Non-Covalent Complexes of the Peptide Fragment Gly-Asn-Asn-Gln-Gln-Asn-Tyr in the Gas-Phase. Photodissociative Cross-Linking, Born-Oppenheimer Molecular Dynamics, and Ab Initio Computational Binding Study

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-11-2018-006893.R1
Article Type:	Paper
Date Submitted by the Author:	12-Dec-2018
Complete List of Authors:	Huang, Shu; University of Washington, Department of Chemistry Liu, Yang; University of Washington, Department of Chemistry Turecek, Frantisek; University of Washington, Department of Chemistry

Journal Name

ARTICLE

Non-Covalent Complexes of the Peptide Fragment Gly-Asn-Asn-Gln-Gln-Asn-Tyr in the Gas-Phase. Photodissociative Cross-Linking, Born-Oppenheimer Molecular Dynamics, and Ab Initio Computational Binding Study

Shu R. Huang, Yang Liu, František Tureček*

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Non-covalent complexes of the short amyloid peptide motif Gly-Asn-Asn-Gln-Gln-Asn-Tyr (GNNQQNY) with peptide counterparts that were tagged with a diazirine ring at the N-termini (*GNNQQNY) were generated as singly charged ions in the gas phase. Specific laser photodissociation (UVPD) of the diazirine tag in the gas-phase complexes at 355 nm generated transient carbene intermediates that underwent covalent cross-linking with the target GNNQQNY peptide. The crosslinking yields ranged between 0.8 and 4.5%, depending on the combinations of peptide C-terminal amides and carboxylates. The covalent complexes were analyzed by collision-induced dissociation tandem mass spectrometry (CID-MS^3), providing distributions of cross-links at the target peptide amino acid residues. A general preference for cross-linking at the target peptide Gln-4-Gln-5-Asn-6-Tyr-7 segment was observed. Born-Oppenheimer molecular dynamics calculations were used to obtain 100 ps trajectories for nine lowest free-energy conformers identified by ωB97X-D/6-31+G(d,p) gradient geometry optimizations. The trajectories were analyzed for close contacts between the incipient carbene atom and the X–H bonds in the target peptide. The close-contact analysis pointed to the Gln-5 and Tyr-7 residues as the most likely sites of cross-linking, consistent with the experimental CID-MS^3 results. Non-covalent binding in the amide complexes was evaluated by DFT calculations of structures and energies. Although antiparallel arrangements of the GNNQQNY and *GNNQQNY peptides were favored in low-energy gas-phase and solvated complexes, the conformations and peptide-peptide interface surfaces were found to differ from the secondary structure of the dry interface in GNNQQNY motifs of amyloid aggregates.

1. Introduction

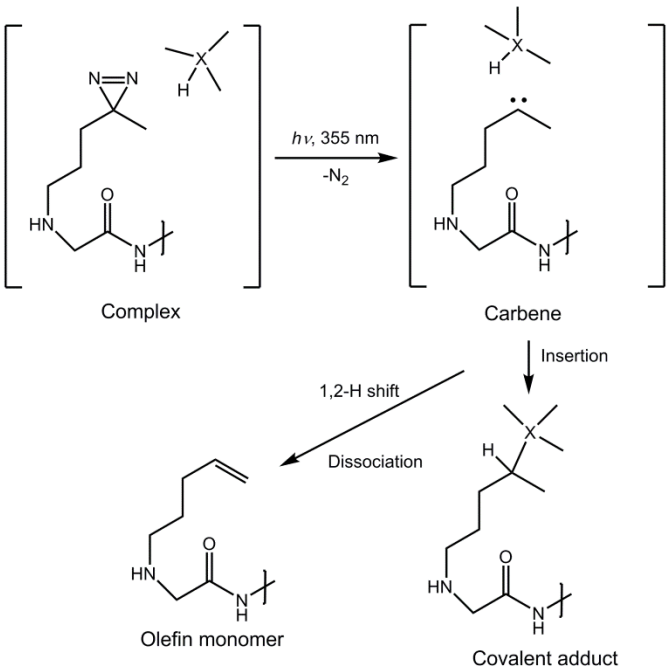
In vivo protein conformational changes and aggregation involving prions[1,2] and other amyloidogenic proteins[3] have been linked to several diseases such as type II diabetes, Alzheimer's and Parkinson's diseases. Numerous in vitro studies addressed the sequence motifs have been identified in amyloid proteins. One of these motifs, Gly-Asn-Asn-Gln-Gln-Asn-Tyr (GNNQQNY), has been found to commonly occur in the cross-β spine that is involved in the formation of amyloid fibrils.[3] In addition to standard biochemical studies in solution, there have been several reports using mass spectrometry to unravel the structure and dynamics of protein aggregation with smaller peptide motifs.[4] For example, peptide complexes involving the VEALYL and SSTNVG sequences were studied by ion mobility mass spectrometry (IM-MS) in combination

[a] Department of Chemistry, Bagley Hall, Box 351700, University of Washington, Seattle, WA, 98195-1799, USA
[b]

† Electronic Supplementary Information (ESI) available: Tables S1-S9, Figures S1-S10. See DOI: 10.1039/x0xx00000x

with force-field molecular dynamics calculations to bear on gas-phase ion structures.[5] Other peptide sequences related to amyloid proteins that have been studied by IM-MS comprised NFGAIL,[6] mutants of NNQQNY,[7] KCNTATCA fragment of human islet amyloid polypeptide,[8,9] and amyloid-β fragments.[10-12] Whereas IM-MS is a powerful technique for monitoring peptide aggregation in solution,[4] the structural information it provides on the complexes is limited to projections used to estimate collision cross sections that can often be realized by several structures. In addition, force-field molecular dynamics used to follow conformation trajectories has not been calibrated to provide thermodynamic properties of gas-phase ions. In addition to IM-MS, H/D exchange has been used to study aggregation of precursors of bacterial curli amyloid protein fibers.[13] In another approach to gas-phase ion structures, cold ion spectroscopy has been used to track structure changes in multiply charge peptide clusters.[14]

We have shown previously that even weakly binding peptide-peptide complexes that are difficult to study in solution at micromolar concentrations can be produced by electrospray ionization as singly or multiply charged ions in the gas phase and isolated in an ion trap mass spectrometer.[15-18] This provides a non-traditional method for generating a variety of non-covalent biomolecular complexes while utilizing tandem mass spectrometric techniques to determine complex structure and binding.[15] Our strategy for investigating non-covalent binding in gas-phase complexes relies on the specific tagging of one peptide component with a photocleavable diazirine group.[19] Consistent with the low chemical reactivity of diazirine derivatives in solution,[20] diazirine-tagged peptides are stable under the conditions of increased acidity in electrospray droplets,[21] so that tagged peptide ions can be efficiently produced and stored in the gas phase.[22] The ground-state chemistry,[22] redox properties,[23,24] and photochemistry[25] of diazirine-tagged peptide ions have been studied in detail and are well understood. Selective photodissociation of the diazirine ring at 355 nm results in N_2 expulsion, forming a highly reactive singlet carbene intermediate (Scheme 1). In the presence of a proximate functional group in the complex, the carbene can undergo fast insertion into an X-H bond forming new C-X and C-H bonds and converting the non-covalent complex into a covalently bound ion. This photodissociative cross-linking competes with exothermic carbene isomerization by 1,2-hydrogen migrations from the adjacent alkyl groups, generating unreactive olefins. According to fast kinetic studies[26-29] and computational analysis,[22] singlet carbenes generated form diazirines have submicrosecond half-lives. This provides a kinetic constraint for the cross-linking reaction that has to occur on the same time scale to be competitive. We utilize this kinetic constraint for Born-Oppenheimer molecular dynamics trajectory calculations of thermal conformational motion in non-covalent gas-phase complexes that occurs on a comparable (10^{-10}-10^{-9} s) time scale. Conversely, in the absence of cross-linking the carbene-olefin isomerization has been calculated to be >200 kJ mol^{-1} exothermic.[22] The generated vibrational energy, when combined with the thermal internal energy of the ion, can drive the dissociation of the non-covalent complex, marking the absence of cross-linking. We utilize these thermodynamic properties of photochemically generated carbenes to assess the binding in non-covalent complexes. We now apply the experimental and computational methods to study binding in non-covalent dimer complexes of peptide GNNQQNY in the carboxyl and amide forms. In contrast to previous IM-MS studies that targeted multiply charged peptide oligomers under low resolution conditions, our goal is to identify peptide-peptide interactions in gas-phase complexes at an atomic-resolution level and quantify the interaction energies at high levels of quantum theory.



Scheme 1. Photodissociative crosslinking of diazirine-tagged peptides

2. Experimental Section

2.1. Materials

Heptapeptide GNNQQNY was synthesized on bare Wang and H-PAL ChemMatrix resin (Sigma-Aldrich) using a LibertyBlue peptide synthesizer, resulting in the carboxyl, $GNNQQNY_{OH}$, and amide, $GNNQQNY_{NH2}$, forms respectively. Standardized Fmoc peptide cleaving procedures were followed.[15] For diazirine derivatization in GNNQQNY,[16,17] the *N*-terminus of the resin-tethered peptide was selectively deprotected and the sample was incubated with 4,4-azi-1-iodopentane in *N,N*-dimethylformamide for 24 h at room temperature. Standard Fmoc peptide cleaving procedures were then used to release the tagged peptide. The *N*-terminal derivatization process was applied to both Wang and H-PAL ChemMatrix resin, resulting in $*GNNQQNY_{OH}$ and $*GNNQQNY_{NH2}$, respectively. The asterisk before the heptapeptide sequence indicates *N*-terminal 4,4-azipentyl group. Custom peptide sequences of >90% purity, $PAGGYYQNY_{NH2}$ and $PQGGYQQYN_{NH2}$, were purchased from GenScript (Piscataway, NJ, USA).

2.2. Methods

Mass spectra were acquired on a ThermoFisher LTQ XL ETD linear ion trap mass spectrometer (ThermoFisher, San Jose, CA) that was coupled to an EKSPLA NL 301 HT Nd:YAG laser (Altos, Bozeman, MT) operating at 20 Hz with a third harmonic frequency generator to produce the 355 nm beam. The typical parameters and experimental set-up for coupling the laser to the mass spectrometer have been reported previously.[30] Peptide dimer complexes were electrosprayed at 250 to 300 μM concentrations in 50:50:1 methanol:water:acetic acid solution using a home-built microelectrospray ion source with a 363 ± 10 μm o.d. fused silica capillary, where the electrospray tip was pulled to a 5 μm o.d. and positioned 2 mm from the LTQ sampling cone. A 2.2 kV voltage was

applied to the capillary and the flow rate was 1 µL/min. Care was taken to avoid sample carryover by carefully washing the ESI source with solvent and running the MS^n experiments between complexes ***I–IV*** at least one-week apart.

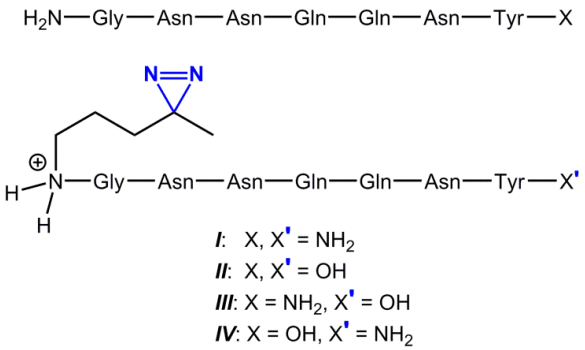
2.3. Calculations

Initial ion structures of the (*GNNQQNY$_{NH2}$ + GNNQQNY$_{NH2}$ + H)$^+$ complex were built in GaussView 6.0 software and subjected to Born-Oppenheimer molecular dynamics (BOMD) trajectories that were run for 20 ps using the Berendsen thermostat method.[31] The calculations were run with 1-fs steps, using the all-valence-electron semiempirical PM6 method[32] that was augmented by including dispersion interactions (PM6-D3H4)[33] that are suitable for handling non-covalent complexes. These calculations were run with MOPAC 16[34] coupled to the Cuby4 platform.[35] The bath temperature in these initial trajectory calculations was set at 400 K to generate a diversity of structures for conformational sampling. Two hundred snapshot structures were extracted from the 20 ps BOMD trajectory, reoptimized with PM6-D3H4 and ranked by energy. Nine of the lowest energy PM6-D3H4 reoptimized structures, within a 25 kJ mol^{-1} of the global energy minimum were subjected to 100 ps trajectory calculations at 310 K, corresponding to the temperature in the ion trap. To obtain thermodynamic data, the PM6-D3H4 geometries were reoptimized with density functional theory (DFT) calculations using the Gaussian 16 (revision A.03) suite of programs.[36] Harmonic frequencies from B3LYP/6-31G(d,p) calculations[37,38] were used to calculate 310 K enthalpies and entropies of fully optimized complexes. In separate runs, structures were fully optimized with ωB97XD/6-31+G(d,p) to capture dispersion interactions in the complexes.[39] The ωB97XD/6-31+G(d,p) energies were used to provide the electronic terms for isomer energy ranking and calculations of binding energies. Solvent effects on the complexes' structure and energetics were investigated by self-consistent reaction field ωB97XD/6-31+G(d,p) calculations using the polarizable continuum model[40] with standard parameters from Gaussian 16.[36]

3. Results and Discussion

3.1. Photodissociation of (GNNQQNY + *GNNQQNY + H)$^+$

Non-covalent peptide-photopeptide ion complexes were generated in the gas phase by electrospray ionization of peptide mixtures containing equimolar concentrations of an amyloid-motif peptide and its photoactive counterpart. The amyloid-motif sequence,[3] GNNQQNY$_X$, was realized with a free carboxyl (X = OH) or an amide group (X = NH_2). The photoactive counterparts, *GNNQQNY$_X$', had the N-terminal Gly amine tagged with the diazirine-containing 4,4-azipentyl group (Scheme 2).



Scheme 2. Peptide components of the complexes

This gave rise to four combinations of complexes ***I-IV*** with (X, X') = (NH_2, NH_2), (OH, OH), (NH_2, OH), and (OH, NH_2), respectively, that were generated in the gas phase and subjected to photodissociative crosslinking (Scheme 2). Electrospray ionization produced singly charged gas-phase complexes in relatively low yields when compared to the ions of the monomeric components. For example, the singly charged non-covalent dimer ion (GNNQQNY$_{NH2}$ + *GNNQQNY$_{NH2}$ + H)$^+$, complex ***I*** at *m/z* 1767, was formed at ca. 0.2% relative to the monomer ion intensities (Figure 1a). The other (X, X') combinations gave similar yields of dimeric ions. The gas-phase complexes are denoted by (mM+H)$^+$, where m stands for the target peptide and M is the N-terminally photo-labeled peptide. No doubly charged complexes were detected, in contrast to previous studies of larger aggregates where multiply charged ions were produced.[5,11] Despite the modest electrospray yields, complexes ***I-IV*** were isolated by their *m/z* ratios in the ion trap and subjected to further investigation. Collision-induced dissociation of the complexes resulted in dissociation to monomers, with predominant charge retention on the photopeptides (Supplementary Figures S1-S4). Photodissociation of mass-selected complexes ***I-IV*** was performed at 355 nm to selectively target the diazirine chromophore. Photodissociation was carried out at a high conversion, as illustrated by the tandem mass spectrum (UVPD-MS^2) of ***I*** that showed major (mM–N_2+H)$^+$ photoproducts and only

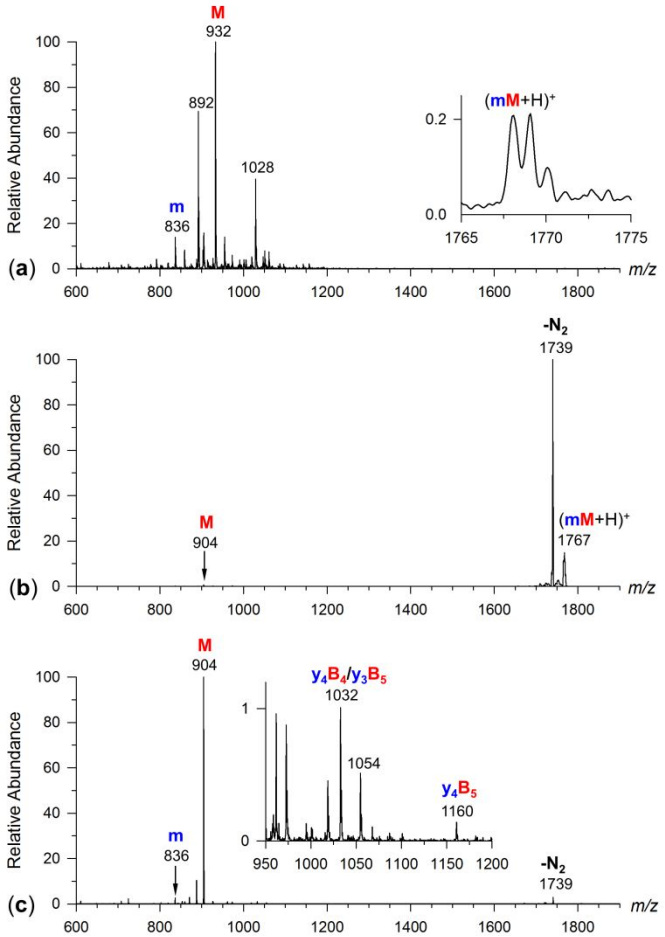


Figure 1. (a) Electrospray mass spectrum of complex ***I***, (GNNQQNY$_{NH2}$+*GNNQQNY$_{NH2}$+H)$^+$, denoted as (mM+H)$^+$, *m/z* 1767. (b) UVPD-MS^2 of complex ***I***. (c) CID-MS^3 of the (mM–N_2+H)$^+$ ion at *m/z* 1739.

weak residual $(mM+H)^+$ precursor ions (Figure 1b). The other complexes, *II*-*IV* gave very similar results. To achieve a high conversion despite the low molar absorptivity of the diazirine chromophore (ε_{max} < 50), the ion population was exposed to 19 laser pulses at 13-14 mJ per pulse. It should be noted that the products of N_2 loss are transparent at 355 nm and do not undergo further photodissociation. Significantly, photodissociative loss of N_2 from the diazirine tag was accompanied by only a minor (<3%) dissociation of gas-phase *I* to monomeric units in ions trapped at 310 K (Table 1). Photodissociation of complexes ***II–IV*** yielded similar results (Table 1, Supplementary Figure S5).

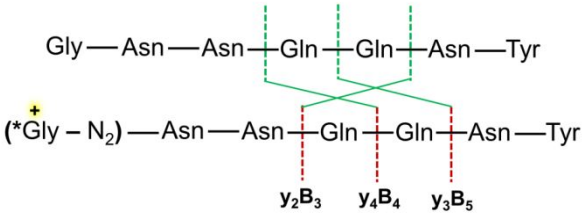
Table 1. Survival and cross-linking efficiencies for GNNQQNY complexes.

	Photopeptide					
	*GNNQQNY$_{NH2}$			*GNNQQNY$_{OH}$		
Target peptide	MS^2	MS^3	Overall	MS^2	MS^3	Overall
GNNQQNY$_{NH2}$	99	2.4	2.4	99	0.8	0.8
GNNQQNY$_{OH}$	99	4.5	4.5	98	4.5	4.4

To further investigate the $(mM–N_2+H)^+$ complexes, the ions were isolated by mass and subjected to collision-induced dissociation, CID-MS^3. The CID-MS^3 spectrum of $(mM–N_2+H)^+$ from ***I*** revealed that a major fraction of the complexes dissociated to monomers, predominantly forming $(M–N_2+H)^+$ ions at *m/z* 905 and neutral m (Figure 1c). The low relative abundance of the complementary (m + H)$^+$ ion at *m/z* 837 reflects the lower basicity of the target peptide.

3.2. CID-MS^3 Spectra Following Photodissociation

The facile dissociation of $(mM–N_2+H)^+$ upon CID indicated that the peptide components in the majority fraction of the complexes were not covalently crosslinked and remained bound by non-covalent interactions. More interestingly, the CID-MS^3 spectrum of the $(mM–N_2+H)^+$ complex ***I*** revealed a small fraction of photo-crosslinked complexes that were identified as backbone fragment ions containing amino acid residues from both M and m. These photo-crosslinked ions corresponded to fragments formed by loss of ammonia and *b*/*y*-type cleavage of the peptide sequences (Figure 1c). The backbone fragment ions were denoted according to a previously introduced nomenclature.[15] Briefly, the capital letters M, B, and Y referred to the photopeptide and sequence fragments originating therefrom, while the small letters m, b, y referred to the target peptide and its corresponding *b*/*y* fragments. The B, Y, b, y fragments followed the Roepstorff-Fohlman-Biemann nomenclature for the dissociation of peptide ions.[41,42] Referring to the Figure 1c and Table S1 spectrum, the photo-crosslinked ion at *m/z* 648 was assigned as y_2B_3, indicating the retention of the y_2 moiety (Asn-Tyr) from the target peptide and the B_3 moiety ((*Gly–N_2)-Asn-Asn) from the photo-labeled monomer (Scheme 2). Some of the photo-crosslinked fragment ions were mass-degenerate by combinations of Q and N residues originating from from m and M. For example, the *m/z* 961 ion had the same nominal mass for mB_1 and b_1M, and likewise for the *m/z* 1018 ion (y_6B_2, y_5B_3, and y_2B_6) and the *m/z* 1032 ion (y_4B_4 and y_3B_5, Scheme 3). According to Savitski et al.,[43] the preference for a peptide bond cleavage ranked from high to low was N-Q > G-N ≈ Q-Q > N-Y > N-N > Q-N, indicating no strong preference for CID bond cleavage at these residues. Hence it is possible that both y_4B_4 and y_3B_5 contribute equally to the ion intensity of *m/z* 1032. A similar argument can be made for the degeneracy at *m/z* 1018. Backbone dissociations in the target peptide part of the cross-linked complex allowed us to approximate the location of the cross-links. For example, the intense backbone fragment ion assigned to y_4B_4/y_3B_5 (*m/z* 1032) suggested that crosslinking has occurred between the photo-labeled peptide and residues 4 to 7 (QQNY) of the target peptide. The backbone fragment ion intensities assignable to identified cross-links (Supplementary Table S1) were normalized[15] and are given in Table 2. Note that cross links to Tyr-7 are specifically identified by the y_1X fragment ions.



Scheme 3. Fragment ion assignment and nomenclature.

Table 2. Normalized distribution of cross-links in complex ***I***.

Target peptide residue	Normalized cross-link fraction (%)
Gly-1	5.5
Asn-2	3
Asn-3	6
Gln-4	13.5
Gln-5	20.5
Asn-6	25
Tyr-7	26.5

However, dissociation along the target peptide backbone can produce y_nX fragment ions from crosslinks at Tyr-7, as well as from Asn-6 and Gln-5, forming the y_2X and y_3X, fragment ions, respectively. To circumvent this redundancy, we counted the pertinent fragment ions (Supplementary Tables S1-S4) for all logical crosslinks shown in Supplementary Table S5. Interestingly, CID of cross-linked ions did not produce any b_nX type of fragment ions. According to Table 1 data, Asn-6 and Tyr-7 at 25% to 27%, respectively, appeared to be the most frequently found residues in the cross-linked ions. Gln-5 appeared at 21% crosslinks followed by Gln-4 at 14%, whereas Gly-1 to Asn-3 only accounted for 3-6%. This suggested that Gln-4 to Tyr-7 were the predominant sites for crosslinking.

UVPD-CID-MS^3 experiments were also carried out for complexes ***II***, ***III***, and ***IV***. Similar to complex ***I***, the CID-MS^3 spectra of ***II***-***IV*** revealed that the majority of the complex ions did not become covalently cross-linked by UVPD-MS^2 (Table 1). However, the fractions of photo-crosslinked ions produced from the four complexes were quite different. In ***II***, the fraction of photo-crosslinked ions amounted to 4.5% (Figure 2a). A significant amount of these backbone fragment ions were of the y_kM series, which provided insight into the possible sites of crosslinking in the target peptide. The mB_n series ions were also observed but can provide little information about the site of crosslinking. Several backbone fragment ions from ***II*** could be unequivocally assigned (Supplementary Table S2), and the normalized cross-links distribution is given in Table 3.

Unlike complex ***I***, where only the y_kM ion series were observed, both the mB_k and b_kM series were observed for ***II*** in addition to the

y_kM series. However, only the b_kM and y_kM series gave us possible insight into the crosslinked residues. Since both the mB_n and b_kM series are isobars, when tabulating the y_kM and b_kM series for cross-links distribution the mB_n series were included as well. As a result, the normalized cross-link fraction for all seven residues ranges from 12% to 16%, indicating non-specific cross-links in ***II***.

Table 3. Normalized distribution of cross-links in complexes ***II*** and ***IV***.

Target peptide residue	Normalized cross-link fraction (%) *GNNQQNY$_{OH}$	*GNNQQNY$_{NH2}$
Gly-1	14.5	0
Asn-2	12.5	0
Asn-3	14	6
Gln-4	14	10
Gln-5	14	17
Asn-6	16	28.5
Tyr-7	15	38.5

Complex ***III*** gave a very small fraction of cross-linked backbone fragment ions (0.8%) (Figure 2b). The sole photo-crosslinked ion at

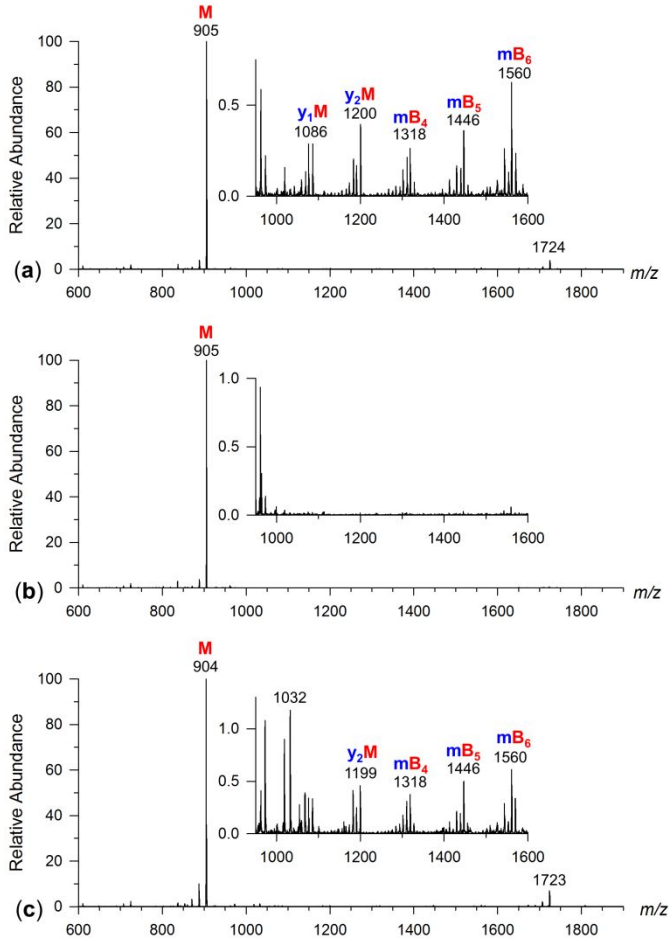


Figure 2. CID-MS^3 of $(mM–N_2+H)^+$ photoproduct ions from (a) complex ***II***, *m/z* 1741; (b) complex ***III***, *m/z* 1740; (c) complex ***IV***, *m/z* 1740.

m/z 961 was unequivocally assigned as a mB_1 fragment because of the different C-termini of the monomers. Since no other backbone fragment ions were observed for this complex, we could not draw any information about binding in complex ***III***. Multiple photo-crosslinked ions were observed for complex ***IV*** (Figure 2c), where fragment-ion assignment was facilitated by the different C-termini of the target peptide (COOH) and the photo-label peptide ($CONH_2$, Scheme 1). The peak assignments and relative intensities of the identified backbone fragment ions (Supplementary Table S4) were used to analyze the cross-linking to amino acid residues in the target peptide (Table 3). Since no b_1X and y_6X fragment ions were observed (Figure 2c) cross-linking at Gly-1 or Asn-2 can be excluded. In addition, Asn-6 and Tyr-7 appeared to be the prevalent sites of photo-crosslinking, at 29% and 38%, respectively. From the CID-MS^3 spectra of complexes ***I*** - ***IV***, it was apparent that the target peptide *C*-terminal residues were most frequently attacked by the photo-produced carbene. To summarize the binding and cross linking efficiencies, the photodissociative loss of N_2 without the peptide ion pair dissociation accounted for 86%, 91%, 88%, and 85% of the product ion channels from the parent ion complexes ***I***, ***II***, ***III***, and ***IV***, respectively (Table 1). The covalently crosslinked fragments identified by CID-MS^3 accounted for 2.3% for ***I***, 4.5% for ***II***, 0.8% for ***III***, and 4.5% for ***IV***. The overall efficiency, which is a product of the UVPD-MS^2 and CID-MS^3 efficiencies, was 2.0%, 4.1%, 0.7%, and 3.8% for ***I***, ***II***, ***III***, and ***IV***, respectively. Thus, ***II*** was the most efficient complex regarding both binding and cross-linking.

3.3. *GNNQQNY$_{NH2}$ Crosslinking with Other Peptides

In addition to GNNQQNY homodimers, we were also interested in heterodimers with another peptide motif, PQGGYQQYN, which occurs in the imperfect oligopeptide repeats subdomain of the yeast prion model Sup35.[1] According to NMR spectroscopy[2], the subdomain containing GNNQQNY can expand into the adjacent oligopeptide repeats of PQGGYQQYN. To determine whether GNNQQNY can crosslink in the gas-phase with sequences from the oligopeptide repeats, nonapeptides PAGGYYQNY$_{NH2}$ (residues 41-49) and PQGGYQQYN$_{NH2}$ (residues 75-83) were selected as target peptides in complexes with *GNNQQNY$_{NH2}$. Competitive binding of three target peptides, GNNQQNY$_{NH2}$, PAGGYYQNY$_{NH2}$ and PQGGYQQYN$_{NH2}$ with *GNNQQNY$_{NH2}$ was investigated by electrospray ionization of a solution mixture containing the photopeptide and equimolar concentrations of the target peptides. ESI produced relatively low yields of each gas-phase complexes. The singly-charged dimer ion (PAGGYYQNY$_{NH2}$ + *GNNQQNY$_{NH2}$ + H)$^+$, *m/z* 1962, denoted as $(nM+H)^+$, was formed at 0.3% relative to the combined intensities of the monomer (PAGGYYQNY$_{NH2}$ + H)$^+$, or $(n+H)^+$, and (*GNNQQNY$_{NH2}$ + H)$^+$ ions (Figure 3a). A similar relative intensity was observed for the (PQGGYQQYN$_{NH2}$ + *GNNQQNY$_{NH2}$ + H)$^+$ complex ion, *m/z* 1984, which is denoted as $(pM+H)^+$. Under the same conditions, the $(mM+H)^+$ ion had a slightly higher relative intensity of 0.5%, indicating a slightly preferred affinity to the photopeptide. The complexes were characterized by CID-MS^2 spectra that showed predominant dissociation to monomers (Supplementary Figures S6 and S7). UVPD-MS^2 photodissociation at 355 nm was used to generate carbene intermediates from *GNNQQNY$_{NH2}$ in these heterogenous dimer complexes and probe covalent cross-linking (Figure 3b). The photodissociative loss of N_2 was accompanied by partial dissociation of the dimer complex into its corresponding monomers. The surviving $(nM–N_2+H)^+$ complex from UVPD accounted for 59%, which was significantly lower than for the $(mM–N_2+H)^+$ complex ***I*** (99%). For the $(pM+H)^+$ complex, the loss of N_2 was also accompanied by less dissociation of the dimer complex (Figure 3c), whereby the surviving $(pM–N_2+H)^+$ complexes accounted for 90% of ions. These figures indicated the stabilizing effect of the polar Gln-2 and Gln-6 residues on increasing the

stability of the gas phase complex when compared with the Ala-2 and Tyr-6 analogs.

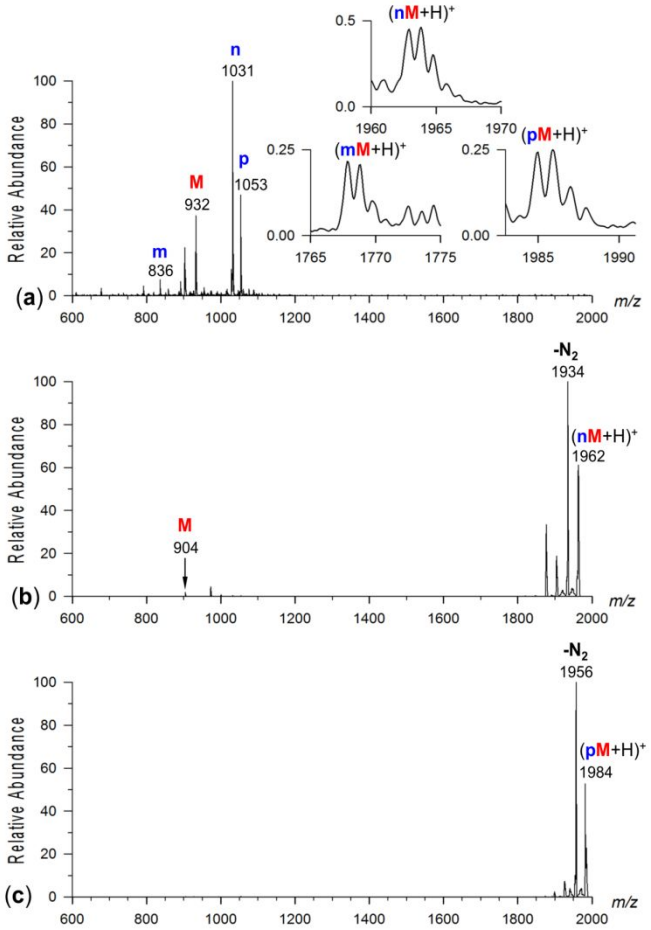


Figure 3. (a) Electrospray mass spectrum of the mixture of $PAGGYYQNY_{NH2}$, $PQGGYQQYN_{NH2}$, $GNNQQNY_{NH2}$, and $*GNNQQNY_{NH2}$. Insets show the ion profiles of the $(mM+H)^+$, $(nM+H)^+$, and $(pM+H)^+$ complexes at *m/z* 1963, 1985, and 1767, respectively. (b) UVPD-MS^2 of $(nM+H)^+$; (c) UVPD-MS^2 of $(pM+H)^+$.

At the same time, replacing the Asn-3 residue in GNNQQNY with Gly-3 in PQGGYQQYN had only a minor effect on the complex stability.

To further investigate both $(nM–N_2+H)^+$ and $(pM–N_2+H)^+$ complexes, these ions were mass-selected and subjected to CID-MS^3. The CID-MS^3 spectrum of $(nM–N_2+H)^+$ predominantly yielded monomer $(n+H)^+$ and $(M–N_2+H)^+$ ions. The $(n+H)^+$ monomer underwent further dissociation forming the b_8 and the $b_8–NH_3$ ions. We also observed the B_6 fragment ion at *m/z* 724 from the photo-labeled monomer which was commonly present in all CID-MS^3 spectra. However, CID-MS^3 revealed no logical peptide backbone fragment ions (Supplementary Table S6) that would match the theoretically generated list of all possible sequence combinations (Supplementary Table S5). The majority of the peaks in the MS^3 spectrum that were assigned corresponded to the loss of NH_3. Thus, no information on site-specific cross-linking was obtained. The CID-MS^3 spectrum of the $(pM–N_2+H)^+$ complex resulted in predominant dissociation to monomers (Supplementary Figure S8). No logical fragment ions were observed that could be assigned to photo-crosslinked complexes, as judged from the CID-MS^3 spectra analysis (Supplementary Table S7) and the list of theoretical fragment ions (Supplementary Table S8). These results indicated that binding of $*GNNQQNY_{NH2}$ to its sequence analogue (complex ***I***) was stronger than to $PAGGYYQNY_{NH2}$ or $PQGGYQQYN_{NH2}$ in the gas-phase. In addition, the binding in complex ***I***, as revealed by CID-MS^3, showed a higher degree of sequence specificity than those for the heterodimer complexes.

3.4. Complex Ion Structures and Molecular Dynamics Simulations

The cross-linking results indicated preferential interactions between the N-terminus of the $*GNNQQNY_X$ photopeptides and the C-terminal residues of the $GNNQQNY_X$ targets. However, because of the limited backbone cleavage in the target peptide moiety upon CID, the experimental data were insufficient to determine the specific residue and X-H bonds that underwent carbene insertion. To improve resolution and further specify the non-covalent interactions in the peptide complexes, we carried out combined Born-Oppenheimer molecular dynamics (BOMD) and density functional theory (DFT) calculations. In this approach[15], we used BOMD to generate multiple families of 180,000 conformers of peptide ion complexes and identified by DFT low free energy conformers in both the gas phase and aqueous solution. This provided the thermodynamically most likely conformers to be formed in solution and transferred to the gas phase. Thermal motion in the gas-phase complexes at the ion trap temperature was then treated by ten BOMD trajectories for 100 ps to provide 10^6 structure snapshots for each complex conformer that were further analyzed for close contacts between the incipient carbene and X-H bonds in the target peptide. To carry out this extensive analysis, we chose $(*GNNQQNY_{NH2} + GNNQQNY_{NH2} + H)^+$ (complex ***I***) as a representative model for molecular dynamics simulations. Complex ***I*** was selected because the C-terminal amidation on both peptide monomers made it more representative of the corresponding peptide sequence from Sup35. At the same time, complex ***I*** simplified modeling and computational work by avoiding complications with multiple zwitterionic forms of the monomers. This eliminated possible salt-bridge interactions, leaving hydrogen bonding and dispersion interactions as the dominant non-covalent interactions between the monomers. Another advantageous feature of ***I***, followed from its CID-MS^2 (Supplementary Figure S1) and CID-MS^3 spectra (Figure 1c) indicating that the photo-label monomer was preferentially protonated. This allowed us to locate the protonation site at the secondary amine of the alkylated N-terminus of the Gly residue in the the photo-label monomer. Nine lowest-energy complex structures were selected from semi-empirical PM6-D3H4 optimizations and subjected to 100 ps trajectory calculations at a bath temperature of 310 K, which corresponds to the temperature in the ion trap. The 100 ps time period is compatible with the 0.1-5 ns half-life of a carbene intermediate.[26-29] These nine lowest-energy structures were then reoptimized with DFT optimizations to evaluate their relative free energies. For the relatively large 232-atom system of complex ***I***, B3LYP calculations allowed us to obtain harmonic frequencies in a time-effective manner for evaluating ion enthalpies and entropies. To evaluate electronic energies, we used ωB97XD/6-31+G(d,p) geometry optimizations that captured dispersion interactions[39] between the monomers that were deemed important for evaluating relative free energies of complexes **1**-**9**. It should be noted that some differences in the relative energies were produced by the two functionals (Supplementary Table S9). Overall, the B3LYP functional, which does not account for dispersion interactions and is encumbered by the self-electron interaction error,[44] showed a smaller range of 29 kJ mol^{-1} for the free energies

of the five low-energy structures, with the second lowest energy complex **2** being practically isoenergetic with the global minimum **1**. When dispersion interactions were accounted for with the ωB97XD functional, the free energy range increased to 113 kJ mol^{-1} and the second lowest energy complex **2** was 27 kJ mol^{-1} above **1**. This difference diminished when solvation energies in water and methanol were included that brought **1** and **2** within 14-15 mol^{-1}, with **1** remaining the lower energy complex. Solvation effects also played a role in reordering the relative free energies of conformers **3**-**9** (Table S9). In particular, solvated structures **3** and **4** were within 25-33 kJ mol^{-1} of solvated **1** which according to our[45] and others[46] experience with peptide calculations we considered being within the limits of the energy uncertainties in ωB97XD calculations. These results indicated that including both dispersion interactions, and solvation energies was important for the evaluation of the relative stabilities of these non-covalent complexes.

The gas-phase structures of **1**-**4** are shown in Figure 4. The calculated relative enthalpies and 310 K free energies of complex ***I*** conformers **1**-**9** are summarized in Table S9 along with the ωB97XD optimized structures for **5**-**9** (Supplmentary Figure S9). The low-energy structures differed in the hydrogen bonds within and between the peptide moieties. In complex **1**, the attractive interactions between the peptide monomers involved the *Gly-1 secondary ammonium group that developed hydrogen bonds to the Gln-4 and Gln-5 side-chain amide carbonyls (Figure 4). In addition, there were multiple inter-fragment hydrogen bonds in **1** involving the Asn-2, Asn-3, Gln-4, and Tyr-7 side chains of the target peptide and Gly-1, Asn-2, Asn-3, Gln-5, Asn-6, and C-terminal amide of the photopeptide. Because of the inter-fragment hydrogen bonds, the peptide units in structure **1** were extremely entangled. In contrast, the *Gly-1 secondary ammonium group in complex **2** was internally solvated by the *Asn-6 side-chain amide, but was not involved in hydrogen bonding to the target peptide. The non-covalent bonding between the peptide residues in complex **2** was mediated by multiple neutral hydrogen bonds between the Gln-1, Asn-2 and and C-terminal amide of the target peptide with Asn-2, Gln-4, and Gln-5 residues of the photopeptide, that delineated the binding surface of the peptide units (Figure 4).

Structure **3** displayed a globular photopeptide whereas the target peptide had an extended chain. The main hydrogen bonds connecting the peptide units involved Gly-1, Asn-2, Asn-3, Gln-4, and Gln-6 of the target peptide that were bound to the C-terminal amide, Gln-5, and Asn-2 side chain amide groups. Ionic H-bonds to the alkylated N-terminal ammonium by the Asn-2 backbone amide and Asn-3 side chain were the other stabilizing and conformation forming elements. Structure **4** represented a yet different complex conformation. The Gly-1, Asn-2, and Asn-3 residues of the target peptide did not engage in H bonding to the photopeptide and made a hairpin turn at the Asn-2 residue that was enforced by intramolecular hydrogen bonding to the Gln-5 side chain of the same peptide. Bonding between the peptide units was mediated by hydrogen bonds of the target peptide Gln-4 side chain with the C-terminal amide and Gln-4 side chain of the photopeptide, Gln-5 with the Asn-6 backbone amide N-H, Asn-6 NH_2 with the Asn-2 terminal amide with the Asn-6 side-chain NH_2 and C-terminal amide carbonyl and Asn-6 C=O with the alkylated ammonium ion, and C-C=O. The barrel-like structure of complex **4** was capped at one end by the target Asn-2 residue and at the other end by the N-alkylated Gly-1 residue of the photopeptide The latter cap was reinforced by ionic H-bonding with the target peptide Asn-6 side chain and Gln-5 backbone amides. Amongst the complexes, intermolecular H-bonding in structure **4** showed the closest analogy with the dry

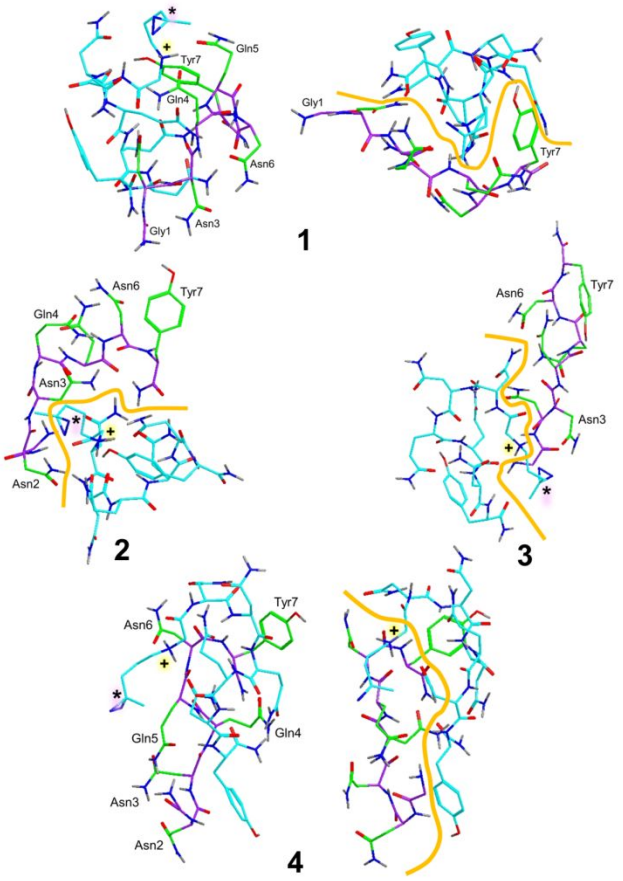


Figure 4. ωB97X-D optimized structures of complexes **1**-**4**. Backbone carbon atoms in the target peptide are shown in magenta, side-chain carbons are light green. Carbon atoms in the photopeptide are shown in cyan. Oxygens are in red, nitrogens in blue. Only exchangeable hydrogens are shown in gray. Structures **1** and **4** on the left are presented in orientations showing the diazirine ring labeled with an asterisk and the charged ammonium group. Yellow curves in structures **1** and **4** on the right and **2** and **3** indicate the interface between the peptide units.

interface of β-sheets in amyloid aggregates.[3] Similar to **4**, the sheets are composed of antiparallel backbone chains whereby the attractive interactions between the sheets are mediated by H-bonding between the Asn-2…Asn-6 side chains and the side-chains of Asn-3 to Gln-4 backbone amides. One Asn-2 residue in complex **4** also binds to the Asn-6 of the counterpart, but this interaction is to a large extent enforced by the strong H bonding of the N-alkyl ammonium (Figure 4). The Asn-3…Gln-4 backbone interactions in the β-sheets were replaced by the Asn-6 and Gln-5…Asn-6-backbone interactions in **4**. The other feature distinguishing dimer **4** from the larger aggregates was that the structurally most closely related structure was not the thermodynamically most stable one. The global energy minimum, both in the gas phase and in solution was the globular complex **1**. Complexes **2** and **3** also showed partially globular structures of the photopeptide units that were enforced by hydrogen bonding of the alkylated *N*-terminal ammonium group (Figure 4).

Ten trajectories were generated by Born-Oppenheimer molecular dynamics (BOMD) calculations for the four lowest free-energy optimized structures of $(*GNNQQNY_{NH2} + GNNQQNY_{NH2} + H)^+$ and examined for close contact points between the diazirine carbon (C-225) and the C—H, N—H, and O—H bonds of the target peptide for

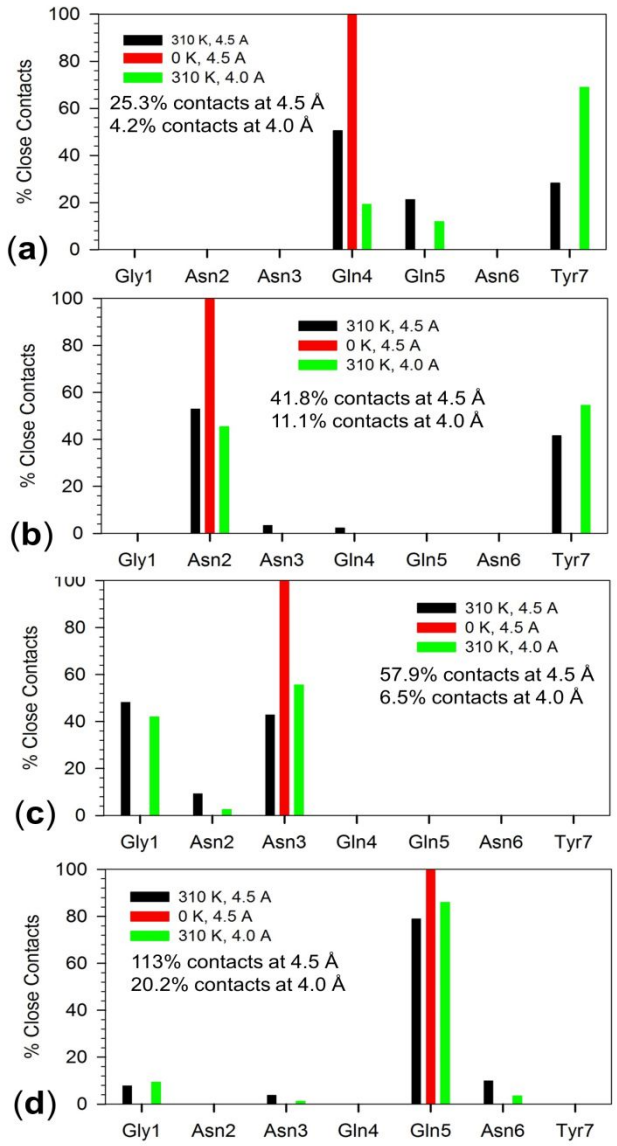
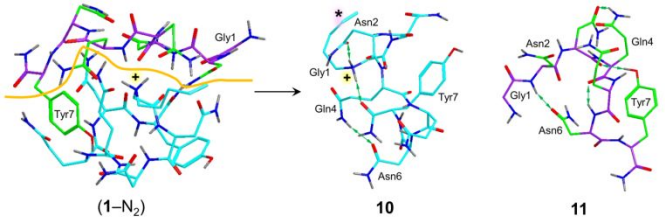


Figure 5. Normalized close contacts (average of ten trajectories) in (a) **1**, (b) **2**, (c) **3**, and (d) **4**. Black and green bars indicate contacts within 4.5 and 4.0 Å, respectively, due to thermal motion at 310 K. Red bars indicate 4.5 Å contacts in the optimized 0 K structures.

possible sites of crosslinking upon UVPD. According to Shaffer *et al.*,[14] the close contact distance between the incipient carbene carbon and the C, N, O of the target was estimated between 4-5 Å from the van der Waals radii of 1.85, 1.2, 1.4, and 1.5 Å for C, H, N, and O, respectively. Taking 4.5 Å as the closest approach, analysis of **1** showed predominant contacts with the Gln-4, Gln-5 and Tyr-7 residues of the target peptide (Figure 5a). The fraction of all 4.5 Å contacts, averaged over 10 trajectories, was 25%. This exceeded the experimental estimate of 2.6% that was based on cross-link analysis (Table 2). The number of contacts was substantially reduced to 4% when considering a stricter closest approach limit of 4.0 Å, which produced a result that was more consistent with the cross-link analysis. However, the distribution of 4.0 Å contacts in **1** changed only slightly by diminishing the participation of Gln-5 and emphasizing contacts with Tyr-7 (Figure 5a). The red bars in Figure 5 indicate the contacts that were present in the fully optimized structures at 0 K and t = 0. The black and green bars indicate the 4.5 Å and 4.0 Å contacts that developed over 100 ps upon thermal motion at 310 K. For example, the incipient carbene represented by the diazirine carbon in conformer **1** at 0 K and t = 0 was closest to the Gln-4 residue of the target peptide, forming a diazirine-Gln-4 contact. In the course of thermal motion at 310 K, the carbene could reach to Gln-5 and Tyr-7 in addition to Gln-4 while the other residues in the target peptide remained inaccessible during the 100 ps trajectory. A similar trend was observed for conformer **2**, where the initial structure primarily formed a carbene-Asn-2 contact (Figure 5b). As a result of thermal motion at 310 K, the diazirine carbon exploited the region near the C-terminus of the target peptide while still maintaining Asn-2 contacts throughout the BOMD trajectory. Out of the other two low-energy conformers (**3** and **4**), complex **3** showed most contacts at the Gly-1 and Asn-3 residues with the latter being in close contact in the local energy minimum (0 K) structure (Figure 5c). In contrast, complex **4** displayed close contacts with the Gln-5 residue while only a few new contacts were achieved by thermal motion for 100 ps (Figure 5d). The results of close contact analysis for the other, higher energy, conformers are displayed in Supplementary Figure S10. The results of contact analysis for **1** and **4** are consistent with the CID-MS^3 sequence analysis of cross-linked complexes that pointed to the Gln-4-Gln-5-Asn-6-Tyr-7 segment as the most likely target (Table 2).

3.6. Binding Energy and Dissociation of Complex *I*

To further characterize peptide-peptide ion binding in complex ***I*** following the photodissociative loss of N_2, we investigated the structure and dissociation energy of the $(mM–N_2+H)^+$ ion derived from **1**, and its dissociation products which were the neutral peptide $GNNQQNY_{NH2}$ (**11**) and its $(M–N_2+H)^+$ counterpart (**10**, Scheme 4). The internal energy of the $(mM–N_2+H)^+$ ion, denoted as (**1** – N_2), can be viewed as being composed of the mean enthalpy of the precursor ion **1** ($\Delta H(310)$ = 361 kJ mol^{-1}, Supplementary Figure S11) and the combined reaction enthalpies of photodissociation and carbene isomerization. Photodissociation of the diazirine ring at 355 nm comprises 337 kJ mol^{-1} excitation ($E_{h\nu}$) followed by an endothermic loss of N_2 that has been calculated to require $\Delta H_0(N_2$ loss) = 52-81 kJ mol^{-1} at different levels of theory.[22] An unknown factor was the kinetic and rotational energy of the departing N_2



Scheme 4. Dissociation of the $(mM–N_2+H)^+$ complex to monomers.

molecule which may be hyperthermal, especially if the dissociation proceeded on the excited (A) singlet electronic state of $(mM+H)^+$. The carbene intermediate that had not undergone cross-linking isomerized to an olefin in a highly exothermic unimolecular reaction,[22] providing additional vibrational excitation of E_{isom} = 200 kJ mol^{-1} in ion (**1** – N_2). The combined energy terms allowed us to estimate the upper bound of the internal energy in (**1** – N_2) as $E_{int} \approx \Delta H(310) + E_{h\nu} - \Delta H_0(N_2$ loss) + E_{isom} = 361 + 337 – 52 + 200 = 846 kJ mol^{-1}. The threshold energy for the dissociation of $(mM–N_2+H)^+$ into the peptide monomers was calculated from the electronic and vibrational energies of the $(mM–N_2+H)^+$ reactant (**1** – N_2) and

products **10** and **11** as $\Delta H_{0,diss}$ = 218 kJ mol^{-1}, including a counterpoise correction for the basis set superposition error[47,48] (Supplementary Table S9). Hence, peptide monomers **10** and **11** experienced substantial attractive non-covalent interactions to be broken in order to dissociate the complex in the gas phase. We note that the dissociation was favored by entropy, $\Delta S_{310,diss}$ = 311 J mol^{-1} K^{-1}, which was chiefly caused by the increased vibrational entropies of the monomers. This lowered the dissociation free energy in the gas phase to $\Delta G_{310,,diss}$ = 127 kJ mol^{-1}. Complex (**1** – N_2) of the above-estimated 846 kJ mol^{-1} internal energy was expected to undergo fast collisional cooling ion the ion trap; according to previous kinetic analysis, the internal energy drops below the dissociation threshold within 20-30 ms,[49] which will result in the stabilization of the complex. To estimate the survival time of (**1** – N_2), we employed RRKM calculations of unimolecular rate constants for dissociation to **10** and **11** (Supplementary Figure S12). The rate constant showed a large kinetic shift, reaching 10^{-12} s^{-1} for E_{int} = 900 kJ mol^{-1} which was insufficient for dissociation to proceed on the experimental time scale. This result was qualitatively consistent with the efficient survival of the (mM–N_2+H)$^+$ ions, as revealed by UVPD-MS^2 data (Table 1).

In contrast to the gas phase, the stability of (**1** – N_2) in solution was critically affected by the entropy increase upon dissociation and, especially, the solvation energies of the reactant and products. The latter term for solvation in water was obtained as $\Delta H_{solv,diss}$ = ΔH_{solv}(**10**) + ΔH_{solv}(**11**) – ΔH_{solv}(**1**–N_2) = –115 kJ mol^{-1} from ωB97X-D/6-31+G(d,p) calculations with the polarizable continuum model.[40] The combined entropy and solvation effects lowered the complex dissociation free energy in water to $\Delta G_{298,,aq,diss}$ = 12 kJ mol^{-1}, corresponding to a dissociation constant of 8 mmol L^{-1}. These values can be taken as estimates of the binding propensity of protonated *GNNQQNY and GNNQQNY in water. Although the calculated $\Delta G_{298,aq,\ diss}$ depends on the level of theory, the value reported here is in a very good qualitative agreement with the low relative yield of complexes produced by electrospray ionization. Conversely, once formed in the gas phase, the substantial binding energy in the complexes prevents their photodissociative breakup to monomers.

4. Conclusions

The experimental results and computational data allowed us to arrive at the following conclusions. Non-covalent dimer complexes of Gly-Asn-Asn-Gln-Gln-Asn-Tyr and its N-terminally tagged analogue can be produced by electrospray ionization as singly charged ions in the gas phase to model interactions in analogous peptide motifs in amyloid protein aggregates. Photodissociation of the N-terminal diazirine resulted in covalent crosslinking to amino acid residues near the C-terminus of the target peptide. Low free-energy complexes were identified that had structures with head-to-tail arrangements of the peptide monomer units and underwent thermal motion, allowing close contacts between the peptide monomers that were consistent with the cross-linking results. However, the interfaces between the peptide monomers due to side chain and backbone hydrogen bonding in the gas-phase complexes were different from those in the GNNQQNY motifs in condensed phase amyloid aggregates. This indicated that non-covalent interactions between heptapeptide motifs alone are insufficient to capture the complexity of interactions leading to protein aggregation in solution. DFT calculations revealed that the peptide complexes were thermodynamically and kinetically stable in the gas phase in agreement with photodissociation results. In contrast, solvation energies in water substantially lowered the complex stability and adversely affected complex formation by electrospray.

Conflicts of interest

There are no conflicts to declare.

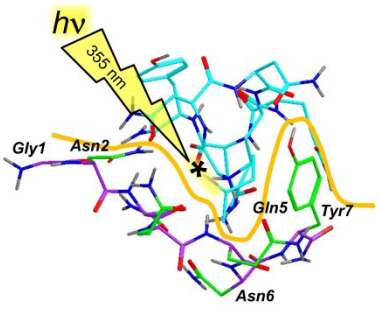
Acknowledgements

Support from the Chemistry Division of the National Science Foundation (Grants CHE-1661815 and CHE-1624430) is gratefully acknowledged. F.T. thanks the Klaus and Mary Ann Saegebarth Endowment for general support. Thanks are due to Dr. Huong T. H. Nguyen for technical assistance with calculations.

References

(1) R. Krishnan, S. L. Lindquist, *Nature* 2005, **435**, 765-772.
(2) B. H. Toyama, M. J. S. Kelly, J. D. Gross, J. S. Weissman, *Nature* 2007, **449**, 233-237.
(3) R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe, D. Eisenberg, D. *Nature* 2005, **435**, 773-778.
(4) W. Hoffmann, G. von Helden, K. Pagel, *Sci. Direct* 2017, **46**, 7-15.
(5) C. Bleiholder, N. F. Dupuis, T. Wyttenbach, M. T. Bowers, *Nature Chem*. 2011, **3**, 172-177.
(6) W. Hoffmann, K. Folmert, J. Moschner, X. Huang, X.; H. von Berlepsch, B. Koksch, M. T. Bowers, G. von Helden, K. Pagel, *J. Am. Chem. Soc.* 2018, **140**, 244-249.
(7) T. D. Do, N. J. Economou, N. E. LaPoine, W. M. Kincannon, C. Bleiholder, S. C. Feinstein, N. E. Teplow, S. K. Buratto, M. T. Bowers, *J. Phys. Chem. B* 2013, **117**, 8436-8446.
(8) A. I. Iltchev, M. J. Giammona, T. D. Do, A. G. Wong, S. K. Buratto, J.-E. Shea, D. P. Raleigh, M. T. Bowers, *J. Am. Soc. Mass Spectrom.* 2016, **27**, 1010-1018.
(9) N. F. Dupuis, C. Wu, J.-E. Shea, M. T. Bowers, *J. Am. Soc. Chem.* 2011, **133**, 7240-7245.
(10) T. N. Le, J. C. Poully, F. Lecomte, N. Nieuwjaer, B. Manil, C. Desfrancois, F. Chirot, J. Lemoine, P. Dugourd, G. van der Rest, G. Gregoire, *J. Am. Soc. Mass Spectrom.* 2013, **24**, 1937-1949.
(11) J.-H. Seo, E. Cha, H.-T. Kim, *Int. J.Mass Spectrom.* 2017, **415**, 55-62.
(12) C. E. Heo, T. S. Choi, H. I. Kim, *Int. J. Mass Spectrom.* 2018, **428**, 15-21.
(13) H. Wang, Q. Shu, D. L. Rempel, C. Frieden, M. L. Gross, *Int. J. Mass Spectrom.* 2017, **420**, 16-23.
(14) J. Ujma, V. Kopysov, N. S. Nagornova, L. G. Migas, M. G. Lizio, E. W. Blanch, C. MacPhee, O. V. Boyarkin, P. E. Barran, *Angew. Chem. Int. Ed.* 2018, **57**, 213-217.
(15) C. J. Shaffer, P. C. Andrikopoulos, J. Rezáč, L. Rulíšek, F.Tureček, *J. Am. Soc. Mass Spectrom*. 2016, **27**, 633-645.
(16) R. Pepin, C. J. Shaffer, F. Tureček, *J. Mass Spectrom*. 2017, **52**, 557-560.
(17) Y. Liu, Z. Ramey, F. Tureček, *Chem. Eur. J.* 2018, **24**, 9259-9263.
(18) H. T. H. Nguyen, P. C. Andrikopoulos, L. Rulíšek, C. J. Shaffer, F. Tureček, *J. Am. Soc. Mass Spectrom.* 2018, **29**, 1706-1720.
(19) J. Das, *Chem. Rev.* 2011, **111**, 4405-4417.
(20) M. T. H. Liu, *Chemistry of Diazirines*, Vols. I and II. CRC Press, Boca Raton, FL (1987)

(21) C. L. Gatlin, F. Tureček, *Anal. Chem.* 1994, **66**, 712-718.
(22) A. Marek, F. Tureček, *J. Am. Soc. Mass Spectrom.* 2014, **25**, 778-789.
(23) A. Marek, C. J. Shaffer, R. Pepin, K. Slováková, K. J. Laszlo, M. F. Bush, F. Tureček, F. *J. Am. Soc. Mass Spectrom.* 2015, **26**, 415-431.
(24) C. J. Shaffer, A. Marek, H. T. H. Nguyen, F. Tureček, *J. Am. Soc. Mass Spectrom.* 2015, **26**, 1367-1381.
(25) C. J. Shaffer, J. Martens, A. Marek, J. Oomens, F. Tureček, *J. Am. Soc. Mass Spectrom.* 2016, **27**, 1176-1185.
(26) R. A. Seburg, R. J. McMahon, *J. Am. Chem. Soc.* 1992, **114**, 7183-7189.
(27) J. E. Jackson, N. Soundararajan, W. White, M. T. H. Liu, R. Bonneau, M. S. Platz, *J. Am. Chem. Soc.* 1989, **111**, 6874-6875.
(28) J. P. Pezacki, P. Couture, J. A. Dunn, J. Warkentin, *J. Org. Chem.* 1999, **64**, 4456-4464.
(29) I. D. R. Stevens, M. T. H. Liu, N. Soundararajan, N. Paike, *Tetrahedron Lett.* 1989, **30**, 481-484.
(30) C. J. Shaffer, A. Marek, R. Pepin, K. Slováková, F. Tureček, *J. Mass. Spectrom.* 2015, **50**, 470-475.
(31) H. J. Berendsen, J. V. Postma, W. F. van Gunsteren, A. R. H. DiNola, J. R. Haak, *J. Chem. Phys.* 1984, **81**, 3684-3690.
(32) J. J. P. Stewart *J. Mol. Model.* 2007, **13**, 1173-1213.
(33) *J.* Řezáč, J. Fanfrlík, D. Salahub, P. Hobza, *J. Chem Theory Comput*. 2009, **5**, 1749-1760.
(34) J. J. P.Stewart MOPAC 16, Stewart Computational Chemistry. CO, USA; Colorado Springs.
(35) J. Řezáč, *J. Comput. Chem.* 2016, **37**, 1230-1237.
(36) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 16, revision A.03. Gaussian, Inc., Wallingford CT, 2016.
(37) A. D. Becke, *Phys. Rev.* A 1988, **38**, 3098–3100.
(38) C. Lee, W. Yang, R. G. Parr, *Phys. Rev.* 1988, **B37**, 785–789 .
(39) J. D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* 2008, **10**, 6615–6620.
(40) J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev*. 2005, **105**, 2999-3093.
(41) P. Roepstorff, J. Fohlman, *Biomed. Mass Spectrom*. 1984, **11**, 601.
(42) K. Biemann, K. *Methods Enzymol.* 1990, **93**, 886-887.
(43) M. M. Savitski, F. Kjeldsen, M. L. Nielsen, R. A. Zubarev, *Angew. Chem. Int. Ed.* 2006, **45**, 5301-5303.
(44) P. Mori-Sanchez, A. J. Cohen, W. Yang, W. *J. Chem. Phys.* 2006, **125**, 201102-1-4.
(45) R. Pepin, K. J. Laszlo, B. Peng, A. Marek, M. F. Bush, F. Tureček, *J. Phys. Chem. A* 2014, **118**, 308-324.
(46) J. Řezáč, D. Bím, O. Guttten, L. Rulíšek, *J. Chem. Theory Comput.* 2018, **14**, 1254-1266.
(47) S. F. Boys, F. Bernardi, *Mol. Phys*., 1970, **19**, 553.
(48) F. B. Van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, J. H. van Lenthe, *Chem. Rev.* 1994, **94**, 1873–1885.
(49) R. Pepin, F. Tureček, *J. Phys. Chem. B*, 2015, **119**, 2818-2826



Photodissociative cross-linking in the Gly-Asn-Asn-Gln-Gln-Asn-Tyr peptide ion complex revealed non-bonding interactions between the peptide units