A Single N-Acetylgalactosamine Residue at Threonine 106 Modifies the Dynamics and Structure of Interferon α2a around the Glycosylation Site

Enzymatic addition of GalNAc to isotopically labeled IFNα2a produced in *Escherichia coli* yielded the *O*-linked glycoprotein GalNAcα-[13C,15N]IFNα2a. The three-dimensional structure of GalNAcα-IFNα2a has been determined in solution by NMR spectroscopy at high resolution. Proton-nitrogen heteronuclear Overhauser enhancement measurements revealed that the addition of a single monosaccharide unit at Thr-106 significantly slowed motions of the glycosylation loop on the nanosecond time scale. Subsequent addition of a Gal unit produced Gal(β1,3)GalNAcα-[13C,15N]IFNα2a. This extension resulted in a further decrease in the dynamics of this loop. The methodology used here allowed the first such description of the structure and dynamics of an *O*-glycoprotein and opens the way to the study of this class of proteins.

Cells infected by viruses produce type I interferons to signal the presence of the invader to uninfected neighboring cells (1). These cytokines are called interferon because they bind to specific cell-surface receptors to activate specific genes that will interfere with virus replication. The discovery of their biological functions led to the development of these proteins as therapeutics. In fact, type I interferons are among the most widely used human recombinant protein therapeutics for the treatment of several cancers and viral infections. Within the 13 α-subtypes, only two, IFNα2a (*Roferon-A™*) and IFNα2b (*Intron-A™*), are marketed for therapeutic use. Recombinant IFNα2 is a polypeptide of 165 amino acids. Subtypes a and b differ only by one residue at position 23 (Lys to Arg). The three-dimensional structure was determined by Klaus et al. (2) using NMR spectroscopy in solution. The fold consists of five α-helices (labeled A–E), four of which are arranged in an antiparallel fashion to form a left-handed four-helix bundle. Two disulfide bonds stabilize the conformation. During the development of recombinant versions of IFNα2 from *Escherichia coli* for therapeutic use, Adolf et al. (3) discovered that three subtypes (a–c) of native IFNα2 are *O*-glycosylated at Thr-106. This residue is located in the CD loop at one end of the helix bundle and may not interact with the receptor. In fact, the recently reported structures of complexes between IFNα2a and its receptor by NMR (4) and x-ray (5) do not shed light on such possible interactions. This may explain the similar biological activity observed for both the recombinant IFNα2 produced in bacteria and the glycosylated version purified from human leukocytes. However, Loignon et al. (6) reported that *O*-glycosylated IFNα2b produced in HEK293 cells inhibits viral activity more than IFNα2b prepared in *E. coli*. This suggests that the glycan does modulate biological activity.

In general, glycosylation of serine/threonine residues of proteins plays important roles in protein localization and trafficking, protein solubility, antigenicity, and cell–cell interactions (7). Recent evidence suggested that the glycan at Thr-106 might not improve the stability of the conformation under thermal stress. Using far-UV circular dichroism and calorimetry to monitor thermal unfolding, Johnston et al. (8) showed that glycosylated IFNα2b is less thermostable by ~2 °C than the nonglycosylated form. In contrast, the opposite effect was observed for glycosylated granulocyte colony-stimulating factor. The latter is more thermostable by ~5 °C compared with the nonglycosylated *E. coli*-derived version (9). These reports thus suggest that glycosylation modulates conformational stability. It may also play a role in protein solubility and half-life in serum (10). In addition, IFNα2 with covalently linked PEG exhibits a significantly increased bioavailability. One injection per week...
Effects of Glycosylation on IFNα2a Structure and Dynamics

of the N-terminally PEGylated version of IFNα2b achieved the same therapeutic result as three injections per week of the non-PEGylated protein (11). Moreover, a PEGylated version of GalNAcα-IFNα2b, in which the PEG moiety was linked to the sugar, had a significantly greater plasma concentration and half-life and 25-fold greater antiviral activity than E. coli-derived IFNα2b (12). The role played by the glycan in modulating these biophysical properties (thermostability and solubility) may be explained in terms of perturbations of the conformation and/or dynamics of the polypeptide chain. Therefore, we aimed to apply NMR spectroscopy techniques to study the effects of glycosylation of IFNα2a on its structure and dynamics.

The main challenge of using NMR spectroscopy lies in the production of glycoprotein samples enriched with stable isotopes such as carbon-13 and nitrogen-15, which are required to determine the three-dimensional structure and measure parameters sensitive to protein dynamics. Recently, a eukaryotic expression system based on human cell lines (HEK293) has been reported to produce high levels of glycosylated IFNα2b (6). Although isotope incorporation may be feasible using defined serum-free media with such eukaryotic expression systems, it is a very expensive endeavor. In addition, this approach produces glycosylated IFNα2b with a heterogeneous glycan composition. Adolf et al. (3) detected three major glycan structure variants in naturally produced IFNα2: the core-1 (or T antigen) structure Gal(β1,3)GalNAc and the latter substituted with either lactosamine linked to the GalNAc unit or sialic acid linked to the Gal moiety. Less abundant structures may incorporate fucose (3). Therefore, to address both challenges of iso- tope incorporation and glycan heterogeneity, we chose a strategy based on the production of labeled IFNα2a in E. coli, followed by an in vitro enzymatic stepwise addition of the various monosaccharide units. This approach, first proposed by Wong and co-workers (13), provides a flexible labeling scheme and facilitates the production of a homogeneous glycoprotein for NMR studies. A priori, one may consider making a fully 13C, 15N-labeled O-glycoprotein to take full advantage of triple resonance NMR techniques. Because UDP-[13C, 15N]GalNAc is not commercially available, its synthesis from the very expensive [13C, 15N]GlcNAc (14, 15) would be required. The many steps involved would further increase the final cost of producing a fully labeled sample. Instead, we opted for the preparation of a fully labeled polypeptide chain bearing unlabeled glycans. This labeling scheme was more practical in view of our intention to first study the effects of a single GalNAc residue on the protein structure and dynamics.

**EXPERIMENTAL PROCEDURES**

**Preparation of GalNAcα-[13C, 15N]IFNα2a**—The procedure for the preparation of [13C, 15N]IFNα2a has been described (16). Enzymatic addition of GalNAc to [13C, 15N]IFNα2a was carried out using human polypeptide GalNAc-transferase 2 (GalNAc-T2). The enzyme was produced in E. coli as a fusion protein with maltose-binding protein. The reaction was performed at 25 °C for 3 h with 1 mg/ml [13C, 15N]IFNα2a in 20 mM MES buffer (pH 6.0) containing 4 mM MnCl2, 150 mM NaCl, 0.005% NaN3, 0.005% Tween 80, 1 mM UDP-GalNAc, and 20 milliunits/ml MalE-GalNAc-T2. The reaction was followed by capillary electrophoresis analysis using a Beckman Coulter P/ACE MDQ system equipped with a photodiode array detector. The aliquots were denatured at 50 °C for 5 min in 0.1% SDS, 1 mM EDTA, and 2.5 mM DTT, and 2.5 mM DTT at 25 °C. GalNAcα-[13C, 15N]IFNα2a was purified as described below before galactosylation.

**Preparation of Gal(β1,3)GalNAcα-13C, 15N]IFNα2a**—Enzymatic addition of a Gal unit to purified GalNAcα-[13C, 15N]IFNα2a was carried out using a MalE fusion of the Campylobacter jejuni β1,3-galactosyltransferase (MalE-GgtB) (17). The reaction was performed at 25 °C for 4 h with 1 mg/ml GalNAcα-[13C, 15N]IFNα2a in 50 mM MES buffer (pH 6.0) containing 6 mM MnCl2, 1.0 mM DTT, 2.5 mM UDP-Gal, and 100 milliunits/ml MalE-GgtB. The reaction was followed by capillary electrophoresis as described above.

**Purification of Labeled Glycoproteins**—The reaction mixtures were desalted on a HiPrep 26/10 desalting column (GE Healthcare). The column was eluted at 5 ml/min with 1.5 column volumes of 25 mM NaOAc buffer (pH 5.0). The desalted fractions were pooled and loaded on a 5-ml HiTrap SP HP column (GE Healthcare) equilibrated in 25 mM NaOAc buffer (pH 5.0). The material was eluted at 5 ml/min using a 0–0.8 M NaCl gradient over 16 column volumes. Fractions (2 ml) absorbing at 280 nm were analyzed by SDS-PAGE (15% acrylamide) to determine which fractions contained the labeled glycoprotein.

**NMR Spectroscopy**—NMR measurements were performed at 25 °C on Bruker AVANCE III 600-MHz and AVANCE 700-MHz spectrometers equipped with a triple resonance cryogenic probe head. Samples were made of 0.7–0.9 mM labeled proteins ([15N]IFNα2a, [13C, 15N]IFNα2a, GalNAcα-[13C, 15N]IFNα2a, or Gal(β1,3)GalNAcα-[13C, 15N]IFNα2a) in buffer containing 25 mM deuterated acetic acid (pH 3.5), 0.1 mM 3-(trimethylsilyl)-1-propanesulfonic acid sodium, and 5% deuterium oxide. Resonance assignments of the backbone atoms were obtained from two-dimensional 1H, 15N heteronuclear single quantum correlation (HSQC) and 1H, 13C heteronuclear multiple quantum correlation and three-dimensional HNCO, HN(CO)CA, and CBCA(CO)NH experiments. Side chain resonances were assigned using three-dimensional HCCH total correlation spectroscopy, three-dimensional HCCH COSY, Aro-CB(CGCD)-HD, and Aro-CB(CGCD-CE)HE. For structure calculation purposes, three-dimensional 15N- and 13C-edited NOESY-HSQC, four-dimensional CC NOESY, and four-dimensional CN NOESY experiments were recorded with mixing times of 150 ms. Resonances for GalNAc were assigned using two-dimensional filtered NOESY with 0.7 mm GalNAcα-[13C, 15N]IFNα2a in buffer containing 100% D2O. All spectra were referenced relative to 3-(trimethylsilyl)-1-propanesulfonic acid sodium, processed using NMRPipe (18), and analyzed with NMRViewJ (19) and Sparky (30). The protein chemical shifts were deposited in the Biological Magnetic Resonance Data Bank (BMRB) under accession numbers 2...
TABLE 1

Statistics of the 20 conformers of GalNAcα-IFNα2a

| NOE category | Number of restraints | r.m.s.d. to averaged coordinates (Å) | Ramachandran plot quality (%) |
|--------------|---------------------|------------------------------------|-----------------------------|
| Intraresidual (i = j) | 731 | 0.9 ± 0.3 | 93.2 |
| Sequential (i = j – 1) | 712 | 1.4 ± 0.3 | 68.0 |
| Medium-range (1 < i – j ≤ 4) | 721 | 0.0 | 0.0 |
| Long-range (i – j > 4) | 660 | 6.8 | 6.8 |

For r.m.s.d., For covalent angles (Å) 0.001 0.2°

*For each hydrogen bond, two distance restraints were defined: 1.8 Å ≤ r_Cα-Cα ≤ 2.0 Å, and 2.7 Å ≤ r_Cα-Cα ≤ 3.0 Å. For each disulfide bridge, three distance restraints were used: 2.0 Å ≤ r_Cα-Cα ≤ 2.1 Å, 3.0 Å ≤ r_Cα-Cα ≤ 3.1 Å, and 3.0 Å ≤ r_Cα-Cα ≤ 3.1 Å.

18135 and 18344 for GalNAcα-IFNα2a and Gal(β1-3)GalNAcα-IFNα2a, respectively. The 15N-[2H] steady-state heteronuclear Overhauser enhancement experiment, described by Farrow et al. (20), was used to measure backbone dynamics at 298 K. The pulse sequence from the Bruker library was modified to add a relaxation delay of 3 s at the start of the pulse sequence. Three measurements at both fields (600 and 700 MHz) were performed for error estimation.

Structure Calculation—Structure calculations were performed using CYANA 2.1 (21). The amino acid library of CYANA had been modified to incorporate torsion angle parameters for an O-glycosylated threonine residue, where the hydroxyl oxygen of threonine was covalently attached to the reducing end (C1) of the GalNAc moiety in an axial orientation, resulting in an α-glycosidic bond. Cartesian coordinates of the GalNAcα-Thr residue were obtained by linking a threonine residue and GalNAc using the Chimera software package (22).

A new CYANA entity was then created using the generated atomic coordinates and editing torsion angle definitions following the template for threonine (from the CYANA standard library) and the template for GalNAc. A total of 258 backbone torsion angles were derived from chemical shifts using TALOS+ (23). In addition, calculations were performed with 66 hydrogen bond constraints inferred from the sequential short- and medium-range NOE pattern (Table 1). Structure validation was performed using the following packages: the Protein Structure Validation Software Suite (PSVS; version 1.4) from the Northeast Structural Genomics Consortium (24), the structure validation program PROCHECK (25), and the structure validation software suites PDBsum (26) and WHATIF (27). No further refinement was necessary to obtain a well resolved NMR ensemble of structures. The NMR structure ensemble was deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics (RCSB), under code 2LMS. Visualization of the structure was performed with the Chimera software package (22).

RESULTS

Sample Preparation—The preparation of glycosylated IFNα2a was carried out in two steps. First, the polypeptide was produced with carbon-13 and nitrogen-15 enrichment in E. coli. Then, an unlabeled GalNAc unit was added enzymatically in vitro to the fully labeled folded protein. This labeling scheme has two advantages: the cost effectiveness of using unlabeled monosaccharide units and the possibility to use isotope-filtered NMR experiments that can distinguish between protein and glycan resonances.

The polypeptide contains 14 serine and 10 threonine residues, half of which are well exposed to the bulk water. Considering that there are no consensus sequences for the addition of O-GalNAc at Ser/Thr residues, the choice of the appropriate glycosyltransferase isoform was determined from the work of DeFrees et al. (12). The authors subjected a synthetic peptide (residues 100–113) covering the potential glycosylation site of IFNα2a to six isoforms of GalNAc-T. Analysis of the reaction products by matrix-assisted laser desorption ionization time-of-flight mass spectrometry suggested that GalNAc-T2 added a single GalNAc unit at Thr-106. The addition of the first unit of the core-1 glycan, also referred to as the Tn antigen, was quantitative as showed by SDS-PAGE (supplemental Fig. 1). The initial analysis of the glycosylation reaction product was carried out using two-dimensional 1H,15N HSQC experiments to verify that GalNAc-T2 added the GalNAc unit at Thr-106. Comparison of the proton-nitrogen NMR correlation spectra of GalNAcα-IFNα2a and non-glycosylated IFNα2a showed that the small glycan induced the largest chemical shift perturbations at the glycosylation site and surrounding residues (Fig. 1A). The overlay revealed that both Thr-106 and Thr-108 showed large chemical shift differences. Considering that amide chemical shifts are very sensitive to the smallest change in the local magnetic environment, the large differences indicated either that Thr-108 was highly perturbed by the glycosylation of Thr-106 or that both residues had been glycosylated. Fortunately, carbon chemical shifts are better probes to detect changes in the covalent structure of the side chains of amino acids after post-translational modification. The large change in chemical shift (+9.92 ppm) observed for the β-carbon of Thr-106 in the HNCA CB and CBCA(CO)NH experiments confirmed this residue as the glycosylation site (Fig. 1B). Indeed, an average difference of 0.17 ppm was measured for the β-carbon chemical shifts of all other Ser/Thr residues.

Resonance Assignment—Protein resonance assignment was carried out following the standard protocol based on the usual array of triple resonance NMR techniques. Resonance assignment of the GalNAc unit was carried out using 13C,15N-filtered
Effects of Glycosylation on IFNα2a Structure and Dynamics

A

B

Glycosylated T106 Unglycosylated T106 CB106a

Glycosylated T108 Unglycosylated T108 CB107

15N (ppm) 1H (ppm)

13C (ppm)

1H (ppm)

13N: 116.30 ppm 15N: 116.84 ppm

1H (ppm)

13N: 114.35 ppm 15N: 116.32 ppm
possible to measure unambiguously the unique anomeric proton chemical shift were assigned using two-dimensional $^{1}H,^{13}C$ heteronuclear multiple quantum correlation (HMQC) experiments that select glycan (unlabeled) resonances by filtering out all carbon- and nitrogen-bound proton signals. After several attempts, direct observation of the GalNAc anomeric proton using filtered experiments was not successful. However, spectra from filtered experiments did show what could be attributed to a correlation between H1 and H2, from which we could obtain the chemical shift of H1. However, the symmetrical correlation (H2 to H1) was not observed (supplemental Fig. 2A). Using a two-dimensional $^{1}H,^{13}C$ heteronuclear multiple quantum correlation experiment recorded over 5 days, it was possible to measure unambiguously the unique anomic $^{1}H$ and $^{13}C$ chemical shift pair at natural abundance. Although this proton-carbon correlation resonates in a specific region of the spectrum that is sufficiently distant from the intense protein signals (supplemental Fig. 2B), it is not the case for all other proton-carbon correlations of the pyranose ring. These resonate at the edge of the fully labeled $\alpha$-carbon region and are masked by the latter. The remaining GalNAc proton chemical shifts were assigned using two-dimensional $^{13}C,^{15}N$-filtered NOESY (supplemental Fig. 2A).

Backbone Dynamics—The steady-state heteronuclear Overhauser effect between nitrogen-15 and proton was measured at two different field strengths to assess the effects of glycosylation on the protein backbone dynamics. The addition of a single GalNAc unit at Thr-106 resulted in a reduction of the backbone motions of the neighboring residues (Fig. 2), whereas the motions of all other residues were not significantly changed (supplemental Fig. 3). Elongation of the glycan chain with a galactose unit further reduced the motions of the loop surrounding the glycosylation site. Although the effects of the disaccharide on the dynamics were small, they were nevertheless significant. However, we did not pursue the determination of the three-dimensional structure of Gal(\(\beta_1,3\))GalNAc\(\alpha\)-IFN\(\alpha\).2a.

Three-dimensional Structure—The ensemble of structures of GalNAc\(\alpha\)-IFN\(\alpha\)2a shown in Fig. 3A was calculated using NOE-derived distance constraints and backbone torsion angle constraints derived from chemical shifts. The calculations were carried out with a fully extended polypeptide chain as a starting point to eliminate the possibility of any bias. The ensemble of structures shows the same overall conformation as reported by Klaus \textit{et al.} (2) for the non-glycosylated protein in solution (Protein Data Bank code 1ITF) with two differences. The first nine residues at the N-terminal end of GalNAc\(\alpha\)-IFN\(\alpha\)2a adopt a well defined pseudo-helical (twisted coil) conformation, whereas these residues are not as well defined in the IFN\(\alpha\)2a structure. The glycosylation loop (residues 101–108) is also better defined in the GalNAc\(\alpha\)-IFN\(\alpha\)2a ensemble compared with the IFN\(\alpha\)2a ensemble (Fig. 3B). The orientation of the GalNAc unit (Fig. 3C) results from the measurement of 20 distance constraints between the glycosylation loop residues and the methyl and anomic protons of the GalNAc moiety. Among the sugar-protein NOE contacts, most constraints are derived from chemical shifts. The calculations were carried out with a fully extended polypeptide chain as a starting structure. The glycosylation loop (residues 101–108) is also better defined in the GalNAc\(\alpha\)-IFN\(\alpha\)2a ensemble compared with the IFN\(\alpha\)2a ensemble (Fig. 3B). The orientation of the GalNAc unit (Fig. 3C) results from the measurement of 20 distance constraints between the glycosylation loop residues and the methyl and anomic protons of the GalNAc moiety. Among the sugar-protein NOE contacts, most constraints are derived from chemical shifts. The calculations were carried out with a fully extended polypeptide chain as a starting structure and between the anomeric proton and residues Glu-107 and Thr-108 on the other.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Overlay of two-dimensional HSQC results for IFN\(\alpha\)2a (black) and GalNAc\(\alpha\)-IFN\(\alpha\)2a (red).}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2.png}
\caption{Plot of the normalized $^{15}N{\{^{1}H\}}$ NOE measured at 25 °C and 700 MHz for $[^{13}C,^{15}N]$IFN\(\alpha\)2a (red bars), GalNAc\(\alpha\)$[^{13}C,^{15}N]$IFN\(\alpha\)2a (green bars), and Gal(\(\beta_1,3\))GalNAc\(\alpha\)$[^{13}C,^{15}N]$IFN\(\alpha\)2a (black bars) plotted for residues 90–120. The decrease in mobility is prominent for residues 105, 106, 108, and 112 (green bars). It is noteworthy that although the alteration on a nanosecond time scale motion was unaffected upon addition of the second saccharide moiety for these residues, residues 90 and 95 displayed a further decrease in mobility when the second saccharide unit was attached (black bars). For the complete plot, see supplemental Fig. 3.}
\end{figure}
Effects of Glycosylation on IFNα2a Structure and Dynamics

DISCUSSION

The structure of IFNα2a is a five-helix bundle fold with mobile segments. These are the first eight residues in N-terminal residues 42–52 and the glycosylation loop comprising residues 105–107 linking helices 3 and 4. The dynamic behavior of the N terminus is somewhat surprising considering that the first residue in the sequence is a cysteine involved in a disulfide bridge with Cys-98 located in helix 3. In contrast, residues 22–44 adopt an elongated peptide stretch terminated by a one-turn helix that displays a similar rigidity to all α-helices. There-fore, the presence of backbone motions produced an averaging of interproton distances, thus facilitating the measurement of more distance constraints involving flexible segments. The ensemble of structures (Fig. 3A) determined in solution by NMR for GalNAcα-IFNα2a shows that the glycosylation loop and the N-terminal residues have a lower r.m.s.d. compared with the ensemble obtained for IFNα2a (Fig. 3B). This suggests that glycosylation-induced rigidification of the mobile segment neighboring Thr-106 may have contributed to obtaining a better defined structure; it is most probably not the major factor. The total number of restraints used by the software CYANA to calculate the structure of GalNAcα-IFNα2a is almost double the number used for IFNα2a. Although the same three- and four-dimensional versions of NOESY experiments were recorded on both proteins, the use of cryogenic probe head technology in this study was certainly the major factor explaining the difference. Indeed, a plot of the number of restraints per residues for each amino acid in the sequence (Fig. 4) revealed that the N terminus and the glycosylation loop both show a similar increase in the number of restraints. Although it is not possible to tease apart the respective contributions of the more sensitive instrumentation and the glycosylation to this increase in the number of constraints, the latter probably benefited from the less mobile glycosylation loop.

A first glance at Fig. 3A suggests that the conformation of the GalNAc unit is not well defined. This is partly due to the fact that the glycosylation loop experiences mobility, and it is not as well defined as the helices. Therefore, we assessed the conformation of the GalNAc unit in relationship to the protein backbone of tripeptide 105–107. This was accomplished with a superimposition of the backbone atoms from residues 105–107. This was accomplished with a superimposition of the backbone atoms of residues 105–107 and the atoms of the pyranose rings are 0.5 and 1.3 Å, respectively. Therefore, the presence of backbone motions produced an averaging of interproton distances, thus facilitating the measurement of more distance constraints involving flexible segments. The ensemble of structures (Fig. 3A) determined in solution by NMR for GalNAcα-IFNα2a shows that the glycosylation loop and the N-terminal residues have a lower r.m.s.d. compared with the ensemble obtained for IFNα2a (Fig. 3B). This suggests that glycosylation-induced rigidification of the mobile segment neighboring Thr-106 may have contributed to obtaining a better defined structure; it is most probably not the major factor. The total number of restraints used by the software CYANA to calculate the structure of GalNAcα-IFNα2a is almost double the number used for IFNα2a. Although the same three- and four-dimensional versions of NOESY experiments were recorded on both proteins, the use of cryogenic probe head technology in this study was certainly the major factor explaining the difference. Indeed, a plot of the number of restraints per residues for each amino acid in the sequence (Fig. 4) revealed that the N terminus and the glycosylation loop both show a similar increase in the number of restraints. Although it is not possible to tease apart the respective contributions of the more sensitive instrumentation and the glycosylation to this increase in the number of constraints, the latter probably benefited from the less mobile glycosylation loop.

A first glance at Fig. 3A suggests that the conformation of the GalNAc unit is not well defined. This is partly due to the fact that the glycosylation loop experiences mobility, and it is not as well defined as the helices. Therefore, we assessed the conformation of the GalNAc unit in relationship to the protein backbone of tripeptide 105–107. This was accomplished with a superimposition of the backbone atoms from residues 105–107 (Fig. 3C) to use these atoms as a common reference for all conformers. Considering that few distance restraints were measured between the sugar (anomeric and methyl protons of the acetyl group only) and protons on the glycosylation loop, the conformation of the GalNAc unit was well defined with respect to the backbone of this loop (average r.m.s.d. of 1.3 Å). The orientation of the GalNAc moiety brings the amide proton of acetyl group only) and protons on the glycosylation loop. The conformation of the GalNAc unit was well defined with respect to the backbone of this loop (average r.m.s.d. of 1.3 Å). The orientation of the GalNAc moiety brings the amide proton of 

Attenuation of backbone dynamics was induced upon addition of a single GalNAc unit at Thr-106 (Fig. 2). A priori, this attenuation should reduce the averaging of interproton distances, thus facilitating the measurement of more distance constraints involving flexible segments. The ensemble of structures (Fig. 3A) determined in solution by NMR for GalNAcα-IFNα2a shows that the glycosylation loop and the N-terminal residues have a lower r.m.s.d. compared with the ensemble obtained for IFNα2a (Fig. 3B). This suggests that glycosylation-induced rigidification of the mobile segment neighboring Thr-106 may have contributed to obtaining a better defined structure; it is most probably not the major factor. The total number of restraints used by the software CYANA to calculate the structure of GalNAcα-IFNα2a is almost double the number used for IFNα2a. Although the same three- and four-dimensional versions of NOESY experiments were recorded on both proteins, the use of cryogenic probe head technology in this study was certainly the major factor explaining the difference. Indeed, a plot of the number of restraints per residues for each amino acid in the sequence (Fig. 4) revealed that the N terminus and the glycosylation loop both show a similar increase in the number of restraints. Although it is not possible to tease apart the respective contributions of the more sensitive instrumentation and the glycosylation to this increase in the number of constraints, the latter probably benefited from the less mobile glycosylation loop.
A similar conformation has been observed in synthetic O-GalNAc glycopeptides (28), suggesting that the acetyl group modulates the orientation of the GalNAc unit with the peptide backbone. In addition, the 3-hydroxyl, the attachment point of the next sugar, is pointing away from the protein, toward the bulk of the solvent. This indicates that the sugar unit adopts an appropriate conformation to allow recognition by glycoprotein-GalNAc 3β-galactosyltransferase (C1GALT1). This enzyme is responsible for elongation of the chain with galactose to make the core-1 glycan. Because a three-dimensional structure of C1GALT1 was not available, it was not possible to model the core-1 glycan. Because a three-dimensional structure of GalNAc with the GalNAc unit revealed that a number of side chain interactions in the receptor area susceptible to interaction with the hydrophobic side of the GalNAc residue. Examination of the receptor area susceptible to interaction with the GalNAc unit revealed that a number of side chain atoms of GalA near the glycosylation loop of GalNAcα-IFNa2a were missing from the crystal structure of the ternary complex. Of particular interest, tryptophan at position 183 was missing, but it was manually added. The resulting model suggests that an interaction between the hydrophobic side of the pyranose ring of the sugar and Trp-183 is possible. The ideal geometry to properly position the hydrophobic face of the sugar on top of the aromatic ring is not present. However, a minor reorientation of the backbone would achieve this accommodation. In fact, among the 20 conformers of the NMR ensemble, the model was built using the one conformer that best positioned the sugar near the tryptophan side chain. We refrained from manually fitting the interactions, but a model from a theoretical simulation might produce a realistic model. In addition, the heteronuclear Overhauser enhancement data indicate that this loop does remain sufficiently mobile and could accommodate the appropriate conformation with a negligible energy penalty. Furthermore, this reorientation of GalNAc would still leave the 3-hydroxyl pointing away toward the solvent, thus indicating that the interaction between GalNAc and Trp-183 could exist with the core-1 glycan or longer chains. This interaction may also explain the higher potency of O-glycosylated IFNa2b versus the non-glycosylated version. The additional contact between the GalNAc unit and Trp-183 would provide an increase in affinity between the cytokine and its receptor. Therefore, additional sugar residues on the glycan that increase glycan complexity and heterogeneity would display a similar potency if the longer chains did not possess excessive dynamics or steric interference. Moreover, the lower thermostability observed by Johnston et al. (8) for O-glycosylated IFNa2b may be explained by intermolecular interactions involving hydrophobic surfaces of the protein exposed during the unfolding process and the hydrophobic faces of sugar units of the glycan. The absence of such carbohydrate-protein interactions in non-glycosylated IFNa2b might explain the higher melting temperature observed.

In summary, the application of the reported enzymatic synthesis approach and the simple labeling scheme utilized to produce GalNAc-αIFNa2a allowed the determination of a three-dimensional structure at high resolution and the study of the dynamics. It is noteworthy that a single GalNAc unit suffices to modulate the backbone dynamics near the glycosylation site. To our knowledge, this is the first example of the characterization of the structure and dynamics of an O-GalNAc glycoprotein in solution. In conjunction with the recently reported structures of IFN complexes, our results shed light on the biological role of glycans.

Acknowledgments—We thank Drs. Jeremy Kunkel and Michael Johnston for critical reading of the manuscript and Dr. Warren W. Wakarchuk for useful discussions.

REFERENCES

1. Parham, P. (2005) The Immune System, 2nd Ed., p. 250, Garland Science, New York
2. Klaus, W., Gsell, B., Labhardt, A. M., Wipf, B., and Senn, H. (1997) The three-dimensional high resolution structure of human interferon α2a determined by heteronuclear NMR spectroscopy in solution. J. Mol. Biol. 274, 661–675
3. Adolf, G. R., Kalsner, I., Ahorn, H., Maurer-Fogy, I., and Cantell, K. (1991) Natural human interferon α2 is O-glycosylated. Biochem. J. 276, 511–518
4. Nudelman, I., Akabayov, S. R., Scherf, T., and Anglitster, J. (2011) Observation of intermolecular interactions in large protein complexes by 2D double-difference nuclear Overhauser enhancement spectroscopy: application to the 44 kDa interferon-receptor complex. J. Am. Chem. Soc. 133, 14755–14764
5. Thomas, C., Moraga, I., Levin, D., Kruztzik, P. O., Podoplelova, Y., Trejo, A., Lee, C., Yarden, G., Vleck, S. E., Glenn, J. S., Nolan, G. P., Piehler, J., Schreiber, G., and Garcia, K. C. (2011) Structural linkage between ligand discrimination and receptor activation by type 1 interferons. Cell 146, 621–632
6. Loignon, M., Perret, S., Kelly, J., Boulais, D., Cass, B., Bisson, L., Afkhamizareh, F., and Durocher, Y. (2008) Stable high volumetric production of glycosylated human recombinant IFNa2b in HEK293 cells. BMC Biotechnol. 8, 65–80
7. Brockhausen, I., Schachter, H., and Stanley, P. (1999) in Essentials of Glycobiology (Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., and Marth, J., eds) pp. 115–128, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
8. Johnston, M. J., Frahm, G., Li, X., Durocher, Y., and Hefford, M. A. (2011) O-Linked glycosylation leads to decreased thermal stability of interferon α2b as measured by two orthogonal techniques. Pharm. Res. 28, 1661–1667
9. Oh-eda, M., Hasegawa, M., Hattori, K., Kubonita, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T., and Ochi, N. (1990) O-Linked sugar chain of the granulocyte colony-stimulating factor protects it against polymerization and denaturation, allowing it to retain its biological activity. J. Biol. Chem. 265, 11432–11435
10. Youngster, S., Wang, Y. S., Grace, M., Bausch, J., Bordens, R., and Wyss D. F. (2002) Structure, biology, and therapeutic implications of pegylated interferon α2b. Curr. Pharm. Des. 8, 2139–2157
11. Glue, P., Fang, J. W., Rouzier-Panis, R., Raffanel, C., Sabo, R., Gupta, S. K.
Salfi, M., and Jacobs, S. (2000) Pegylated interferon α2b: pharmacokinetics, pharmacodynamics, safety, and preliminary efficacy data. Clin. Pharmacol. Ther. 68, 556–567.

12. DeFrees, S., Wang, Z.-G., Xing, R., Scott, A. E., Wang, J., Zo¨pf, D., Gouty, D. L., Sjoberg, E. R., Panneerselvam, K., Brinkman-Van der Linden, E. C., Bayer R. J., Tarp, M. A., and Clausen, H. (2006) GlycoPEGylation of recombinant therapeutic proteins produced in Escherichia coli. Glycobiology 16, 833–843.

13. Brik, A., Ficht, S., and Wong, C.-H. (2006) Strategies for the preparation of homogenous glycoproteins. Curr. Opin. Chem. Biol. 10, 638–644.

14. Zhao, G., Guan, W., Cai, L., and Wang, P. G. (2010) Enzymatic route to preparative-scale synthesis of UDP-GlcNAc/GalNAc, their analogues and GDP-fucose. Nat. Protoc. 5, 636–646.

15. Yamamoto, K., Kawai, H., and Tochikura, T. (1981) Preparation of uridine diphosphate-N-acetylgalactosamine from uridine diphosphate-N-acetylglucosamine by using microbial enzymes. Appl. Environ. Microbiol. 41, 392–395.

16. Panjwani, N., Hodgson, D. J., Sauv´e, S., and Aubin, Y. (2010) Assessment of the effects of pH, formulation and deformation on the conformation of interferon α2 by NMR. J. Pharm. Sci. 99, 3334–3342.

17. Bermatchez, S., Gilbert, M., Blanchard, M.-C., Karwaski, M.-F., Li, J., DeFrees, S., and Wakarchuk, W. W. (2007) Variants of the β1,3-galactosyltransferase CgtB from the bacterium Campylobacter jejuni have distinct specificities. Glycobiology 17, 1333–1343.

18. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRpipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277–293.

19. Johnson, B. A., and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data. J. Biomol. NMR 4, 603–614.

20. Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1995) Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. Biochemistry 34, 868–878.

21. Güntert, P. (2004) Automated NMR protein structure calculation with CYANA. Methods Mol. Biol. 278, 353–378.

22. Petersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

23. Shen, Y., Delaglio, F., Cornilescu, G., Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J. Biol. Mol. NMR 4, 213–223.

24. Bhattacharya, A., Tejero, R., and Montellione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia. Proteins 66, 778–795.

25. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26, 283–291.

26. Laskowski, R. A. (2007) Enhancing the functional annotation of PDB structures in PDBsum using key figures extracted from the literature. Bioinformatics 23, 1824–1827.

27. Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. 8, 52–56.

28. Colta, D. M., Royyuru, A. K., Williams, L. J., Glunz, P. W., Sames, D., Kuduk, S. D., Schwarz, J. B., Chen, X.-T., Danishefsky, S. J., and Live D. H. (2002) Principles of mucin architecture: structural studies on synthetic glycopeptides bearing clustered mono-, di-, tri-, and hexasaccharide glycodomains. J. Am. Chem. Soc. 124, 9833–9844.

29. Loris, R., De Greve, H., Dao-Thi, M. H., Messens, J., Imberty, A., and Wyna, L. (2000) Structural basis of carbohydrate recognition by lectin II from Ulex europaeus, a protein with a promiscuous carbohydrate-binding site. J. Mol. Biol. 301, 987–1002.

30. Goddard, T. D., and Kneller, D. G. (2008) SPARKY 3, University of California, San Francisco.