J|A|C|S ARTICLES

Published on Web 12/30/2008

Peptide Secondary Structures in the Gas Phase: Consensus Motif of N-Linked Glycoproteins

Emilio J. Cocinero,[†] E. Cristina Stanca-Kaposta,[†] David P. Gamblin,[‡] Benjamin G. Davis,*[,‡] and John P. Simons*[,†]

Department of Chemistry, Physical and Theoretical Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QZ, United Kingdom, and the Department of Chemistry, Chemical Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 4TA, United Kingdom

Received November 5, 2008; E-mail: john.simons@chem.ox.ac.uk; ben.davis@chem.ox.ac.uk

Abstract: The possibility of secondary structure acting as a primary determinant in nature's choice of the consensus sequon, NXS/T in all N-linked glycoproteins, has been addressed by determining the intrinsic secondary structures of the capped oligopeptide, Ac-NGS-NHBn, and two "mutants", Ac-QGS-NHBn and Ac-NPS-NHBn, by use of infrared laser ion dip spectroscopy in the gas phase coupled with ab initio and density functional theory calculation. Their global minimum energy conformations, exclusively or preferentially populated in all three peptides, display marked differences. NGS adopts an open, S-shaped backbone conformation rather than the C_{10} "Asx" turn structure that all previous measurements have identified in solution; the difference can be related to the high dipole moment of the "Asx" conformation and structural selection in a polar environment. QGS adopts a similar but more rigid backbone structure, supported by markedly stronger hydrogen bonds. NPS adopts an Asx turn coupled with a C_{10} β-turn backbone conformation, a structure also adopted in a crystal environment. These and other more subtle structural differences, particularly those involving interactions with the carboxamide side chain, provide strong evidence for the operation of structural constraints, and a potential insight into the unique reactivity of the asparagine side chain toward enzymatic glycosylation.

Introduction

The cotranslational N-glycosylation of proteins is of immense importance in a wide array of biological processes including signaling, trafficking, adhesion, immune response, and many others.[1−7] The glycan can exert a profound influence on their physical behavior, not least the kinetics of folding,[8] the resulting population of preferred conformations, and their subsequent conformational stability.[9−16] N-Linked glycoproteins found in eukaryotes are created by similar enzyme complexes, oligosaccharyltransferases (OSTs), that attach a conserved glycan ($Glc_3Man_9GlcNAc_2$), directly to the δ-NH_2 group of asparagine in the consensus motif, Asn-Xaa-Ser/Thr (N-X-S/T, where X $\neq$ proline).[3−5] Why has nature settled on this particular motif? Might it present biosynthetic advantages? Does the evolutionary pressure that prevents its loss through mutation arise because it is too difficult to change or because it is too valuable to lose?

In addressing such questions, the selection of Asn as the unique N-glycosylation site has been associated with the presence of a carboxamide side chain that can form turn structures involving a hydrogen-bond interaction with the peptide backbone, $(NH)_{S/T} \rightarrow (O{=}C)_N$.[17] The resulting 10-membered ring, the so-called "Asx" turn, analogous to a backbone β-turn, brings the side chain into close proximity with the hydroxyl group located on the Ser or Thr residue. This could facilitate hydrogen-bond interactions that enhance the nucleophilicity of the normally poorly reactive carboxamide δ-NH_2 group in the side chain of Asn, providing a possible explanation for its unique reactivity. Another factor, the incompatibility of proline (Pro) at the $i + 1$ site, is strikingly suggestive of a key conformational constraint: incorporation of Pro with its ring structure would introduce a severe limit on the flexibility of the local peptide backbone. Its exclusion by nature and the maintenance of specifically conserved peptide (and glycan) motifs implicates conformation and secondary structure as

[†] Physical and Theoretical Chemistry Laboratory.
[‡] Chemistry Research Laboratory.

(1) Varki, A. *Glycobiology* **1993**, *3*, 97–130.
(2) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
(3) Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364–2369.
(4) Trombetta, E. S. *Glycobiology* **2003**, *13*, 77R–91R.
(5) Helenius, A.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019–1049.
(6) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113–159.
(7) Crocker, P. R. *Curr. Opin. Struct. Biol.* **2002**, *12*, 609–615.
(8) Brockwell, D. J.; Radford, S. E. *Curr. Opin. Struct. Biol.* **2007**, *17*, 30–37.
(9) Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759–12761.
(10) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1996**, *3*, 803–812.
(11) Lee, K.-C.; Falcone, M. L.; Davis, J. *J. Org. Chem.* **1996**, *61*, 4198–4199.
(12) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1998**, *5*, 427–437.
(13) Dempski, R. E., Jr.; Imperiali, B. *Curr. Opin. Chem. Biol.* **2002**, *6*, 844–850.
(14) Bosques, C. J.; Tschampel, S. M.; Woods, R. W.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421–8425.
(15) Ben-Dor, S.; Esterman, N.; Rubin, E.; Sharon, N. *Glycobiology* **2004**, *14*, 95–101.
(16) Hindley, M. L.; Lee, K.-C.; Davis, J. T. *J. Carbohydr. Chem.* **2005**, *24*, 261–273.
(17) Abbadi, A.; Mcharfi, M.; Aubry, A.; Premilat, S.; Boussard, G.; Marraud, M. *J. Am. Chem. Soc.* **1991**, *113*, 2729–2735.

10.1021/ja808687j CCC: $40.75

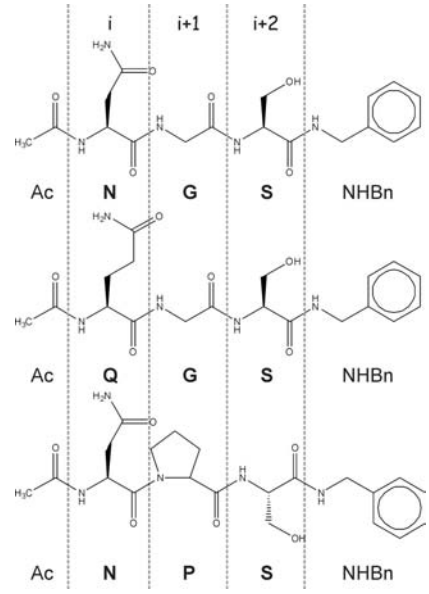
important control factors in selection during biosynthesis or in subsequent biological interactions.[18]

This well-conserved peptide motif is therefore unique and its role is not, as yet, fully explained. Not surprisingly, there have been many investigations[9−16] of the local secondary structure of oligopeptides incorporating the tripeptide sequence N-X-S/T, both crystallographic and spectroscopic, through 2D-NMR measurements coupled with molecular dynamics (MD) simulations. The NMR measurements have also addressed their secondary structures in solution before and after glycosylation, as well as the potential correlation between the efficiency of their OST-mediated glycosylation and their initial secondary structures.[19−23] The measurements have been conducted in both aqueous solution and in polar, aprotic solvents, which may provide a better simulation of a membrane environment.[16]

Glycosylation of the side chain disrupts the Asx interaction with consequences, presumably, for the secondary structure of the backbone. Synthetic oligoglycopeptides incorporating a central N-X-S/T sequon and an appropriate glycan can adopt β-turns across the Asn residue[10,12,14] promoted in part, by attractive interactions between the peptide backbone and the adjacent N-acetyl group found at C-2 in the reducing-terminal residue of the glycan[14] and perhaps also by repulsive steric constraints.[14,16] A statistical analysis of the structures of glycosylated N-X-S/T sequons in natural glycoproteins revealed unexpectedly, an overwhelming preference for extended N-X-S/T secondary structures.[23,24]

NMR and crystallographic measurements can, potentially, be influenced by environmental effects. In solution at ~300 K, many conformations may be accessed, and NMR measurements of nuclear Overhauser effects and coupling constants, which provide a set of spatial constraints, will, in combination with molecular dynamical calculations, lead to averaged conformations rather than precise atomic coordinates; hydrogen bonds are not easily detected. In an attempt to understand the key role of the N-X-S/T motif, we have adopted an alternative approach that is free from environmental constraints, based upon infrared ion dip vibrational spectroscopy[25] of capped NGS and mutant oligopeptides isolated at low temperatures in the gas phase. When coupled with ab initio and density functional theory calculation, this provides a powerful way of determining their *intrinsic* secondary structures. Individually resolved backbone conformations are reflected in the vibrational frequencies of their characteristic NH (amide A) and OH stretch bands, which are extraordinarily sensitive to local hydrogen-bond environments. Displacement of the NH or OH stretch modes toward lower wavenumbers, coupled with line broadening and enhanced relative intensities, reflect the locale and strength of hydrogen-bond interactions and provides direct, *bond-specific* structural information. This strategy is well suited to assessing the possibility of a structural basis for nature's evolutionary choice of the N-X-S/T motif and addressing related issues. Why, for example, does Asn provide the unique glycosylation site when glutamine (Gln) also offers a carboxamide side chain, albeit an extended one with an extra methylene group? Why is a neighboring proline residue (X = P) not found in natural N-linked glycoproteins? Does the intrinsic global minimum energy structure of the N-X-S/T motif incorporate an Asx turn?

Scheme 1. Diagrammatic Oligopeptide Structures



Answers to these questions have been sought by comparing the intrinsic secondary structure of the glycosylation-compatible oligopeptide containing sequon NGS, with those of two incompatible "mutants", QGS (N → Q) and NPS (G → P); see Scheme 1. Their global minimum energy conformations, which are exclusively or preferentially populated in all three peptides, display marked differences. NGS adopts an open, S-shaped backbone conformation rather than the C_{10}, Asx turn structure that all previous measurements have identified in solution. QGS adopts a similar but more rigid backbone structure that is supported by markedly stronger hydrogen bonds. NPS does adopt an Asx turn together with a C_{10} β-turn backbone conformation. The ability to identify these and other more subtle structural differences, particularly those involving interactions with the carboxamide side chain, provides a striking illustration of the power of the experimental strategy. Not only do the intrinsic secondary structures change dramatically upon mutation of key residues, but they *all* differ from those adopted in solution or within the conserved N-X-S/T sequon in naturally occurring N-linked glycoproteins.

Methods

Spectroscopy. Detailed descriptions of the experimental strategy have been published previously.[26] Peptide samples were vaporized by a laser desorption system into a supersonic jet of argon that passed through a 2 mm skimmer to create a collimated molecular beam; this was crossed by one or two tunable laser beams in the extraction region of a linear time-of-flight mass spectrometer (Jordan). Mass-selected resonant two-photon ionization (R2PI) spectra were recorded by use of a frequency-doubled pulsed Nd:YAG-pumped dye laser (Continuum Powerlite II/Sirah PS-G, 3 mJ/

(18) Ali, M. M. N.; Aich, U.; Varghese, B.; Perez, S.; Imberty, A.; Loganathan, D. *J. Am. Chem. Soc.* **2008**, *130*, 8317–8325.
(19) Imperiali, B.; Shannon, K. L. *Biochemistry* **1991**, *30*, 4374–4380.
(20) Yan, A.; Lennarz, W. J. *J. Biol. Chem.* **2005**, *280*, 3121–3124.
(21) Chen, M. M.; Glover, K. J.; Imperiali, B. *Biochemistry* **2005**, *46*, 5579–5585.
(22) Reddy, A.; Gibbs, B. S.; Liu, Y.-L.; Coward, J. K.; Changchien, L.-M.; Maley, F. *Glycobiology* **1999**, *9*, 547–555.
(23) Petrescu, A. J.; Milac, A.-L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103–114.
(24) Petrescu, A.-J.; Wormald, M. R.; Dwek, R. A. *Curr. Opin. Struct. Biol.* **2006**, *16*, 600–607.
(25) Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1–18.
(26) Çarçabal, P.; Patsias, Th.; Hünig, I.; Liu, B.; Kaposta, E. C.; Snoek, L. C.; Gamblin, D. P.; Davis, B. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 129–136.

pulse UV) operating at 10 Hz. Conformer-specific UV and IR spectra were recorded through UV−UV and IR ion dip (IRID) double-resonance spectroscopy. The IRID experiments employed radiation in the range 3100−3800 cm^{-1}, generated by difference frequency mixing of the fundamental of a pulsed Nd:YAG laser with the output of a dye laser in a $LiNbO_3$ crystal (Continuum Powerlite 8010/ND6000/IRP module). The delay between the pump and the probe laser pulses was ∼150 ns in both the IRID and UV-hole burning experiments.

Computational Strategies. The calculations began with exhaustive and unrestricted surveys of the conformational landscapes of each peptide with a wide range of molecular force fields to minimize the risk of missing some of the many conformations that could be adopted. They were identified by the Monte Carlo multiple minimization procedure as implemented in MacroModel software (MacroModel v 8.5, Schrödinger, LLC21).[27] The most stable 25 or so structures—that is, those that might have a significant population in the cooled adiabatic expansion—were reoptimized through density functional theory calculations (B3LYP/6-31+G*) by use of the Gaussian 03 program package[28] to provide a new energy ranking of the lowest energy structures and their associated harmonic vibrational spectra. Zero-point-corrected relative energies were computed through subsequent single-point ab initio calculations (MP2/6-311++G**), and final optimizations were based upon comparisons with the experimental spectra themselves, to provide feedback and guide the fine-tuning of the predicted structures.

Results

Qualitative Overview. Novel oligopeptides containing the NGS, QGS, and NPS motifs were prepared in N-acetyl-capped form; a C-terminal benzylamide ultraviolet chromophore facilitated their spectroscopic detection via resonant two-photon ionization (R2PI) (full details are provided as Figures S1−S3 in Supporting Information).

The vibrational spectra of the three peptides, isolated at low temperature (<10 K) in a supersonic molecular beam and recorded by the IR ion dip technique,[25] are shown in Figure 1, together with their corresponding R2PI spectra recorded close to the $S_0 \rightarrow S_1$ origin band. The spectra of Ac-NGS-NHBn and Ac-NPS-NHBn were unchanged when the UV probe was switched from the origin band to the weaker neighboring bands or to the underlying continuum at higher wavenumbers, confirming in each case their association with one predominant conformer. A second minor conformer was populated in Ac-QGS-NHBn; its IR spectrum is presented as Figure S4 in Supporting Information.

IR bands, associated with X−H stretching vibrations, include backbone NH (amide A) modes, Asn or Gln side-chain NH_2 modes (symmetric and antisymmetric), and the Ser OH mode. In the absence of hydrogen bonding, the bands would be located at ∼3480 cm^{-1} (NH), ∼3440 and 3550 cm^{-1} (NH_2), and ∼3680 cm^{-1} (OH). Since the observed bands are all shifted toward lower wavenumber, lying in the range ∼3250 − 3550 cm^{-1}, hydrogen-bond interactions must be ubiquitous. Each peptide presents a distinct vibrational spectrum that reflects differences in the pattern of interactions and consequently, their associated secondary structures. The band at highest wavenumber, located at ∼3545 cm^{-1} in NGS and QGS, is most likely to be associated with the Ser OH group, but its displacement to ∼3480 cm^{-1} in NPS indicates enhanced hydrogen bonding. The broad cluster

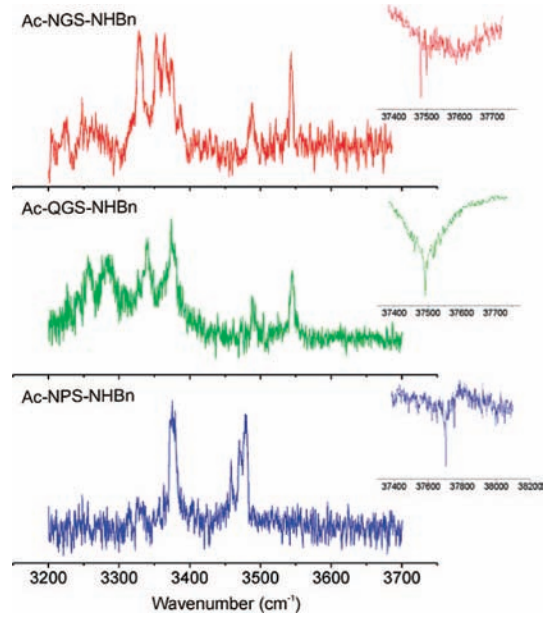


Figure 1. IR ion-depletion spectra of Ac-NGS-NHBn and its two "chemical mutants". (Inset) Corresponding UV R2PI spectra. (A second minor conformer of Ac-QGS-NHBn was also populated in the free jet expansion; see Figure S4 in Supporting Information.)

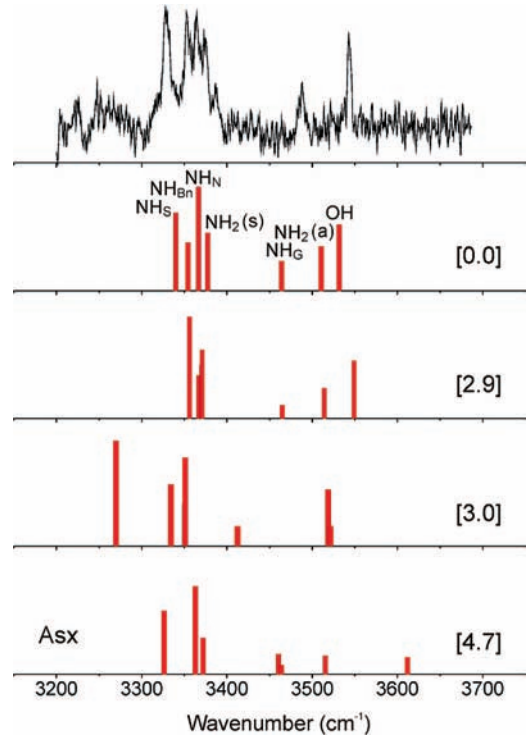


Figure 2. Experimental and computed vibrational spectra of Ac-NGS-NHBn with relative energies, in kilojoules per mole, shown in brackets; the "Asx" turn conformation is located at 4.7 kJ mol^{-1}.

of bands lying between ∼3330 and 3380 cm^{-1} in NGS, associated principally with backbone NH modes, extends downward to ∼3250 cm^{-1} in QGS, again indicating enhanced hydrogen bonding and a more rigid secondary structure.

Secondary Structures. Figures 2, 3, and 4 compare the experimental infrared ion dip spectra of Ac-NGS-NHBn and its two mutants with the computed vibrational spectra associated with their four lowest energy structures [note: in the computed

(27) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrikson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440–467.
(28) Frisch, M. J., et al. *Gaussian 03*, Revision B.03; Gaussian, Inc.: Pittsburgh, PA, 2003.

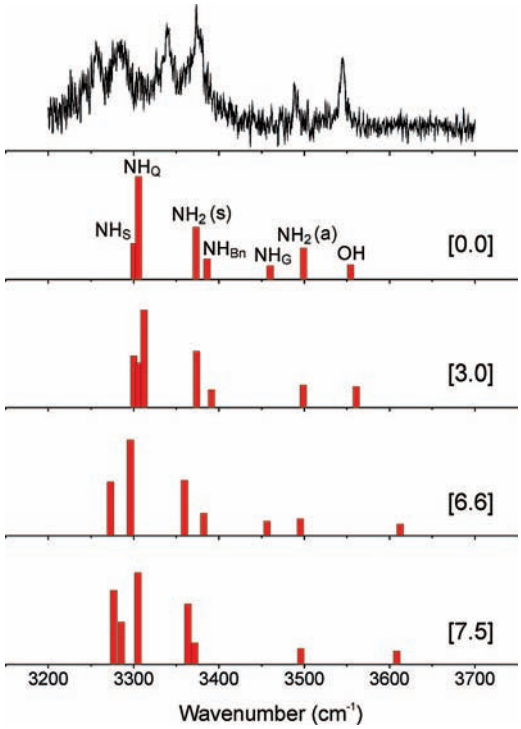


Figure 3. Experimental and computed vibrational spectra of Ac-QGS-NHBn.

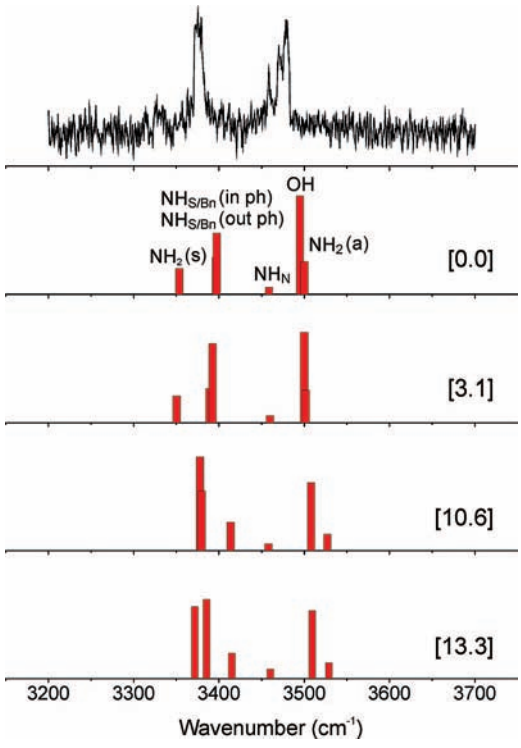


Figure 4. Experimental and computed vibrational spectra of Ac-NPS-NHBn.

spectra, wavenumbers were scaled by 0.9734 (OH) or 0.96 (NH)]. The broad agreement between the experimental spectra and those associated with the calculated global minimum energy structure of each peptide is striking. (In NPS the computed vibrational spectra of the two lowest energy structures are virtually indistinguishable since they differ only in the puckering of the proline.) The computed minimum energy structures of the three peptides, presented in Figure 5 and in Figures S5−S8 in Supporting Information, all display configurations in which the carboxamide side chain of Asn or Gln folds back around the peptide backbone to bring the side-chain nitrogen atom, N^{δ} or N^{ε}, close to the Ser residue, either $[C{=}O]_S$ in NGS and QGS or $[OH]_S$ in NPS. This is supported in NGS and QGS by an additional hydrogen bond to the peptide backbone, $[NH_2^{\delta,\varepsilon}]_{N,Q}{\rightarrow}[O{=}C]_S$, labeled a in Figure 5; there is also a more local hydrogen bond from the backbone to the side chain, $[NH]_{N,Q}{\rightarrow}[O{=}C]_{N,Q}$, labeled c in Figure 5. Despite their similarity however, the latter interaction is much stronger in QGS, where the hydrogen-bond distance $[NH]_Q{\cdots}[O{=}C]_Q$ is 1.90 Å, than in NGS, where the corresponding distance increases to 2.08 Å; see Table 1. Similarly, the $[NH_2^{\delta,\varepsilon}]_{N,Q}{\cdots}[O{=}C]_S$ distance increases from 2.10 Å in QGS to 2.20 Å in NGS. A third hydrogen bond in NGS and QGS, localized on the serine residue and labeled b in Figure 5, links the backbone $[C{=}O]_S$ to the neighboring hydroxyl group, $[OH]_S$. Hydrogen-bond interactions *along* the peptide backbones of NGS and QGS create two sequential C_7 γ-turns supported by strong $[NH]_S{\rightarrow}[O{=}C]_{N,Q}$ bonds with $[NH]_S{\cdots}[O{=}C]_{N,Q}$ distances, 2.06 Å (NGS) and 2.00 Å (QGS), and weaker $[NH]_{Bn}{\rightarrow}[O{=}C]_G$ bonds with $[NH]_{Bn}{\cdots}[O{=}C]_G$ distances of 2.11 Å (NGS) and 2.16 Å (QGS).

In NPS the set of interactions along the side chain and with the backbone are quite different. The carboxamide side chain interacts with the serine residue only, bound through a relatively weak interaction, $[NH]_S{\rightarrow}[O{=}C]_N$ with a hydrogen-bond distance of 2.14 Å, creating a C_{10} Asx turn, and a much stronger interaction linking the Ser and Asn side chains, $[OH]_S{\cdots}[O{=}C]_N$ = 1.94 Å. Along the backbone, a further relatively weak hydrogen bond linking $[NH]_{Bn}$ to $[O{=}C]_N$, with an $[NH]_{Bn}{\cdots}[O{=}C]_N$ distance of 2.13 Å, closes the C_{10} β-turn generated by the conformationally constrained central Pro residue.

Vibrational Assignments. The vibrational assignments indicated in Figures 2−4 are in good accord with the "vital statistics" of the computed peptide secondary structures. The close similarity of the bands assigned to the OH and antisymmetric NH_2 stretch modes, located at ∼3545 and ∼3488 cm^{-1} in both NGS and QGS, reflects the very similar environments of the serine hydroxyl group and the terminus of the carboxamide side chain, which are brought into close proximity in each peptide. NGS and QGS also adopt similar backbone and side-chain configurations but the stronger $[NH]_S{\rightarrow}[O{=}C]_Q$ backbone interaction in QGS shifts the corresponding Ser NH band from ∼3330 cm^{-1} (NGS) to ∼3260 cm^{-1} (QGS). Similarly, a stronger backbone to side-chain interaction, $[NH]_{N,Q}{\rightarrow}[O{=}C]_{N,Q}$ shifts the Asn/Gln NH band from ∼3360 cm^{-1} (NGS) to ∼3280 cm^{-1} (QGS): the two NH bands in QGS lie ∼200 cm^{-1} below the location of a free amide A band. The changed secondary structure in NPS, where proline exerts the principal constraint, is reflected in its very different vibrational signature (Figure 1). The displacements of the OH and antisymmetric NH_2 stretch modes toward lower wavenumber, respectively 3480 and 3335 cm^{-1} compared with 3545 and 3488 cm^{-1} in NGS, bear witness to the altered hydrogen-bonding environment.

Discussion

The intrinsic secondary structures preferentially adopted by the peptide backbones represent a compromise between the conflicting demands imposed by their interactions with the carboxamide side chain, their own internal interactions, and (in

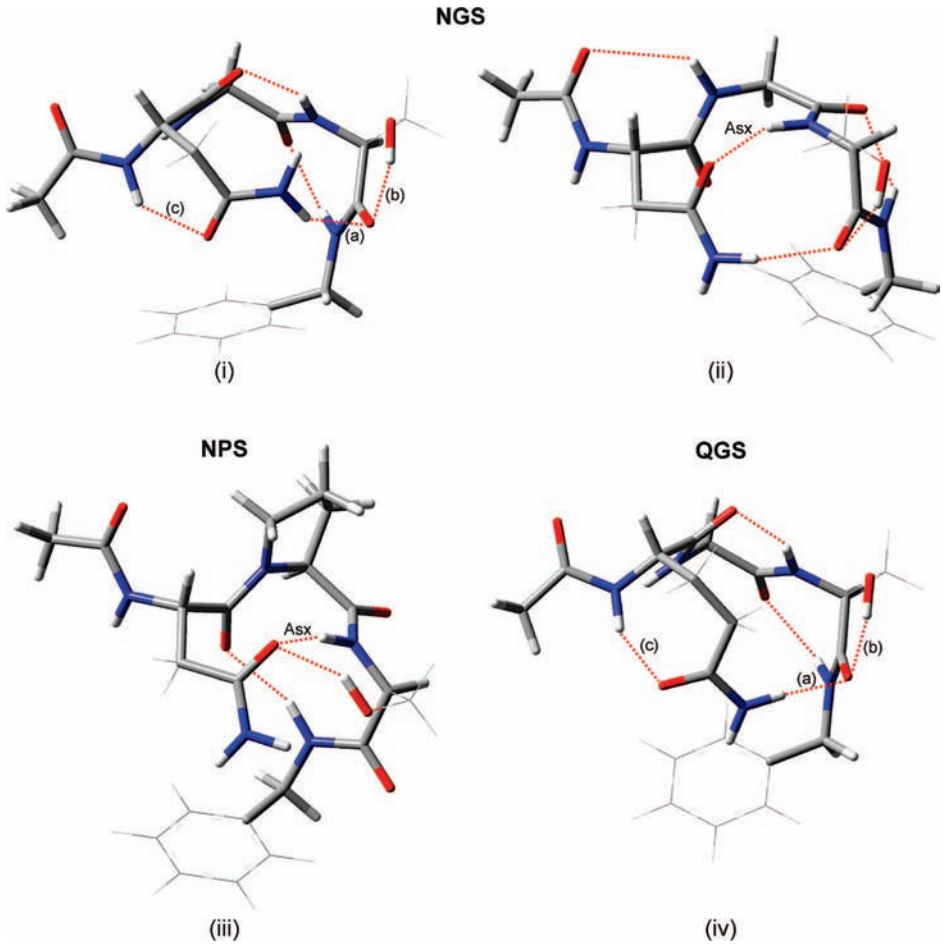


Figure 5. Populated global minimum conformations in (i) Ac-NGS-NHBn, (iii) Ac-QGS-NHBn, and (iv) Ac-NPS-NHBn. (ii) "Asx" turn conformation of Ac-NGS-NHBn, lying at relative energy 4.7 kJ mol^{-1}.

Table 1. Key Structural Parameters in the Observed Global Minimum Configurations of Ac-NGS-NHBn and Its Two Mutants, Ac-QGS-NHBn and Ac-NPS-NHBn[a]

Ac-NGS-NHBn		
$[NH]_N\cdots[\gamma\text{-}O{=}C]_N$	2.08	c
$[NH]_S\cdots[O{=}C]_N$	2.06	
$[NH]_{Bn}\cdots[O{=}C]_G$	2.11	
$[NH_2^{\delta}]_N\cdots[O{=}C]_S$	2.20	a
$[OH]_S\cdots[O{=}C]_S$	1.97	b
$[N^{\delta}]_N\cdots[OH]_S$	3.45	
Ac-QGS-NHBn		
$[NH]_Q\cdots[\delta\text{-}O{=}C]_Q$	1.90	c
$[NH]_S\cdots[O{=}C]_Q$	2.00	
$[NH]_{Bn}\cdots[O{=}C]_G$	2.16	
$[NH_2^{\varepsilon}]_Q\cdots[O{=}C]_S$	2.10	a
$[OH]_S\cdots[O{=}C]_S$	2.01	b
$[N^{\varepsilon}]_Q\cdots[OH]_S$	4.61	
Ac-NPS-NHBn		
$[NH]_S\cdots[\gamma\text{-}O{=}C]_N$	2.14	
$[NH]_{Bn}\cdots[O{=}C]_N$	2.13	
$[NH_2^{\delta}]_N\cdots[OH]_S$	2.16	
$[OH]_S\cdots[\gamma\text{-}O{=}C]_N$	1.94	
$[N^{\delta}]_N\cdots[OH]_S$	2.93	

[a] Distances are given in angstroms. a, b, and c refer to the hydrogen-bond distances identified in Figure 5.

NPS) the conformational constraint associated with the proline residue. In Ac-NGS-NHBn and each of its two mutants, the side chains of Asn or Gln are brought into close proximity with the Ser residue, reducing the distance between the carboxamide and Ser OH group, $N^{\delta,\varepsilon}\cdots OH$ to 3.45 Å in NGS, 4.61 Å in QGS, and 2.93 Å in NPS. Their close proximity and interaction are thought to be a necessary requirement of the enzyme-catalyzed N-glycosylation mechanism, which requires conferment of nucleophilicity on the target nitrogen atom, N^{δ}.[12] The conformers determined here, however, highlight strong roles for Lewis acid/base interactions between the carboxamide side chain of Asn/Gln and the backbone Ser C═O and Asn/Gln NH, labeled a and c in Figure 5. The interaction with the carboxamide, $(NH)_{N,Q}\rightarrow\gamma,\delta\text{-}C{=}O$ (c), is stronger for Gln than Asn, and the side-chain hydroxyl of Ser appears to coordinate, and perhaps modulate, the $(NH_2)\rightarrow(C{=}O)_S$ interactions (a), with the $(O{-}H)_S/(C{=}O)_S/(NH_2)$ diad, b and a in Figure 5, acting together as a bifurcated Lewis acid/base. If the transition states of OST-catalyzed glycosylation of NGS and QGS can be assumed to be similar in energy, then the tighter binding in QGS with respect to NGS, associated with intramolecular Lewis acid interactions that quench the nucleophilicity of the side-chain carboxamide nitrogen atom, will result in lower reactivity.

A recent survey of the local structures of a wide range of natural N-linked N-X-S/T glycosylation sites,[24] which identified a frequency distribution of $N^{\delta}\cdots OH(S/T)$ distances peaking at ~7.3 Å and falling close to zero for distances <5 Å, provides another insight into selection of the consensus motif. The great majority of natural N-linked glycopeptides adopt locally extended backbone structures,[23] but given the *intrinsic* structure of the free NGS sequon, the enzyme-catalyzed cotranslational glycosylation of proteins must have a dramatic effect both on

the disposition of the glycosylated carboxamide side chain and on the local secondary structure of the peptide backbone, releasing the side chain from its proximity to the serine residue and creating an extended local backbone structure. The mutation NGS → QGS would militate against these structural changes because the side chain is much more strongly bound to the backbone and it also leads to stronger $(NH)_S \rightarrow (C{=}O)_Q$ bonding along the backbone. Substitution of Pro for Gly militates against the required structural change because the imposition of a β-turn structure prevents extension of the peptide backbone.

Comparisons with NMR Experiments in Solution. The intrinsic secondary structure adopted by Ac-NGS-NHBn in the gas phase differs from the Asx turn structure assigned to related oligopeptides incorporating the N-X-S/T motif in the condensed phase. NMR,[19] IR,[17] and crystallographic[29] measurements all indicate secondary Asx-turn structures supported by an interaction between C═O on the Asn side chain and the backbone NH two residues further along, that is, located on the Ser or Thr residue. An assay of oligosaccharyltransferase- (OST-) catalyzed N-linked glycosylation of oligopeptides incorporating a series of NXT sequons in dimethyl sulfoxide (DMSO) solution identified the most efficient as those supported by this secondary structure.[19] In the gas phase however, while it is one of the lower-lying predicted structures in NGS (see Figure 5), its computed vibrational spectrum shown in Figure 2 does not correspond to the experimental one, and its computed relative energy (at 0 K) lies 4.7 kJ mol^{-1} above the global minimum (though its free energy at 298 K is calculated to be only ~2 kJ mol^{-1} higher). No low-lying Asx turn structures were identified in QGS. A similar contrast arises for the N-P-S/T sequon. In the gas phase Ac-NPS-NHBn adopts a secondary structure that incorporates an Asx turn and a backbone β-turn, both in the global minimum *and* in the each of the next three lowest-energy structures. Remarkably, it also adopts the same structure in a crystalline environment,[29] although it has been proposed that the ROESY spectrum of Ac-NPT-NH_2 in DMSO solution indicates the absence of an Asx turn.[19]

The apparent conflict between the secondary structures adopted by the consensus motif N-X-S/T in the gas phase and in DMSO solution can be resolved when the environment is taken into account. The NMR measurements have all been conducted in strongly polar solvents, water or DMSO, which will preferentially stabilize peptide configurations associated with high dipole moments. Strikingly, the dipole moment of the Asx turn structure in Ac-NGS-NHBn is calculated to be ~7.9 D but its intrinsic global minimum structure has a much lower dipole moment, ~2.5 D (see Table 2). In DMSO, which has a dielectric constant ~47, the Asx turn structure will be strongly stabilized. If a continuous dielectric is assumed, the relative stabilization energy of the Asx configuration, calculated on the basis of Onsager's PCM model, is ~9.7 kJ mol^{-1} in DMSO, considerably greater than its computed relative energy in vacuo (Table 2). Since the "forbidden" NPT sequon does present an Asx turn, either in the gas phase or in a crystal environment,[29] its inactivity is most likely associated with the enforced β-turn backbone structure rather than the suggested absence[19] of an Asx turn in solution.

Table 2. Dipole Moments and Relative Energies of the Four Lowest-Lying Conformers of Ac-NGS-NHBn and Its Mutants, Ac-QGS-NHBn and Ac-NPS-NHBn[a]

dipole moment (D)	ΔE_{rel} (vac), kJ mol^{-1}	ΔE (vac → DMSO), kJ mol^{-1}
	Ac-NGS-NHBn	
2.52	0.0	−2.0
2.63	2.9	−2.3
2.26	3.0	−0.6
7.85 (Asx)	4.7	−11.7
	Ac-QGS-NHBn	
5.64	0.0	−8.1
3.02	3.0	−2.3
3.94	6.6	−3.4
2.00	7.5	−1.0
	Ac-NPS-NHBn	
7.87 (Asx)	0.0	−12.3
8.07 (Asx)	3.1	−8.3
10.00 (Asx)	10.6	−13.4
8.92 (Asx)	13.3	−12.9

[a] Relative energies were calculated in vacuo and in DMSO (see text for details).

Conclusions

The intrinsic, global minimum secondary structure of the glycosylation-compatible consensus peptide sequon NGS resembles, but is much less tightly bound than, its incompatible "mutant", QGS. Although there are interactions between the peptide backbone and the Asn (and Gln) side chains, these do not create the Asx turn structure identified in polar solutions of NGS. The difference between the intrinsic structure of NGS in the gas phase and its structure in polar solution reflects their different dipole moments. The intrinsic secondary structure of NPS presents both an Asx turn and backbone β-turn, quite different from that of NGS and QGS. The hydrogen-bonded networks in NGS act to support an enhanced nucleophilicity of its side-chain acetamido N^{δ} atom and, perhaps, the unique enzyme-catalyzed glycosylation reactivity of the asparagine residue.

Acknowledgment. We appreciate helpful contributions made by Dr. Timothy Vaden and the support provided by EPSRC (B.G.D. and J.P.S.), the Leverhulme Trust (J.P.S., Grant F/08788G), the Spanish Ministry of Education and Science (E.J.C.), the Oxford Supercomputing Centre, the STFC Laser Loan Pool, and the Physical and Theoretical Chemistry Laboratory.

Supporting Information Available: Detailed description of the synthesis and characterization of Ac-NGS-NHBn, Ac-QGS-NHBn, and Ac-NPS-NHBn; Cartesian coordinates, total energies, and larger-scale displays of vibrational spectra and assignments of the conformations shown in Figures 2−5; and complete ref 28. This material is available free of charge via the Internet at http://pubs.acs.org.

JA808687J

(29) Marraud, M.; Aubry, A. *Biopolymers (Pept. Sci.)* **1996**, *40*, 45–83.