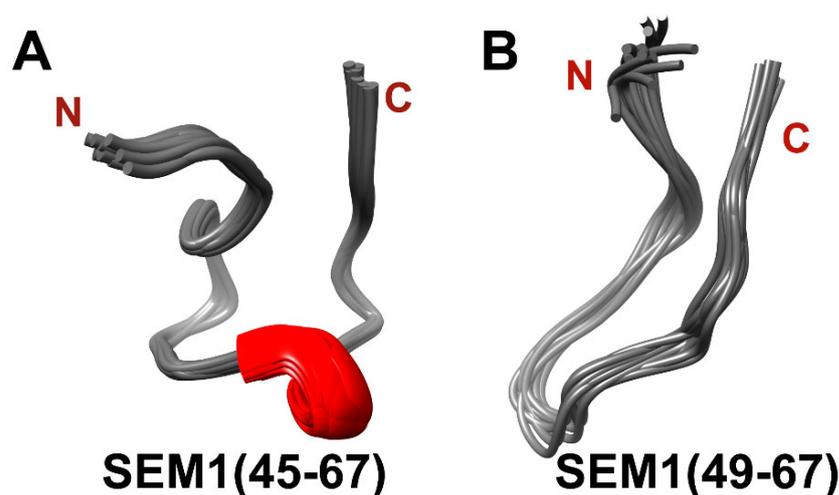


Figure 5. (A) Spatial structure of SEM1(45–67) ensemble (11 structures); (B) spatial structure of SEM1(49–67) ensemble (12 structures).

2.3. Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy was used for the validation of the calculated peptides' structures. The CD spectra of SEM1(45–67) (Figure 6A) and SEM1(49–67) (Figure 7A) are typical for the prevailing of an unordered polypeptide chain. The analysis of the CD spectra (Figures 6B and 7B) using the DichroWeb web service provided the following

secondary structure components: SEM1(45–67) consists of distorted helices (24%), regular β -strands (2%), distorted β -strands (11%), turns (33%) and unordered structural fragments (30%); SEM1(49–67) consists of distorted helices (6%), regular β -strands (3%), distorted β -strands (14%), turns (37%) and unordered structural (40%) fragments. The RMSD values of the CD spectra analysis were 0.008 and 0.017 for SEM1(45–67) and SEM1(49–67), respectively. Thereby CD data also confirmed the mainly disordered structure of both peptides obtained via NMR spectroscopy. The components of regular β -strands and distorted β -strands are in agreement with the horseshoe shape of the peptides obtained via the NMR structures.

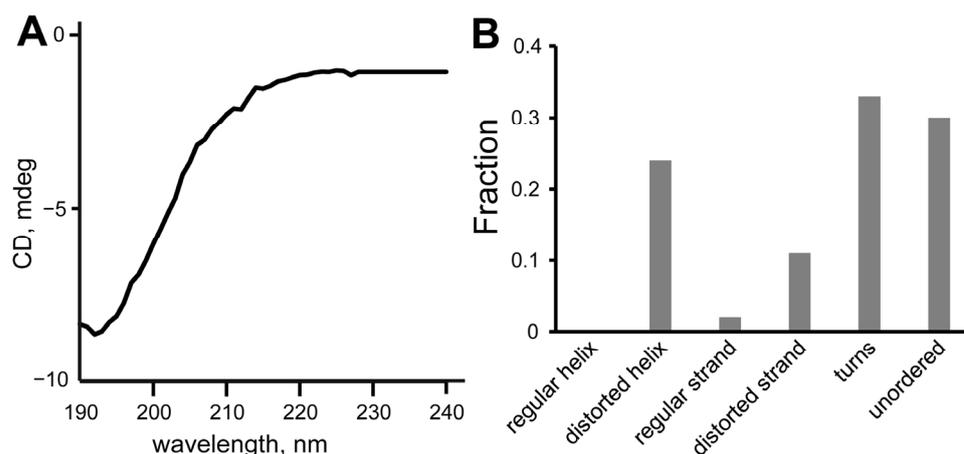


Figure 6. (A) CD spectrum of SEM1(45–67), (B) component analysis of SEM1(45–67) CD spectrum.

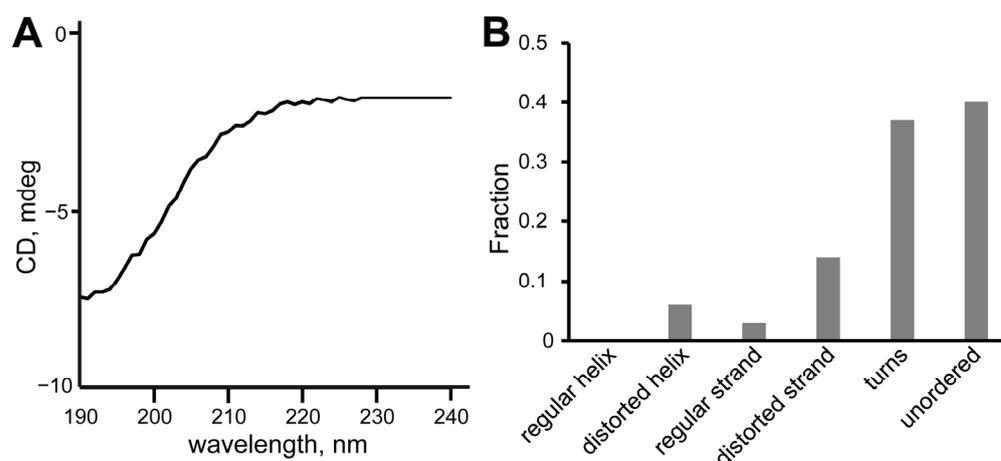


Figure 7. (A) CD spectrum of SEM1(49–67); (B) component analysis of SEM1(49–67) CD spectrum.

2.4. PFG-NMR Spectroscopy of SEM1(45–67) and SEM1(49–67)

DOSY spectra of peptides SEM1(45–67) and SEM1(49–67) were recorded to confirm that the peptides were in monomeric forms. For SEM1(45–67) and SEM1(49–67), the mono exponential diffusion decays of signals at 8.02–7.94 ppm were observed (Figure 8). Diffusion decays were characterized by self-diffusion coefficients D ($2.07 \pm 0.02 \cdot 10^{-10} \text{ m}^2/\text{s}$ SEM1(45–67)) and $2.35 \pm 0.03 \cdot 10^{-10} \text{ m}^2/\text{s}$ (SEM1(49–67)). We calculated the hydrodynamic radius R_h using the Stokes–Einstein relation:

$$R_h = \frac{k_B T}{6\pi\eta D}, \quad (1)$$

where T is the temperature, k_B is the Boltzmann constant and η is the water viscosity. The calculated hydrodynamic radii are 1.2 nm and 1 nm for SEM1(45–67) and SEM1(49–67),

respectively. These values correspond to monomer forms of peptides and are consistent with NMR structures of SEM1(45–67) and SEM1(49–67).

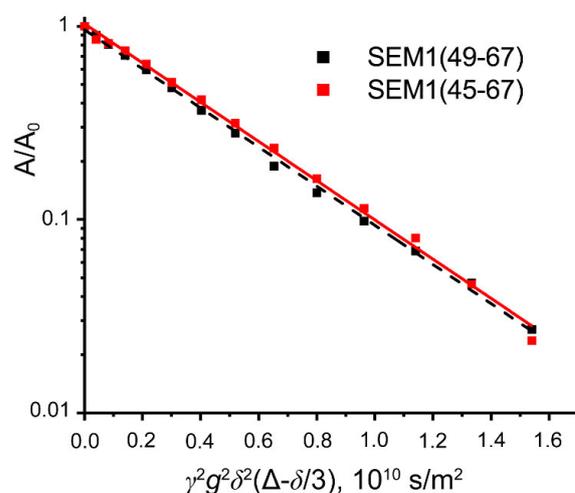


Figure 8. Diffusion decays of signal intensity at 8.02–7.94 ppm for SEM1(45–67) (red) and SEM1(49–67) (black).

2.5. DLS Spectroscopy of SEM1(45–67) and SEM1(49–67)

The oligomerization of SEM1(45–67) and SEM1(49–67) was further monitored via dynamic light scattering (DLS), which enables the observation of NMR-invisible diffusive species since the scattering intensity is very dependent on the particle mass/size [39]. The evolution of the size distribution with time is shown in Figure 9.

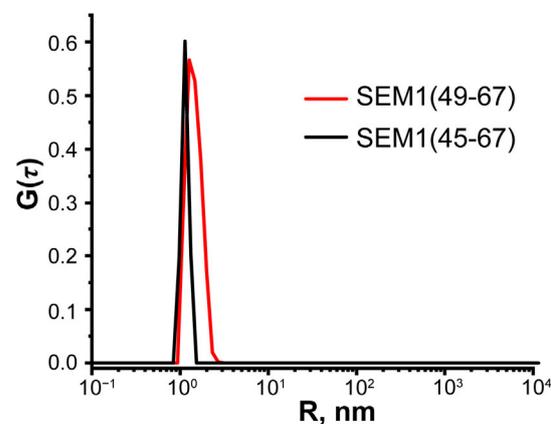


Figure 9. Hydrodynamic radius redistribution for monodisperse water solutions of SEM1(49–67) (red line) and SEM1(45–67) (black line).

The hydrodynamic radii R of SEM1(49–67) and SEM1(45–67) were obtained using the DLS technique. In Figure 9, one can see that SEM1(49–67) and SEM1(45–67) solutions are characterized by single peaks. Meanwhile, the increased peak width of SEM1(49–67) indicates a continuous size distribution of different conformational states. We estimated the hydro-dynamic radius of $R = 1.8$ nm and $R = 1.2$ nm for SEM1(49–67) and SEM1(45–67), respectively. Thus, the R values of peptides are in acceptable agreement with the calculated size of both peptides via the DOSY experiments and their spatial structures.

The hydrodynamic radius R values obtained via the DLS technique are usually strongly influenced by weak intermolecular interactions [40]. The major role of protein–protein interactions is caused by electrostatic interactions. Therefore, the ζ -potential values of the peptide solutions were determined. We did not find principal differences in the ζ -potential values of SEM1(49–67) ($\zeta = +11.84$ mV) and SEM1(45–67) ($\zeta = +17.44$ mV) solutions. On

the basis of ζ -potential values, we calculated the peptide effective charges of SEM1(49–67) ($Q = +2.33e$) and SEM1(45–67) ($Q = +2.28e$), which were found to be close to the calculated ones (<https://www.protpi.ch/Calculator/ProteinTool>).

2.6. MD Simulation of SEM1(45–107) and SEM1(49–107)

All-atom simulations were performed for the evaluation of SEM1(45–107) and SEM1(49–107) spatial structures. The simulation process consisted of several stages. In the first stage, we created files containing experimental spatial constraints (internuclear distances and dihedral angles) collected from restraints of individual peptide fragments (SEM1(45–67) and SEM1(68–107) for SEM1(45–107) and SEM1(49–67) and SEM1(68–107) for SEM1(49–107)). Restraints for SEM1(68–107) were taken from previous work [21]. In the next step of the calculation, SEM1(45–107) and SEM1(49–107) structures were annealed in the XPLOR-NIH (v. 3.6) program, where experimental spatial constraints were used as input parameters. SEM1(45–107) and SEM1(49–107) structures were calculated according to the protocol described above for peptides SEM1(45–67) and SEM1(49–67). From the set of resulting structures, we selected ones that are characterized by the lowest energy. Next, the all-atom molecular dynamics (MD) simulation of these structures was performed using Gromacs (v. 2022) software [41] with NMR distance constraints data for SEM1(45–107) and SEM1(49–107). In the final step of the calculation, MD-trajectories were classified via the clustering analysis tool in GROMACS (gmx cluster) [42] with the corresponding root mean square deviation (RMSD) of C α atoms cut-off to search preferential ensembles of peptide structures. In the Supplementary Materials (Figures S7 and S8), donut plots quantifying the conformational states of SEM1(45–107) and SEM1(49–107) are shown. The diagram represents the fractional occupancies of the four most populated clusters. Each slice represents a distinct state. RMS deviations of C α -C α atom-pair distances were used to define the distance between structures.

Figure 10 shows four populated clusters of SEM1(45–107) containing up to 55% of total protein structures from the MD-trajectory of SEM1(45–107), where Figure 10A–D correspond to 21%, 13%, 11% and 9% of total protein structures, respectively. Figure 11 shows four populated clusters of SEM1(49–107) containing up to 75% of total protein structures, wherein Figures 11A, 11B, 11C and 11D correspond to 32%, 16%, 14% and 13% of them, respectively. All ensembles of SEM1(45–107) contain helical fragments in the N-terminus, which is in good agreement with SEM1(45–67) spatial structure data. At the same time, the C-terminus spatial structures of SEM1(45–107) and SEM1(49–107) are similar. Thus, the differences in the N-terminuses of these peptides do not provide significant effects on the spatial structure of the C-terminuses.

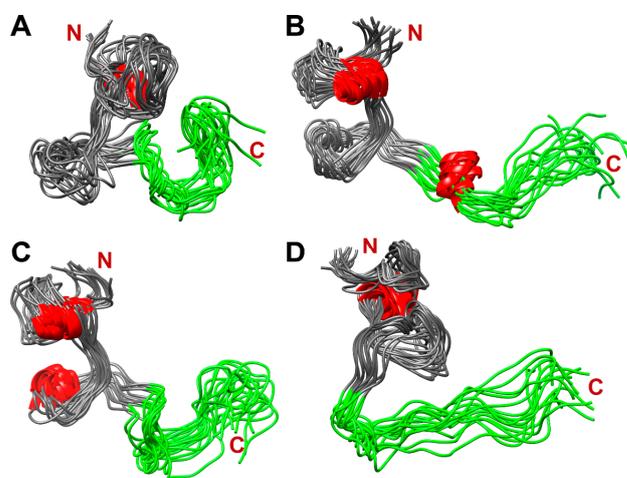


Figure 10. The dominant clusters of SEM1(45–107) holding ~55% of total protein structures: 21% (A), 13% (B), 11% (C) and 9% (D). Red labeled fragments are helical fragments, and green labeled fragments are C-terminuses (86–107) of SEM1(45–107).

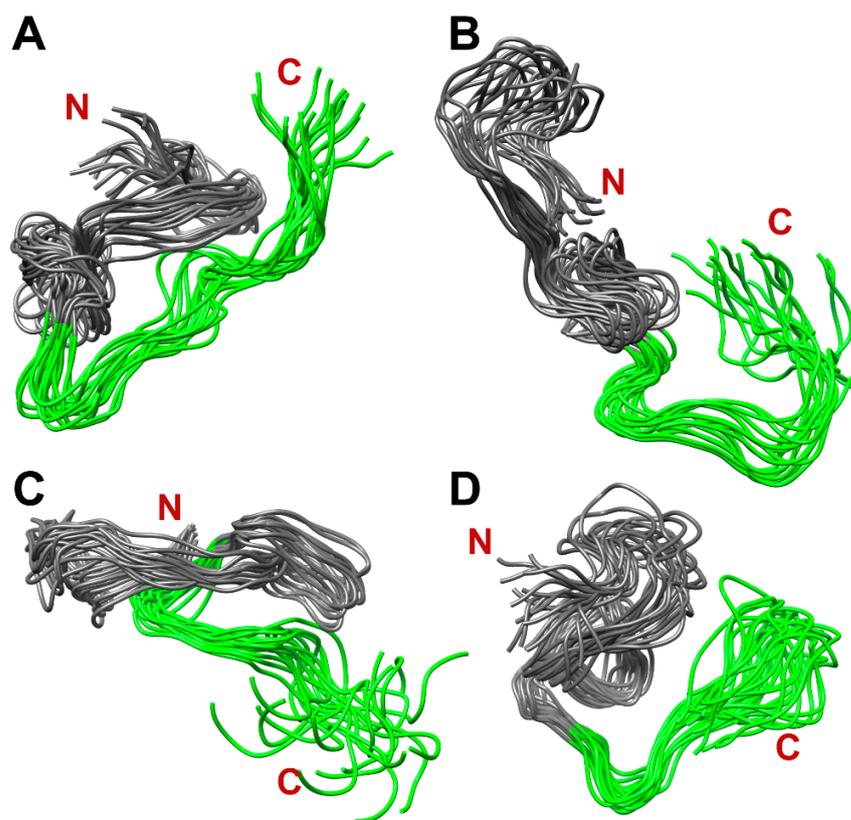


Figure 11. The dominant clusters of SEM1(49–107) holding ~75% of total protein structures: 32% (A), 16% (B), 14% (C) and 13% (D). Green labeled fragments are C-terminuses (86–107) of SEM1(49–107).

3. Discussion

The estimation of the translational mobility of peptides via DLS, DOSY and ζ -potential measurement shows a stable monomeric form of both SEM1(45–67) and SEM1(49–67). For SEM1(45–67) and SEM1(49–67), we obtained a similar distribution of hydrophobic regions (Figure S4) and electrostatic potential surface distribution of SEM1(45–67) and SEM1(49–67) (Figure S5). Thus, SEM1(45–67) and SEM1(49–67) peptides have a similar behavior in water solution (hydrodynamic radius, ζ -potential, hydrophobic surface and electrostatic surface). However, the DLS signal width indicated the less stable conformational state of SEM1(49–67) in comparison with SEM1(45–67).

As it was shown via ThT fluorescence spectroscopy (Chapter 2.1), SEM1(45–107) has a higher amyloid formation rate in comparison to SEM1(49–107) (Figure 2). The various levels of amyloid activity of the studied peptides may be caused by internal differences in their spatial structures. It can be noted that SEM1(45–67) contains helix (E58–K60) and an ordered helix-like structure (S49–Q51) in comparison to SEM1(49–67). In previous articles, the constant secondary structure was observed for individual peptide fragments (PAP(248–261), PAP(262–270) and PAP(274–284)) [43,44], as well as a part of a full-sized PAP(248–286) peptide [45]. Based on the assumption of the secondary structure conservation of individual peptides as well as full-sized protein, the presence of helical fragment SEM1(45–67) in the SEM1(45–107) spatial structure was supposed. Moreover, the MD simulation validates the presence of a helical fragment in the N-terminus of SEM1(45–107) in comparison with SEM1(49–107). The C-terminus (D86–L107) is presented in all amyloidogenic peptides of SEM1. Previously, the C-terminus (D86–L107) spatial structure was published in an article by Sanchugova et al., 2021 [46]. The presence of four extra amino acid residues in the N-terminus of SEM1(45–107) may affect the structure of the C-terminus (D86–L107) with the subsequent difference in the amyloid formation rate of SEM1(45–107) and SEM1(49–107). The peptide structures are necessary to find out this

influence. However, the fast aggregation of SEM1(45–107) and SEM1(49–107) makes it impossible to perform NMR structural studies of these peptides. Therefore, we performed MD simulations of SEM1(45–107) and SEM1(49–107) (Figures 10 and 11) to evaluate the influence of a peptide fragment (G45–Y48) on the C-terminus (D86–L107). The analysis of the MD simulations did not show a principal difference between the spatial structure of the C-terminuses (D86–L107) of the SEM1(45–107) and SEM1(49–107) peptides. The helical fragments were present in all configurations of SEM1(45–107), while SEM1(49–107) obtained all configurations characterized by a disordered structure.

The above results raise the question of what breaks the helix of SEM1(49–107), and how is this all connected to the role of the first four residues in the amyloid formation of SEM amyloidogenic peptides. Residue H47 is positively charged and may form electrostatic contact with residue E58 to stabilize the helix motif of SEM1(45–107). Residue 63 is an hydrophobic phenylalanine that may provide hydrophobic contact with the I65 residue side chain in the helix motif that is not observed in SEM1(49–107). Hence, four extra residues stabilize the helix motif of SEM1(45–107), which contributes to forming the stabilized hydrophobic region. Moreover, for other amyloid peptides, it has been previously shown that hydrophobic regions enhance intermolecular cohesion during amyloid formation [47,48]. The higher hydrophobicity of SEM1(45–107) provides the propensity of peptides to form attractive interactions, leading to a fast aggregation rate.

We carried out the CD analysis of SEM1(45–107) fibrils. The recorded CD spectrum (Figure 12) is typical for the β -sheet of amyloids [49]. We proposed that the so-called discordant helix of SEM1(45–107) in the N-terminus may convert to a β -sheet during fibril formation [50]. A more detailed and complete study of this phenomenon is expected in further studies.

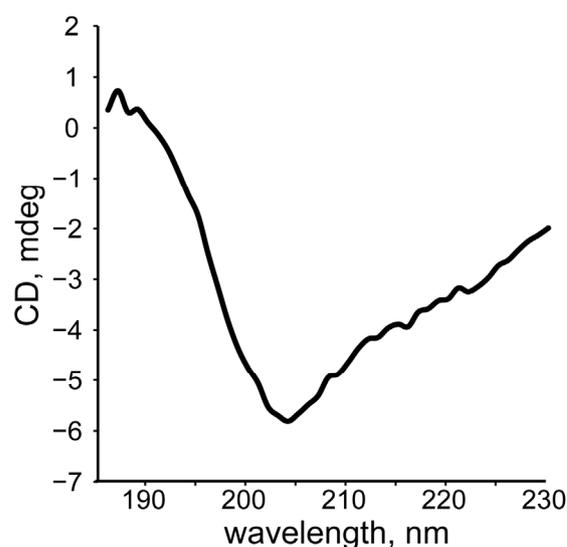


Figure 12. CD spectrum of SEM1(45–107) fibrils.

4. Materials and Methods

4.1. Protein Expression and Purification

The expression vectors were obtained by cloning the SEM1(49–107) and SEM1(45–107) peptide fragments of the *H. sapiens* Semenogelin 1 gene fused with 6xHistidine-tagged GB1 partner protein into pET28a plasmid [51]. Histidine-tagged partner protein was linked with semenogelin fragments via a TEV-protease cleavage site for further separation. The expression and purification protocols for the SEM1(49–107) and SEM1(45–107) peptide fragments (SEM1 fragments) are the same and based on the protocol described before with minor modifications [52]. Protein expression was carried out in an *E. coli* BL21 (DE3) pLysS strain (Novagen, Darmstadt, Germany). Cells were grown in LB-rich nutrient medium, supplemented with 50 $\mu\text{g}/\text{mL}$ kanamycin and 25 $\mu\text{g}/\text{mL}$ chloramphenicol at 37 °C and

180 rpm shaking until the optical density OD_{600} of 0.6–0.8 was reached. The culture was induced to express the SEM1 fragment via the addition of 1 mM isopropyl 1-thio- β -D-galactopyranoside (IPTG) and allowed to grow for 4 h in the same conditions. Then, cells were harvested via centrifugation (5000 rpm, 15 min, 4 °C), frozen and stored at –20 °C.

The frozen cells were thawed, resuspended in buffer 1 (50 mM Tris-HCl, pH 8.8, 0.3 M NaCl) and supplemented with the protease inhibitor cocktail (Roche, Basel, Switzerland) and phenylmethylsulfonyl fluoride (PMSF), followed by sonication on an HD2070 homogenizer (Bandelin, Berlin, Germany) at 4 °C to remove the DNA viscosity. The cell lysate was clarified via 1 h of centrifugation at $100,000\times g$ at 4 °C (Ti-45 rotor, Optima XPN centrifuge, Beckman Coulter Inc., Brea, CA, USA) and loaded on NiNTA resin (QIAGEN, Hilden, Germany) equilibrated by buffer 1. SEM1 peptide fragments fused with GB1 and 6xHis-tag were eluted with buffer 2 (50 mM Tris-HCl, pH 8.8; 0.3 M NaCl; 300 mM imidazole). Fractions after elution were pooled and concentrated via Amicon Ultra-4 (10 K) spin-concentrators (Merck, Burlington, MA, USA) to the concentration allowed for loading on the gel filtration column.

Gel filtration was performed using an NGC Discover chromatographic system and Enrich SEC75 column (BioRad, Hercules, CA, USA) in buffer 3 (50 mM Tris-HCl pH 8.5, 0.5 M NaCl) with a 1 mL/min flow rate. Peak fractions were pooled, and the fusion protein was digested via homemade his-tagged TEV-protease [53] at a ratio TEV:GB1-SEM1 fragment equal to 1:100 (*w/w*). Overnight digestion was carried out in the presence of DTT (1 mM), PMSF (1 mM) and EDTA (0.5 mM) at 4 °C [54]. Then, the reaction mix was loaded on NiNTA-resin again to trap GB1 and TEV-protease. The concentration of the SEM1 fragment in the flow-through fraction reached ~2 mM using Amicon Ultra-0.5 (3K) spin-concentrators (Merck, Burlington, MA, USA). The purity of the samples on each purification step was evaluated via polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE) in pH 8.3 Tris-glycine buffer [55]. Finally, samples of SEM1(49–107) and SEM1(45–107) peptide fragments with purities of more than 95% were obtained.

The SEM1(45–67) and SEM1(49–67) peptides were synthesized and purified to >95% purity by Almabion (Voronezh, Russia).

4.2. ThT Fluorescence

The fibril formation of SEM1(45–107) and SEM1(49–107) was studied using ThT fluorescence intensity measurements. Thioflavin T (ThT) dye fluorescence is regularly used to quantify *in vitro* amyloid fibril formation. Upon binding to amyloid fibrils, ThT gives a strong fluorescence signal at approximately 482 nm [56]. ThT fluorescence probes were observed at 37 °C on a Thermo Scientific Varioskan LUX multimode microplate reader (Waltham, MA, USA) via FluorEssence (v. 6.1) software in a 96-well microplate. ThT fluorescence assays were prepared by mixing 17 μ L of the peptide solution (2mM peptide in 50 mM Tris-HCl pH 8.5, 0.5 M NaCl), 3 μ L of 500 μ M ThT in 50mM Tris-HCl pH 8.5 and 0.5 M NaCl (total volume of 20 μ L) to maintain the fluorescence signal within the linear range of the instrument. The samples were excited at 440 nm, and the fluorescence emission intensity was collected at 482 nm for 90 s and averaged. Fresh peptide solutions of SEM1(45–107) and SEM1(49–107) were prepared for ThT fluorescence measurements. The time interval between the end of synthesis and the registration of fluorescence spectra was less than 10 min. The fluorescence intensity was corrected for lamp intensity fluctuations by dividing the observing fluorescence signal by the lamp intensity. The concentration of SEM1(45–107) and SEM1(49–107) peptides was controlled using NanoDrop One^C (Thermo Fisher Scientific, Waltham, MA, USA).

4.3. PFG-NMR Spectroscopy

Diffusion experiments were carried out using a 700 MHz NMR spectrometer (AVANCE III-HD, Bruker, Billerica, MA, USA) equipped with a quadruple resonance CryoProbe (¹H, ¹³C, ¹⁵N and ³¹P) with a standard z-gradient (a maximum strength of 55.7 G \times cm⁻¹). Diffusion decays were obtained with the help of the stimulated-echo pulse sequence with

water suppression (STEBPGP1S19) containing two field gradient pulses (g) with duration δ , which are separated by interval Δ [57]. Diffusion decays were fitted by:

$$\frac{A(g)}{A(0)} = \exp\left(-D_s \gamma^2 g^2 \delta^2 (\Delta - \delta/3)\right), \quad (2)$$

where $A(0)$ is the spin echo amplitude without gradient pulses, D_s is the self-diffusion coefficient of molecules and γ is the proton gyromagnetic ratio. For all experiments, the amplitude of the field gradient pulses (g) was varied from 2 to 95% of its maximum under a constant diffusion time ($\Delta = 100$ ms) and gradient pulse duration ($\delta = 3.6$ ms). All experiments were performed at 298 K. Data processing and analysis were carried out with Bruker Topspin (v. 3.6) software.

4.4. NMR Spectroscopy

NMR experiments of SEM1(49–67) and SEM1(45–67) were carried out with a 700 MHz NMR spectrometer (AVANCE III-HD, Bruker, Billerica, MA, USA) with a quadruple resonance CryoProbe ($^1\text{H}/^{19}\text{F}$, ^{13}C , ^{15}N and ^{31}P) and a 500 MHz NMR spectrometer (AVANCE-II, Bruker, Billerica, MA, USA) with a triple resonance probe TXI (^1H , ^{13}C and ^{15}N). For the NMR experiment, peptides were dissolved in water solution ($\text{H}_2\text{O} + \text{D}_2\text{O}/90\% + 10\%$): SEM1(49–67) (0.8 mM, pH 3.5); SEM1(45–67) (0.7 mM, pH 3.0) at 298 K. pH values were chosen to prevent fibril formation [58]. To assign the hydrogen chemical shifts, 2D ^1H - ^1H TOCSY [59] (TOCSY mixing time $d9 = 85$ ms) and 2D ^1H - ^1H NOESY [60] (mixing time $t_m = 350$ ms) spectra were obtained. The 2D ^1H - ^{15}N HSQC [57] and 2D ^1H - ^{13}C HSQC [61], 2D ^1H - ^{13}C HMBC [59] and 2D ^1H - ^{13}C HSQC-TOCSY [62] spectra were obtained to perform nitrogen and carbon chemical shift assignments, respectively. Scalar coupling constants $^3J_{\text{H}\alpha\text{-HN}}$ were determined using 2D JRES [63] spectra. The following parameters remained constant in all experiments: the relaxation delay of 1.5s, the number of increments of 2048×1024 and the WATERGATE method [57] for water suppression (if needed). The parameters of 2D NMR spectra are shown in the Supplementary Materials.

Data processing was carried out using Bruker Topspin (v. 3.6) software. All spectra were analyzed with the help of the CCPNMR (v. 2.5) program [64].

4.5. CD Spectroscopy

CD experiments of SEM1(49–67), SEM1(45–67) and SEM1(45–107) fibrils were performed using a Jasco J-1500 spectrometer (Tokyo, Japan) with a scanning speed of 30 nm/min. For CD experiments, the concentrations of peptides were determined spectrophotometrically using the corresponding extinction coefficients of $E_{214}^{1\%} = 28243 \text{ M}^{-1}\text{cm}^{-1}$, $E_{214}^{1\%} = 42598 \text{ M}^{-1}\text{cm}^{-1}$ and $E_{214}^{1\%} = 113084 \text{ M}^{-1}\text{cm}^{-1}$ for SEM1(49–67), SEM1(45–67) and SEM1(45–107), respectively. The concentration was determined as 4 μM for SEM1(49–67), 5 μM for SEM1(45–67) and 1 μM SEM1(45–107) fibrils, respectively. The peptide extinction coefficient was calculated using the web-service Prot.pi (www.protpi.ch/Calculator/ProteinTool, accessed on 1 December 2022). The CD spectra were recorded in 1 cm path length cells in the 190–240 nm wavelength range at 298 K. The peptide secondary structure was analyzed in the frame of the CDSSTR algorithm (protein set 4 and 7) using DichroWeb server (<http://dichroweb.cryst.bbk.ac.uk>, accessed on 1 December 2022) [65]. For the CD analysis of SEM1(45–107) fibrils, we mixed a mature SEM1(45–107) fibril solution (Tris buffer, 100 μL) with water (1000 μL) to decrease the effect of Tris on the CD spectrum.

4.6. Calculation and Visualization

The spatial structure calculations of SEM1(49–67) and SEM1(45–67) were performed using the XPLOR-NIH (v. 3.6) program [66]. The internuclear distances and $^3J_{\text{H}\alpha\text{-HN}}$ alpha coupling constants were determined from the analysis of NOESY and 2D JRES, respectively. The dihedral angles were obtained from the TALOS+ (v. 3.8) program [67] using a chemical shift of backbone chain nuclei: H, N, H_α , C_α , C. The standard non-linear annealing protocol of the XPLOR-NIH (v. 3.6) software [68] was applied for the

calculation of the spatial structure. NMR restraints (internuclear distances and $^3J_{\text{H}\alpha\text{-HN}}$ alpha coupling constants) were used as input parameters in the XPLOR-NIH calculations. Individual peptide structures were minimized, heated to 1000 K for 6000 steps, cooled in 100 K increments to 50 K, and finally, structures were minimized with 1000 steps of steepest descent followed by 1000 steps of conjugate gradient minimization. Thus, a total of 1000 structures were determined, and they were refined in subsequent calculations using the protein.par force field [68,69]. Furthermore, the water refinement of the calculated structures with additional NMR restraints [70] was performed using XPLOR-NIH. Finally, the 12 (for SEM1(49–67)) and 11 (for SEM1(45–67)) structures with the lowest energy were selected.

The peptide spatial structures were visualized using UCSF Chimera (v. 1.16) [71] and MOLMOL (v. 2K.2) software [72]. Ramachandran's plot was calculated using the MolProbity website (molprobity.biochem.duke.edu, accessed on 15 December 2022).

4.7. DLS Spectroscopy

The diffusion coefficient and hydrodynamic radius in mQ water solutions with concentrations of 0.8 and 1 mg/mL for SEM1(49–67) and SEM1(45–67), respectively, were obtained using DLS Photocor Complex (Moscow, Russia) equipped with a He-Ne laser ($\lambda = 632.8$ nm). Cylindrical glass cuvettes with 1 cm radii were used for the DLS measurements. Each correlation function was collected at 90° for 60 s with 20 repeats. We used the CONTIN algorithm to evaluate the corresponding distribution functions [73]. The modified Stokes–Einstein relation was applied to calculate the hydrodynamic radius of molecule R :

$$R = \frac{k_B T q^2}{6\pi\eta} \tau, \quad (3)$$

where k is the Boltzmann constant, T is the absolute temperature, η is dynamic viscosity of solutions, τ are characteristic times, q is the scattering vector and $q = 4\pi n_0 / \lambda \sin(\theta/2)$ ($\lambda = 632.8$ nm, $n_0 = 1.33$ (water), $\theta = 90^\circ$, $q(90^\circ) = 18.5 \times 10^4 \text{ cm}^{-1}$).

Zeta potential (ζ) was analyzed in aqueous solutions by means of a DLS technique on the Photocor Complex. Experiments were conducted in a 1 cm cell with a carbon electrode for electrokinetic measurements under the same conditions as for the hydrodynamic radius measurements.

4.8. MD Simulation

To generate the structural ensembles of the SEM1(45–107) and SEM1(49–107) peptides, we used an integrative approach that incorporates NMR distance constraints into molecular dynamics simulations using a method from Sinelnikova and Spoel's 2021 article [74]. The all-atom MD simulation of peptides was performed using Gromacs (v. 2022) software [41]. The Charmm36 [75] and TIP3P [76] models were used to simulate protein and water molecules, respectively. The initial conformations of SEM1(45–107) and SEM1(49–107) were prepared using the XPLOR-NIH (v. 3.6) program, where we selected structures of peptides with the lowest energy. These structures were solvated in a rhombic dodecahedron box with an initial volume of 282 nm^3 containing 9287 water molecules for SEM1(45–107) and 255 nm^3 containing 8371 water molecules for SEM1(49–107). The neutralization of systems was performed by adding 3 Cl^- ions. The obtained systems were minimized using the steepest descent algorithm with a target maximum force of $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$. In the next step, equilibration was performed in the canonical NVT (constant Number of particles, Volume, and Temperature) ensemble for 100 ps at 300 K using the Berendsen thermostat [77] and for 100 ps at 300 K and 1 bar in the isothermal–isobaric NPT (constant Number of particles, Pressure, and Temperature) ensemble using a Parrinello–Rahman barostat [78]. Furthermore, simulations were carried out for 100 nanoseconds using the same pressure (1 bar) and temperature (300 K) as for the equilibration process. NMR distance constraints were used as input parameters during the MD simulation [74]. After the MD simulation, structural ensembles of SEM1(45–107) and SEM1(49–107) and their MD-trajectory were

obtained. In the final stage, to characterize the convergence of structural ensembles, we performed a cluster analysis of the MD-trajectories (Figures S7 and S8) via the GROMOS clustering algorithm [79]. GROMOS clustering was performed with 0.5 nm RMSD of the C α -C α atom-pair distance cut-off for two structures to be neighbors.

4.9. Transmission Electron Microscopy

The presence of fibrils was shown with the help of transmission electron microscopy (TEM) using the Hitachi HT7700 Exalens scanning electron microscope (Tokyo, Japan). Solutions (10 μ L) of SEM1(45–107) and SEM1(49–107) peptides in Tris-buffer (10^{-3} M) were placed on a 3 mm formvar/carbon-coated copper grid, and drying was carried out at room temperature. The analysis was carried out at an accelerating voltage of 100 kV in TEM mode.

5. Conclusions

In this work, we studied two amyloidogenic peptides of human Semenogelin 1 (residue 45–107, SEM1(45–107) and residue 49–107, SEM1(49–107)) and their N-domains (SEM1(45–67) and SEM1(49–67), respectively). The amyloid fibril formation of SEM1(45–107) and SEM1(49–107) was studied via ThT-fluorescence. It was found that the amyloid formation of SEM1(45–107) starts immediately after purification, whereas the increased ThT fluorescence intensity of SEM1(49–107) solution was observed 30 min after purification, indicating the amyloid formation. The spatial structures of the N-domains of SEM1(45–107) and SEM1(49–107) (SEM1(45–67) and SEM1(49–67)) manifest disordered peptide structures. However, it was found that SEM1(45–67) contains helical fragments in comparison to SEM1(49–67). Previous studies have suggested that helical fragments may rearrange into amyloid β -strands. Thus, the difference in SEM1(45–107) and SEM1(49–107) amyloid-forming behavior may be explained by the presence of the structured helical part of the SEM1(45–107) N-terminus, which contributes to an increased rate of amyloid formation.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24108949/s1>.

Author Contributions: Conceptualization, D.S.B.; methodology, K.S.U. and V.V.K.; validation, A.R.Y.; investigation, D.A.O., A.G.B., E.A.K., A.R.Y., E.A.S. and T.A.M.; writing—original draft preparation, D.S.B. and A.M.K.; writing—review and editing, D.S.B. and V.V.K.; visualization, A.M.K. and D.A.O.; project administration, D.S.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Russian Science Foundation (project no. 20-73-10034).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The ^1H , ^{15}N and ^{13}C chemical shifts of SEM1(45–67) and SEM1(49–67) have been deposited to BioMagResBank, with BMRB ID 51673 and 51691, respectively. The calculated structures of SEM1(45–67) and SEM1(49–67) have been deposited to the Protein Data Bank, with PDB ID 8BOO and 8BVZ, respectively.

Acknowledgments: The authors gratefully acknowledge the Assigned Spectral-Analytical Center of FRC Kazan Scientific Center of RAS for the opportunity to fulfill the DLS experiments. V.K. also acknowledges support from the Kazan Federal University Strategic Academic Leadership Program (PRIORITY-2030).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Simon, V.; Ho, D.D.; Karim, Q.A. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *Lancet* **2006**, *368*, 489–504. [[CrossRef](#)]
2. Dimitrov, D.S.; Willey, R.L.; Sato, H.; Chang, L.J.; Blumenthal, R.; Martin, M.A. Quantitation of human immunodeficiency virus type 1 infection kinetics. *J. Virol.* **1993**, *67*, 2182–2190. [[CrossRef](#)] [[PubMed](#)]

3. Pomerantz, R.J. HIV: Cross-talk and viral reservoirs. *Nature* **2003**, *424*, 136–137. [[CrossRef](#)] [[PubMed](#)]
4. Lilja, H.; Abrahamsson, P.A.; Lundwall, A. Semenogelin, the predominant protein in human semen. Primary structure and identification of closely related proteins in the male accessory sex glands and on the spermatozoa. *J. Biol. Chem.* **1989**, *264*, 1894–1900. [[CrossRef](#)]
5. Malm, J.; Hellman, J.; Magnusson, H.; Laurell, C.B.; Lilja, H. Isolation and characterization of the major gel proteins in human semen, semenogelin I and semenogelin II. *Eur. J. Biochem.* **1996**, *238*, 48–53. [[CrossRef](#)]
6. de Lamirande, E. Semenogelin, the main protein of the human semen coagulum, regulates sperm function. *Semin. Thromb. Hemost.* **2007**, *33*, 60–68. [[CrossRef](#)]
7. Robert, M.; Gagnon, C. Semenogelin I: A coagulum forming, multifunctional seminal vesicle protein. *Cell. Mol. Life Sci.* **1999**, *55*, 944–960. [[CrossRef](#)] [[PubMed](#)]
8. Robert, M.; Gagnon, C. Purification and characterization of the active precursor of a human sperm motility inhibitor secreted by the seminal vesicles: Identity with semenogelin. *Biol. Reprod.* **1996**, *55*, 813–821. [[CrossRef](#)]
9. Vickram, A.S.; Rohini, K.; Anbarasu, K.; Dey, N.; Jeyanthi, P.; Thanigaivel, S.; Issac, P.K.; Arockiaraj, J. Semenogelin, a coagulum macromolecule monitoring factor involved in the first step of fertilization: A prospective review. *Int. J. Biol. Macromol.* **2022**, *209*, 951–962. [[CrossRef](#)]
10. Lilja, H. A kallikrein-like serine protease in prostatic fluid cleaves the predominant seminal vesicle protein. *J. Clin. Investig.* **1985**, *76*, 1899–1903. [[CrossRef](#)]
11. Kim, K.A.; Yolamanova, M.; Zirafi, O.; Roan, N.R.; Staendker, L.; Forssmann, W.G.; Burgener, A.; Dejuq-Rainsford, N.; Hahn, B.H.; Shaw, G.M.; et al. Semen-mediated enhancement of HIV infection is donor-dependent and correlates with the levels of SEVI. *Retrovirology* **2010**, *7*, 55. [[CrossRef](#)] [[PubMed](#)]
12. Roan, N.R.; Müller, J.A.; Liu, H.; Chu, S.; Arnold, F.; Stürzel, C.M.; Walther, P.; Dong, M.; Witkowska, H.E.; Kirchhoff, F.; et al. Peptides released by physiological cleavage of semen coagulum proteins form amyloids that enhance HIV infection. *Cell Host Microbe* **2011**, *10*, 541–550. [[CrossRef](#)] [[PubMed](#)]
13. Roan, N.R.; Liu, H.; Usmani, S.M.; Neidleman, J.; Müller, J.A.; Avila-Herrera, A.; Gawanbacht, A.; Zirafi, O.; Chu, S.; Dong, M.; et al. Liquefaction of Semen Generates and Later Degrades a Conserved Semenogelin Peptide That Enhances HIV Infection. *J. Virol.* **2014**, *88*, 7221–7234. [[CrossRef](#)] [[PubMed](#)]
14. Roan, N.R.; Munch, J.; Arhel, N.; Mothes, W.; Neidleman, J.; Kobayashi, A.; Smith-McCune, K.; Kirchhoff, F.; Greene, W.C. The cationic properties of SEVI underlie its ability to enhance human immunodeficiency virus infection. *J. Virol.* **2009**, *83*, 73–80. [[CrossRef](#)]
15. Roan, N.R.; Sowinski, S.; Munch, J.; Kirchhoff, F.; Greene, W.C. Aminoquinoline surfen inhibits the action of SEVI (semen-derived enhancer of viral infection). *J. Biol. Chem.* **2010**, *285*, 1861–1869. [[CrossRef](#)]
16. Roan, N.R.; Sandi-Monroy, N.; Kohgadai, N.; Usmani, S.M.; Hamil, K.G.; Neidleman, J.; Montano, M.; Ständker, L.; Röcker, A.; Cavrois, M.; et al. Semen amyloids participate in spermatozoa selection and clearance. *eLife* **2017**, *6*, e24888. [[CrossRef](#)]
17. Castellano, L.M.; Shorter, J. The Surprising Role of Amyloid Fibrils in HIV Infection. *Biology* **2012**, *1*, 58–80. [[CrossRef](#)]
18. Jones, R. Interaction of zona pellucida glycoproteins, sulphated carbohydrates and synthetic polymers with proacrosin, the putative egg-binding protein from mammalian spermatozoa. *Development* **1991**, *111*, 1155–1163. [[CrossRef](#)]
19. Cohen, F.S.; Melikyan, G.B. The Energetics of Membrane Fusion from Binding, through Hemifusion, Pore Formation, and Pore Enlargement. *J. Membr. Biol.* **2004**, *199*, 1–14. [[CrossRef](#)]
20. Evans, J.P.; Florman, H.M. The state of the union: The cell biology of fertilization. *Nat. Cell Biol.* **2002**, *4*, s57–s63. [[CrossRef](#)]
21. Sanchugova, D.; Kusova, A.; Bikmullin, A.; Yulmetov, A.; Mukhamezzyanov, T.; Klochkov, V.; Blokhin, D. Conformational ensemble of amyloid-forming semenogelin 1 peptide SEM1(68–107) by NMR spectroscopy and MD simulations. *J. Struct. Biol.* **2022**, *214*, 107900. [[CrossRef](#)] [[PubMed](#)]
22. Blokhin, D.S.; Fayzullina, A.R.; Filippov, A.V.; Karataeva, F.K.; Klochkov, V.V. Spatial structure of fibrinopeptide B in water solution with DPC micelles by NMR spectroscopy. *J. Mol. Struct.* **2015**, *1102*, 91–94. [[CrossRef](#)]
23. Kononova, O.; Litvinov, R.; Blokhin, D.; Klochkov, V.; Weisel, J.; Bennett, J.; Barsegov, V. Mechanistic Basis for the Binding of RGD- and AGDV-Peptides to the Platelet Integrin α IIb β 3. *Biochemistry* **2017**, *56*, 1932–1942. [[CrossRef](#)] [[PubMed](#)]
24. Kazimierczuk, K.; Kasprzak, P.; Georgoulia, P.S.; Matečko-Burmann, I.; Burmann, B.M.; Isaksson, L.; Gustavsson, E.; Westenhoff, S.; Orekhov, V.Y. Resolution enhancement in NMR spectra by deconvolution with compressed sensing reconstruction. *Chem. Commun.* **2020**, *56*, 14585–14588. [[CrossRef](#)] [[PubMed](#)]
25. Lawrence, C.; Grishaev, A. Chemical shifts-based similarity restraints improve accuracy of RNA structures determined via NMR. *RNA* **2020**, *26*, 2051–2061. [[CrossRef](#)]
26. Puthenveetil, R.; Vinogradova, O. Solution NMR: A powerful tool for structural and functional studies of membrane proteins in reconstituted environments. *J. Biol. Chem.* **2019**, *294*, 15914–15931. [[CrossRef](#)]
27. Yao, S.; Babon, J.J.; Norton, R.S. Protein effective rotational correlation times from translational self-diffusion coefficients measured by PFG-NMR. *Biophys. Chem.* **2008**, *136*, 145–151. [[CrossRef](#)]
28. Lucas, L.H.; Larive, C.K. Measuring ligand-protein binding using NMR diffusion experiments. *Concepts Magn. Reson.* **2004**, *20A*, 24–41. [[CrossRef](#)]
29. Kusova, A.M.; Sitnitsky, A.E.; Uversky, V.N.; Zuev, Y.F. Effect of Protein–Protein Interactions on Translational Diffusion of Spheroidal Proteins. *Int. J. Mol. Sci.* **2022**, *23*, 9240. [[CrossRef](#)]

30. Kusova, A.M.; Sitnitsky, A.E.; Zuev, Y.F. The Role of pH and Ionic Strength in the Attraction–Repulsion Balance of Fibrinogen Interactions. *Langmuir* **2021**, *37*, 10394–10401. [[CrossRef](#)]
31. Lorber, B.; Fischer, F.; Bailly, M.; Roy, H.; Kern, D. Protein analysis by dynamic light scattering: Methods and techniques for students. *Biochem. Mol. Biol. Educ.* **2012**, *40*, 372–382. [[CrossRef](#)] [[PubMed](#)]
32. Wang, W.; Lv, H.; Zhao, Y.; Liu, D.; Wang, Y.; Zhang, Y. DLS: A Link Prediction Method Based on Network Local Structure for Predicting Drug-Protein Interactions. *Front. Bioeng. Biotechnol.* **2020**, *8*, 330. [[CrossRef](#)] [[PubMed](#)]
33. Whitmore, L.; Wallace, B.A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* **2008**, *89*, 392–400. [[CrossRef](#)]
34. Liu, G.; Prabhakar, A.; Aucoin, D.; Simon, M.; Sparks, S.; Robbins, K.J.; Sheen, A.; Petty, S.A.; Lazo, N.D. Mechanistic studies of peptide self-assembly: Transient α -helices to stable β -sheets. *J. Am. Chem. Soc.* **2010**, *132*, 18223–18232. [[CrossRef](#)] [[PubMed](#)]
35. Karplus, M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *J. Phys. Chem.* **1959**, *30*, 11–15. [[CrossRef](#)]
36. Lovell, S.C.; Davis, I.W.; Arendall III, W.B.; de Bakker, P.I.W.; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Calpha geometry: Phi, psi and Cbeta deviation. *Proteins* **2003**, *50*, 437–450. [[CrossRef](#)] [[PubMed](#)]
37. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [[CrossRef](#)]
38. Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley-Interscience: New York, NY, USA, 1986.
39. Leite, J.P.; Gimeno, A.; Taboada, P.; Jiménez-Barbero, J.J.; Gales, L. Dissection of the key steps of amyloid- β peptide 1-40 fibrillogenesis. *Int. J. Biol. Macromol.* **2020**, *164*, 2240–2246. [[CrossRef](#)]
40. Kusova, A.M.; Iskhakova, A.K.; Zuev, Y.F. NMR and dynamic light scattering give different diffusion information for short-living protein oligomers. Human serum albumin in water solutions of metal ions. *Eur. Biophys. J.* **2022**, *51*, 375–383. [[CrossRef](#)]
41. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; van der Spoel, D.; et al. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *J. Bioinform.* **2013**, *29*, 845–854. [[CrossRef](#)]
42. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. [[CrossRef](#)] [[PubMed](#)]
43. Blokhin, D.S.; Filippov, A.V.; Antzutkin, O.N.; Karataeva, F.K.; Klochkov, V.V. Spatial structure of oligopeptide PAP (248–261), the N-terminal fragment of the HIV enhancer prostatic acid phosphatase peptide PAP (248–286), in aqueous and SDS micelle solutions. *J. Mol. Struct.* **2014**, *1070*, 38–42. [[CrossRef](#)]
44. Blokhin, D.S.; Filippov, A.V.; Antzutkin, O.N.; Afonin, S.; Klochkov, V.V. Spatial Structures of PAP (262–270) and PAP (274–284), Two Selected Fragments of PAP (248–286), an Enhancer of HIV Infectivity. *Appl. Magn. Reson.* **2015**, *46*, 757–769. [[CrossRef](#)]
45. Nanga, R.P.R.; Brender, J.R.; Vivekanandan, S.; Popovych, N.; Ramamoorthy, A. NMR Structure in a Membrane Environment Reveals Putative Amyloidogenic Regions of the SEVI Precursor Peptide PAP248-286. *J. Am. Chem. Soc.* **2009**, *131*, 17972–17979. [[CrossRef](#)]
46. Sanchugova, D.; Kusova, A.; Bikmullin, A.; Klochkov, V.; Blokhin, D. The Structure of Fibril-Forming SEM1 (86–107) Peptide Increasing the HIV Infectivity. *BioNanoScience* **2021**, *11*, 182–188. [[CrossRef](#)]
47. Ziehm, T.; Buell, A.K.; Willbold, D. Role of Hydrophobicity and Charge of Amyloid-Beta Oligomer Eliminating d-Peptides in the Interaction with Amyloid-Beta Monomers. *ACS Chem. Neurosci.* **2018**, *9*, 2679–2688. [[CrossRef](#)]
48. Itoh, S.G.; Yagi-Utsumi, M.; Kato, K.; Okumura, H. Effects of a Hydrophilic/Hydrophobic Interface on Amyloid- β Peptides Studied by Molecular Dynamics Simulations and NMR Experiments. *J. Phys. Chem. B* **2019**, *123*, 160–169. [[CrossRef](#)]
49. Ye, Z.; French, K.C.; Popova, L.A.; Lednev, I.K.; Lopez, M.M.; Makhatadze, G.I. Mechanism of fibril formation by a 39-residue peptide (PAPf39) from human prostatic acidic phosphatase. *Biochemistry* **2009**, *48*, 11582–11591. [[CrossRef](#)]
50. Kallberg, Y.; Gustafsson, M.; Persson, B.; Thyberg, J.; Johansson, J. Prediction of amyloid fibril-forming proteins. *J. Biol. Chem.* **2001**, *276*, 12945–12950. [[CrossRef](#)]
51. Bogomolovas, J.; Simon, B.; Sattler, M.; Stier, G. Screening of fusion partners for high yield expression and purification of bioactive viscotoxins. *Protein Expr. Purif.* **2009**, *64*, 16–23. [[CrossRef](#)]
52. Bikmullin, A.; Klochkova, E.; Krasnovid, F.; Blokhin, D. The data of heterologous expression protocol for synthesis of ^{15}N , ^{13}C -labeled SEM1 (68–107) peptide fragment of homo sapiens semenogelin 1. *MethodsX* **2021**, *8*, 101512. [[CrossRef](#)] [[PubMed](#)]
53. Tropea, J.E.; Cherry, S.; Waugh, D.S. Expression and purification of soluble His(6)-tagged TEV protease. *High Throughput Protein Expr. Purif.* **2009**, *498*, 297–307. [[CrossRef](#)]
54. Raran-Kurussi, S.; Cherry, S.; Zhang, D.; Waugh, D.S. Removal of affinity tags with TEV protease. *Methods Mol. Biol.* **2017**, *1586*, 221–230. [[CrossRef](#)] [[PubMed](#)]
55. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–685. [[CrossRef](#)]
56. Abelein, A.; Jarvet, J.; Barth, A.; Gräslund, A.; Danielsson, J. Ionic strength modulation of the free energy landscape of A β 40 peptide fibril formation. *J. Am. Chem. Soc.* **2016**, *138*, 6893–6902. [[CrossRef](#)]
57. Sklenar, V.; Piotto, M.; Leppik, R.; Saudek, V. Gradient-Tailored Water Suppression for ^1H - ^{15}N HSQC Experiments Optimized to Retain Full Sensitivity. *J. Magn. Reson. Ser. A.* **1993**, *102*, 241–245. [[CrossRef](#)]
58. Avni, A.; Swasthi, H.M.; Majumdar, A.; Mukhopadhyay, S. Intrinsically disordered proteins in the formation of functional amyloids from bacteria to humans. *Prog. Mol. Biol. Transl. Sci.* **2019**, *166*, 109–143. [[CrossRef](#)]

59. Berger, S.; Braun, S. *200 and More NMR Experiments*; Wiley-VCH: Weinheim, Germany, 2004.
60. Rule, G.S.; Hitchens, T.K. *Fundamentals of Protein NMR Spectroscopy*; Springer: Dordrecht, The Netherlands, 2006.
61. Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S.J.; Sørensen, O.W.; Griesinger, C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* **1994**, *4*, 301–306. [[CrossRef](#)]
62. Nolis, P.; Parella, T. Spin-edited 2D HSQC-TOCSY experiments for the measurement of homonuclear and heteronuclear coupling constants: Application to carbohydrates and peptides. *J. Magn. Reson.* **2005**, *176*, 15–26. [[CrossRef](#)]
63. Huang, Y.; Zhang, Z.; Chen, H.; Feng, J.; Cai, S.; Chen, Z. A high-resolution 2D J-resolved NMR detection technique for metabolite analyses of biological samples. *Sci. Rep.* **2015**, *5*, 8390. [[CrossRef](#)]
64. Vranken, W.F.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, M.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687–696. [[CrossRef](#)] [[PubMed](#)]
65. Sreerama, N.; Woody, R.W. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **2000**, *287*, 252–260. [[CrossRef](#)] [[PubMed](#)]
66. Schwieters, C.D.; Kuszewski, J.J.; Tjandra, N.; Clore, G.M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **2003**, *160*, 65–73. [[CrossRef](#)] [[PubMed](#)]
67. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **2009**, *44*, 213–223. [[CrossRef](#)] [[PubMed](#)]
68. Schwieters, C.D.; Kuszewski, J.J.; Clore, G.M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 47–62. [[CrossRef](#)]
69. Engh, R.A.; Huber, R. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. A* **1991**, *47*, 392–400. [[CrossRef](#)]
70. Nederveen, A.J.; Doreleijers, J.F.; Vranken, W.; Miller, Z.; Spronk, C.A.; Nabuurs, S.B.; Güntert, P.; Livny, M.; Markley, J.L.; Nilges, M.; et al. RECOORD: A REcalculated COORDinates Database of 500+ proteins from the PDB using restraints from the BioMagResBank. *Proteins* **2005**, *59*, 662–672. [[CrossRef](#)]
71. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]
72. Koradi, R.; Billeter, M.; Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **1996**, *14*, 51–55. [[CrossRef](#)]
73. Scotti, A.E.; Liu, W.; Hyatt, J.S.; Herman, E.S.; Choi, H.S.; Kim, J.W.; Lyon, L.A.; Gasser, U.; Fernandez-Nieves, A. The CONTIN algorithm and its application to determine the size distribution of microgel suspensions. *J. Chem. Phys.* **2015**, *142*, 234905. [[CrossRef](#)]
74. Sinelnikova, A.; Spoel, D.V.D. NMR refinement and peptide folding using the GROMACS software. *J. Biomol. NMR* **2021**, *75*, 143–149. [[CrossRef](#)] [[PubMed](#)]
75. Huang, J.; MacKerell, A.D., Jr. CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145. [[CrossRef](#)] [[PubMed](#)]
76. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
77. Berendsen, H.J.; Postma, J.V.; van Gunsteren, W.F.; DiNola, A.; Haak, J. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690. [[CrossRef](#)]
78. Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190. [[CrossRef](#)]
79. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Gunsteren, W.F.V.; Mark, A.E. Peptide folding: When simulation meets experiment. *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 236–240. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.