Structural Characterization by NMR of the Natively Unfolded Extracellular Domain of β-Dystroglycan: Toward the Identification of the Binding Epitope for α-Dystroglycan†

Manuela Bozzi,‡ Marzia Bianchi,‡,⊥ Francesca Sciandra,‡ Maurizio Paci,§,|| Bruno Giardina,‡ Andrea Brancaccio,*,‡ and Daniel O. Cicero*,§

Istituto di Chimica del Riconoscimento Molecolare c/o Istituto di Biochimica e Biochimica Clinica, Consiglio Nazionale della Ricerche, Università Cattolica del Sacro Cuore, Largo Francesco Vito 1, 00168 Rome, Italy, Dipartimento di Scienze e Tecnologie Chimiche, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Rome, Italy, and INFM, Sez. B, Rome, Italy

Received May 22, 2003; Revised Manuscript Received September 22, 2003

ABSTRACT: Dystroglycan (DG) is an adhesion molecule playing a crucial role for tissue stability during both early embriogenesis and adulthood and is composed by two tightly interacting subunits: α-DG, membrane-associated and highly glycosylated, and the transmembrane β-DG. Recently, by solid-phase binding assays and NMR experiments, we have shown that the C-terminal domain of α-DG interacts with a recombinant extracellular fragment of β-DG (positions 654−750) independently from glycosylation and that the linear binding epitope is located between residues 550 and 565 of α-DG. In order to elucidate which moieties of β-DG are specifically involved in the complex with α-DG, the ectodomain has been recombinantly expressed and purified in a labeled (^{13}C,^{15}N) form and studied by multidimensional NMR. Although it represents a natively unfolded protein domain, we obtained an almost complete backbone assignment. Chemical shift index, $^{1}H-^{15}N$ heteronuclear single-quantum coherence and nuclear Overhauser effect (HSQC−NOESY) spectra and $^{3}J_{HN,H\alpha}$ coupling constant values confirm that this protein is highly disordered, but $^{1}H-^{15}N$ steady-state NOE experiments indicate that the protein presents two regions of different mobility. The first one, between residues 659 and 722, is characterized by a limited degree of mobility, whereas the C-terminal portion, containing about 30 amino acids, is highly flexible. The binding of β-DG(654−750) to the C-terminal region of the α subunit, α-DG(485−620), has been investigated, showing that the region of β-DG(654−750) between residues 691 and 719 is involved in the interaction.

Dystroglycan (DG)[1] is an adhesion molecule composed of two subunits that interact tightly in a noncovalent fashion. α-DG, extracellular and highly glycosylated, binds with high affinity a number of basement membrane proteins such as laminins, agrin, perlecan, neurexin, and biglycan (*1*, *2*), whereas the transmembrane β-DG interacts with dystrophin, utrophin, and other cytosolic proteins, such as rapsyn, caveolin-3, and Grb2 (*3*−*5*). DG is an important component of the dystrophin−glycoprotein complex (DGC), formed also by sarcoglycans, dystrobrevins, syntrophins, and sarcospan, which is likely to represent, together with integrins, a major molecular bridge connecting the cytoskeleton to the surrounding basement membrane in skeletal muscle and in a wide variety of tissues (*2*, *6*), including the central and peripheral nervous system and several epithelia (*7*).

DG plays an important role in muscle stability. It was shown that in some muscular dystrophies the integrity of the DG complex is largely altered and a frequent event can be the absence of both DG subunits or only α-DG (*8*). Moreover, the disruption of the DG gene results in lethality during mouse early embryogenesis (*9*). The importance of DG biological function is further stressed by recent studies showing an involvement of DG in signal transduction, as suggested by the presence of potential SH2 and SH3 binding motifs and of a consensus sequence for tyrosine phosphorylation located within the cytodomain of β-DG (*10*, *11*).

Only few structural details are available so far on DG and its binding partners. Electron microscopy revealed that α-DG has a dumbbell-like molecular shape (*12*), whereas the N-terminal globular domain of α-DG has an autonomous folding and it is organized into two separated subdomains (*13*, *14*). On the other hand, extensive crystallographic work has been done on the LG/LNS domains of laminin and of other DG binding proteins, which are know to harbor the

† The financial support of CNR, target project "Biotechnology" and Telethon-Italy (Grant GGP030332) is gratefully acknowledged.

* Corresponding authors. D.O.C. (for NMR studies): Tel +39-06-72594449; fax +39-06-72594328; e-mail cicero@scienze.uniroma2.it. A.B. (for dystroglycan recombinant fragments): Tel +39-06-3057612; fax +39-06-3053598; e-mail a.brancaccio@uniserv.ccr.rm.cnr.it.

‡ Università Cattolica del Sacro Cuore.

§ Università di Roma "Tor Vergata".

|| INFM.

⊥ Dipartimento di Biologia, Università degli Studi "Roma Tre", Viale Guglielmo Marconi 446, 00146 Rome, Italy.

[1] Abbreviations: DG, dystroglycan; α-DG, α-dystroglycan; β-DG, β-dystroglycan; Trx, thioredoxin; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser spectroscopy; CSI, chemical shift index; IPTG, isopropyl β-D-galactopyranoside; PMSF, phenylmethanesulfonyl fluoride; Tris, tris(hydroxymethyl)aminomethane.

10.1021/bi034867w CCC: $25.00
Published on Web 10/25/2003

α-DG binding site. It has been shown that LG/LNS domains are characterized by a series of antiparallel β-strands forming a β-sandwich, showing a structure similar to the one of lectins (*15*, *16*). The molecular details of the interaction with α-DG have not yet been clarified, and although a series of positively charged residues are likely to be involved in the recognition of negatively charged carbohydrate moieties protruding from the α-DG molecule, a direct contribution of protein−protein interactions could not be ruled out (*17*, *18*). As far as β-DG is concerned, it was shown that the WW domain, the DG-binding domain of dystrophin, containing two Trp residues, needs its EF-hands to interact with β-DG (*19*). Furthermore, a synthetic peptide spanning the last 15 C-terminal amino acids of β-DG (881−895) has been cocrystallized with the WW DG-binding domain of dystrophin and the structure of the complex was obtained at the atomic level (*20*).

DG is expressed as a single precursor protein and its two subunits are formed by a posttranslational cleavage (*2*, *21*). It is likely that the direct interaction between α-DG and β-DG is one of the major factors affecting the stability of the DG complex (*22*). The C-terminal region of α-DG has been demonstrated to interact with β-DG (*23*, *24*). By an approach based on solid-phase binding assays with a series of recombinant fragments harboring progressive deletions, it has been shown that the binding epitope for β-DG resides between amino acids 550 and 585 of the C-terminal domain of α-DG (*23*). A further analysis carried out with NMR allowed us to map with higher detail the binding epitope and indicated that a linear sequence between positions 550 and 565 represents the major binding epitope for β-DG(654−750), a recombinant polypeptide corresponding to the N-terminal extracellular region of β-DG (*24*).

Previous studies have shown that β-DG(654−750) can be classified as a natively unfolded protein, showing few elements of classical secondary structures (*25*). In this paper, an extensive NMR study of β-DG(654−750) is reported. In particular, we investigated the propensity of β-DG(654−750) to form secondary structures and the mobility of the entire polypeptide chain. Moreover, an analysis of the interaction between β-DG(654−750) and α-DG(485−620) has shown that a region between residues 691 and 719 is directly involved in the binding with the C-terminal domain of α-DG.

MATERIALS AND METHODS

Preparation of Recombinant Fragments. All the DNA manipulations required for the production of the recombinant proteins under analysis have been already described elsewhere (*23*). Murine DG recombinant fragments were expressed in *Escherichia coli* BL21(DE3) as thioredoxin (Trx) fusion proteins, which also contain a N-terminal His_6 tag and a thrombin cleavage site (*26*). Bacterial cells were cultured at 37 °C in 1.2 L of normal LB medium until an OD at 600 nm of 0.8; cells were induced with IPTG and harvested after 3 h. The cellular pellets were resuspended in a lysis buffer containing 5 mM imidazole, 0.5 M NaCl, 20 mM Tris-HCl and 1 mM PMSF at pH 7.9. Cell lysis was achieved by sonication. The fusion proteins were purified by nickel nitrilotriacetate affinity chromatography. Trx-α-DG(485−620) was used as it is, whereas β-DG(654−750), with the foreign residues Gly-Ser at the N-terminus, was obtained upon thrombin cleavage. The yield of the recombinant protein was approximately 5 mg/L of bacterial culture.

Preparation of Recombinant Labeled [^{15}N,^{13}C]-*β-DG-(654−750).* The labeled recombinant fragment [^{15}N,^{13}C]-β-DG(654−750) was expressed in *E. coli* BL21(DE3) as thioredoxin fusion protein that also contains an N-terminal His_6 tag and a thrombin cleavage site. Bacterial cells were grown at 37 °C in 1.2 L of normal LB medium until an OD at 600 nm of 0.5−0.6. After centrifugation, the cellular pellet was collected and resuspended in 400 mL of M9 1× (containing 48 mM $Na_2HPO_4{\cdot}7H_2O$, 22 mM KH_2PO_4, 8.5 mM NaCl, and 18 mM $^{15}NH_4Cl$) for two times. Cells were then resuspended in 300 mL of minimal medium containing M9 1×, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% (w/v) D-glucose-$^{13}C_6$, and an appropriate amount of BME vitamins 100× solution purchased by ICN Biomedicals. After an hour of growth at 37 °C in this medium, cells were induced with IPTG and harvested after 3 h (*27*). Cell lysis and protein purification were carried out as already described. The yield of the recombinant labeled protein was approximately 1 mg/L of LB medium.

NMR Samples. [^{15}N,^{13}C]-β-DG(654−750) and Trx-α-DG-(485−620) were dialyzed in sample buffer containing 20 mM NaH_2PO_4, 0.15 M NaCl, 1.25 mM $CaCl_2$, and 1 mM $MgCl_2$ at pH 6.5 and then concentrated with Centricon 10000 to a final concentration of 300−400 μM; 5% D_2O was added. We used 0.15 M NaCl to increase the protein solubility, whereas 1.25 mM $CaCl_2$ and 1 mM $MgCl_2$ are essential for the binding between β-DG(654−750) and Trx-α-DG(485−620) and showed no influence on the chemical shifts of [^{15}N,^{13}C]-β-DG(654−750) (data not shown).

NMR Experiments. The NMR data were recorded on Bruker Avance 400 and 700 MHz spectrometers equipped with pulsed field gradient triple-resonance probes. The temperature was held constant at 298 K. Chemical shifts for ^{1}H and ^{13}C are referred to DSS, whereas for ^{15}N calibration NH_4Cl was used as standard.

The following triple resonance experiments were recorded on [^{15}N,^{13}C]-β-DG(654−750) at 400 MHz: the HNCA (*28*) and HN(CO)CA (*29*) spectra were acquired as 34 (^{15}N) × 24 (^{13}C) × 2048 (^{1}H) complex points data set; the acquisition times were 38.5, 10.8, and 319.5 ms respectively. The HNCO (*28*) spectrum was acquired as 34 (^{15}N) × 30 (^{13}C) × 2048 (^{1}H) complex points data set; the acquisition times were 38.5, 24.8, and 319.5 ms respectively. At 700 MHz the CBCANH and CBCA(CO)NH (*30*) spectra were acquired as 36 (^{15}N) × 50 (^{13}C) × 1024 (^{1}H) complex points data set; the acquisition times were 23.0, 5.1, and 90.9 ms respectively. The 3D HNN spectrum (*31*, *32*) was recorded with the following parameters: 48 complex points (^{15}N and ^{15}N) and 1024 complex points (^{1}H), 24 scans for each fid, acquisition times of 30.7 ms (^{15}N and ^{15}N) and 86.9 ms (^{1}H). The relaxation delay was set to 2 s and the total experimental time was 38 h. A water flip-back version was used.

The ^{1}H−^{15}N HSQC−TOCSY experiments (*33*) were recorded as 80 (^{1}H) × 32 (^{15}N) × 512 (^{1}H) complex points dataset with acquisition times of 9.5, 20.5, and 45.1 ms, respectively, and with mixing times ranging from 40 to 60 ms; the ^{1}H−^{15}N HSQC−NOESY (*34*) spectra were recorded as 60 (^{1}H) × 42 (^{15}N) × 1024 (^{1}H) complex points data set with acquisition times of 12.5, 36.3, and 160 ms, respectively, and with mixing times ranging from 100 to 250 ms. Two-dimensional (2D) HSQC spectra were recorded with a water flip-back version (*35*) as 100 (^{15}N) × 1024 (^{1}H) complex

```
m.musculus  485   SGVPRGGEPNQRPELKNHIDRVDAWVGTYFEVKIPSDTFYDNEDTTTDKLKLTLKLREQQ  544
h.sapiens         SGVPRGGEPNQRPELKNHIDRVDAWVGTYFEVKIPSDTFYDHEDTTTDKLKLTLKLREQQ
x.laevis          SGVP-----NTDPELKNHIDKVVAWVGTYFEVKIPPDTFYDREDGTTDNLQLTLVPRIKA
d.rerio           ---------NIKPELRNPIDQVNAYVGTYFEVKVPSDTFFDKEDGTTDKLRLTLRKG-ND
                            *  ***:* **:* *:********:*.***:*.** ***:*:***    :

m.musculus  545   LVGEKSWVQFNSNSQLMYGLPDSSHVGKHEYFMHATDKGGLSAVDAFEIHVHKRPQGDKA  604
h.sapiens         LVGEKSWVQFNSNSQLMYGLPDSSHVGKHEYFMHATDKGGLSAVDAFEIHVHRRPQGDRA
x.laevis          SAGEKMWVMLNSTSQVMYGMPDYIHIGDHEYYLKAADKAGRTAVDALEIQVRNLFQKQPS
d.rerio           VVADDSWIQFNSTSQLLYGLPDQEHAGKHEYFMQATDKGGLYAMDAFEVRVSMWGNSVKP
                   ..:. *: :**.**::**:**  * *.***:::*:**.*  *:**:*::*    :   .

m.musculus  605   PARFKARLAGDPAPVVNDIHKKIALVKKLAFAFGDRNCSSVTLQNITRGSIVVEWTNNTL  664
h.sapiens         PARFKAKFVGDPALVLNDIHKKIALVKKLAFAFGDRNCSTITLQNITRGSIVVEWTNNTL
x.laevis          PVKFHAKFHGDHNAVINDINKKILLVKKLAFAFGDRNSSSITLHNITKGSVVVDWTNNTF
d.rerio           PVLFTAVFDGDARTVTNDIHKKILLVKKLSQSFGDRNSSTITLKSITKGSIIVEWTNNSL
                  *. * * : **   * ***:*** *****: :*****.*::**:.**:**::*:****::

m.musculus  665   PLEPCPKEQIIGLSRRIADENGKPRPAFSNALEPDFKALSIAVTGSGSCRH--LQFIPVA  722
h.sapiens         PLEPCPKEQIAGLSRRIAEDDGKPRPAFSNALEPDFKATSITVTGSGSCRH--LQFIPVV
x.laevis          PTEPCPVEQVESVGKKIYDERGSPRQHFVNSVEPEYKLLNISLSFTGSCKHKNFRYIPMR
d.rerio           QQSPCPKDQIQQLSKKISDPEGKPSSIFKFTMEPDFRPSNITVRGTASCRN--YMFVPLG
                    .*** :*:  :.::* :  *.*   *  ::**:::  .*::  :.**::    ::*:

m.musculus  723   PPSPGSSAAPATEVPDRDPEKSSEDDVY   750
h.sapiens         PPRRVPSEAPPTEVPDRDPEKSSEDDVY
x.laevis          PEEPIPTAVAPTVAADRNPEKSSEDDVY
d.rerio            -EIPDPTPSPGTPAVG—AGRQSTDDVY
                      .:  . * . .  . :.* ****
```

FIGURE 1: Multiple alignment of *M. musculus* and other DG sequences. The presence of an identical residue in all the sequences analyzed is indicated by an asterisk; of conserved substitutions, by two dots; of semiconserved substitutions, by a single dot (*39*). Positions 485−653 and 654−750 correspond to the C-terminal region of α-DG and to the ectodomain of *β*-DG, respectively. The post-translational cleavage carried out by a still unknown protease occurs between residues 653 and 654 (*40*).

points data set, with eight scans for each free induction decay (fid) and with acquisition times of 64 ms (^{15}N) and 163 ms (^{1}H). The delay time between scans was 2 s, and the measuring time was 2 h. The (3D) HNHA spectrum (*36*) was recorded as 40 (^{15}N) × 52 (^{1}H) × 1024 (^{1}H) complex points data set, with acquisition times of 25.6, 9.3, and 90.1 ms, respectively (*36*). The ^{1}H−^{15}N steady-state NOE spectra (*35*), as 100 (^{15}N) × 2048 (^{1}H) complex points data set, with acquisition times of 64.1 and 163.4 ms, respectively and 100 scans for each fid, were recorded with and without proton saturation during relaxation delay, with a 5 s recycle delay. The total experimental time was 62 h. Data were processed with NMRPipe (*37*) and analyzed with the NMRView 5.0 software package (*38*).

RESULTS

Backbone NMR Resonances Assignment of β-DG(654−750). The primary structure, including the C-terminal region of α-DG and the ectodomain of *β*-DG (positions 485−750), is reported in Figure 1. A high degree of identity and similarity among different species is evident, particularly in higher vertebrates (e.g., human and mouse) but also when fish (*Danio rerio*) and amphibians (*Xenopus laevis*) are considered. All our recombinant fragments refer to the mouse sequence.

A selected region of the ^{1}H−^{15}N heteronuclear single-quantum coherence (HSQC) spectrum of the recombinant fragment [^{15}N]-*β*-DG(654−750) at pH 6.5 and 25 °C is shown in Figure 2. The low chemical shift dispersion of ^{1}H between 8.7 and 7.7 ppm is indicative of a largely unfolded polypeptide. Despite the small chemical shift dispersion of ^{1}H, ^{13}C$^{\alpha}$, and ^{13}C$^{\beta}$ nuclei, it was possible to obtain an almost complete assignment of the backbone of *β*-DG(654−750), by a combination of classical 3D NMR experiments on the double-labeled ^{15}N/^{13}C protein (*30*), i.e., HNCO, HNCA, HN(CO)CA, HNCBCA, HN(CO)CBCA, ^{1}H−^{15}N HSQC−TOCSY, and a recently published sequence, HNN, particularly useful for partially folded proteins (*31*, *32*). In fact, this experiment provides interresidue N−N connectivities exploiting the larger chemical shift dispersion of the ^{15}N resonances. All the N−H groups were assigned with the exception of the following amino acids: the two N-terminal foreign residues Gly-Ser, plus S654, I655, C669, Q673, and two stretches 701−704 and 713−715. The lack of their relative peaks in all the 3D spectra is probably due to the fast exchange of their amide protons with the solvent and/or conformational exchange in the millisecond time scale that broadened the peaks. Chemical shift values of backbone resonances were deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under accession number 5743.

The possible occurrence of elements of secondary structure was investigated by means of the chemical shift index (CSI) protocol for ^{13}C$^{\alpha}$ and ^{13}CO (*41*), and the measurement of $^{3}J_{HN,H\alpha}$, obtained by a 3D HNHA spectrum (*36*). Chemical

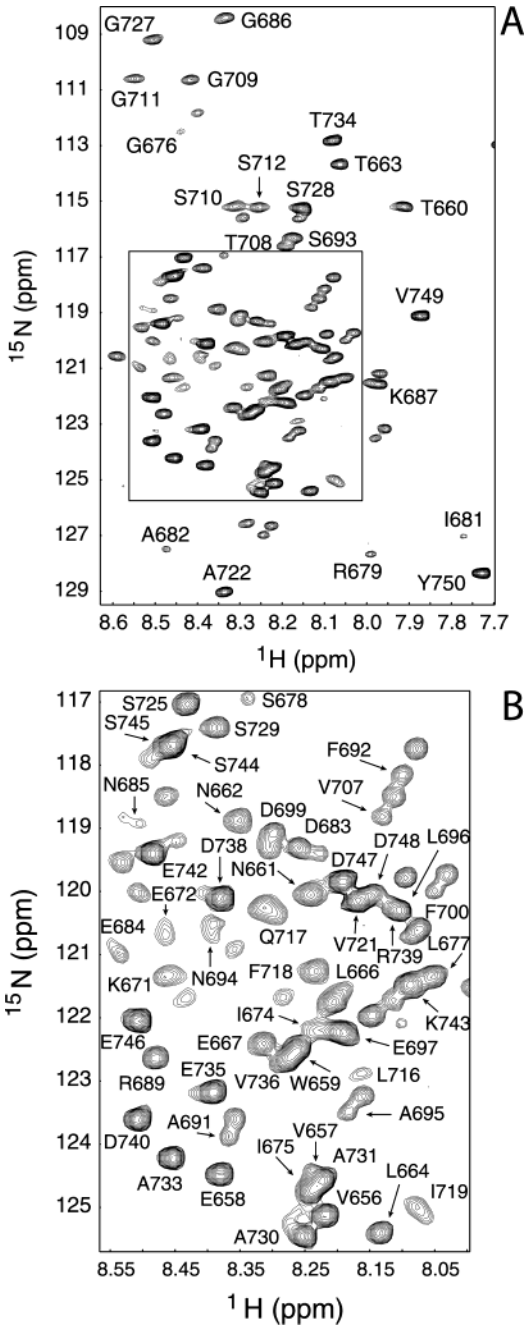

FIGURE 2: (A) ^{1}H−^{15}N HSQC spectrum (700 MHz) of 300 μM [^{15}N]-β-DG(654−750) at 25 °C in 20 mM sodium phosphate buffer at pH 6.5, and (B) its enlarged central region.

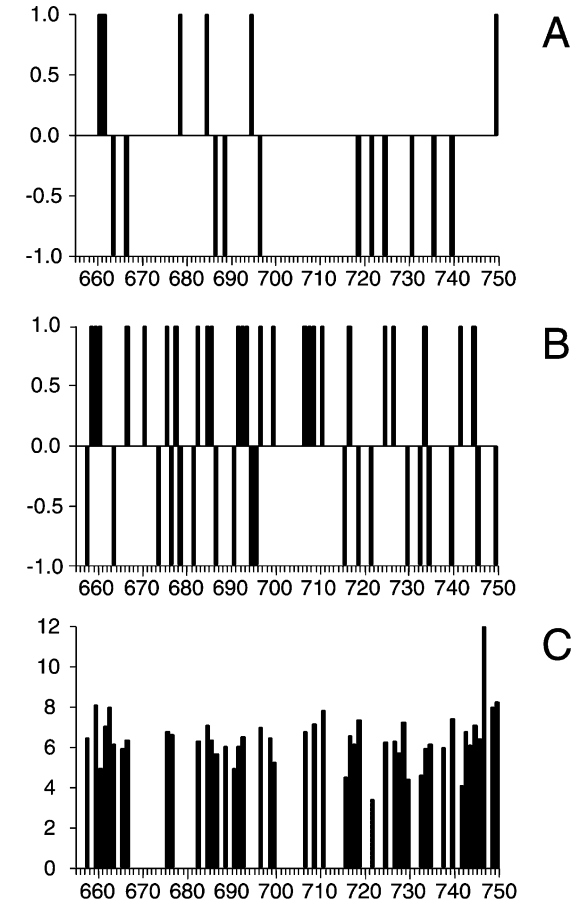

FIGURE 3: CSI for (A) $^{13}C^{\alpha}$ and (B) ^{13}CO and (C) coupling constant values, $^{3}J_{HN,H\alpha}$, versus β-DG(654−750) primary sequence.

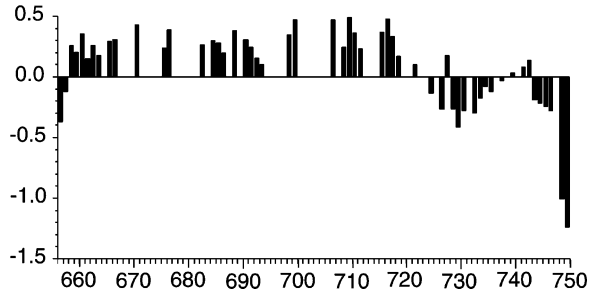

FIGURE 4: Plot of ^{1}H−^{15}N heteronuclear NOEs versus β-DG(654−750) primary sequence. Data were collected at 25 °C under conditions identical to those used for the backbone resonance assignments.

shift values of $^{13}C^{\alpha}$ and ^{13}CO are highly sensitive to their local conformation and their deviation from known values for a random coil conformation are typically indicative of classical secondary structure (*41*). In our case, no stretch with a defined CSI tendency was observed (Figure 3A,B). The same result was obtained through the analysis of the $^{3}J_{HN,H\alpha}$ couplings, which are related to the ϕ dihedral angle. Figure 3C reports the coupling constant values measured for well-resolved peaks of the HSQC spectrum. Values around 4 and 9 Hz are characteristic of α-helices and β-strands, respectively (*42*). Most of the coupling constants are within the interval 5.5−7.5 Hz, representing typical average values for flexible portions of a protein (*42*).

$^{1}H-^{15}N$ Steady-State NOE Experiments. To obtain further information about the flexibility of the polypeptide chain of β-DG(654−750), a ^{1}H−^{15}N steady-state NOE experiment was carried out. This experiment measures the NOE between the amide nitrogen and its attached proton. Structured regions of a protein show typically a ^{1}H−^{15}N NOE on the order of 0.7−0.9, depending on the magnetic field and the global correlation time of the molecule. Figure 4 represents the intensity of the NOE effect between the amide nitrogen and its proton versus the amino acidic sequence of β-DG(654−750). The analysis of the measured effects reveals that β-DG(654−750) shows no peaks with NOEs exceeding 0.5, which is again indicative of a highly mobile conformation.

Nevertheless, the trend of the NOEs suggests the presence of two distinct regions in β-DG(654−750). The first one is between positions 659 and 722 in the primary sequence and is characterized by positive NOEs. The second one, which approximately corresponds to the last 30 amino acids of the

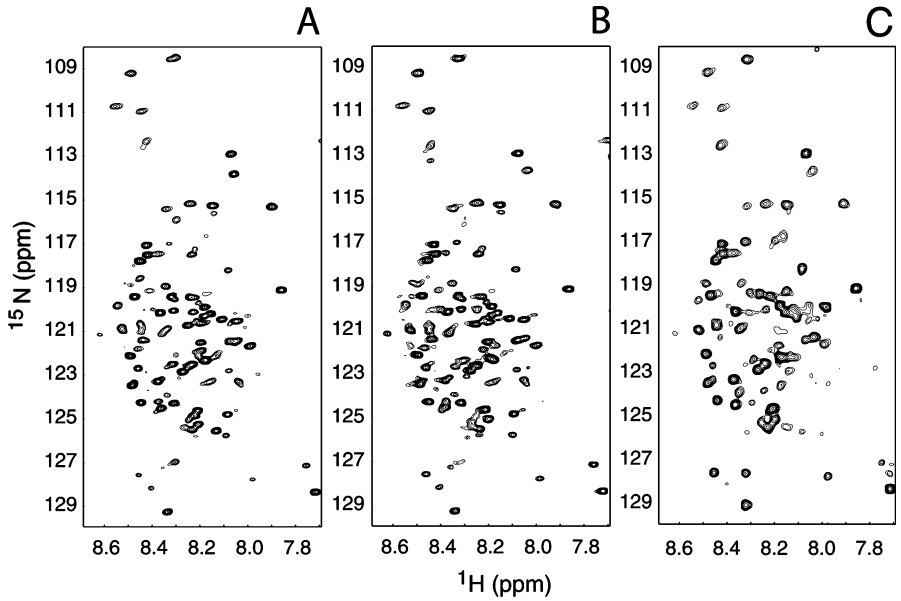

FIGURE 5: $^1H-^{15}N$ HSQC spectra (700 MHz) of 320 μM [^{15}N]-β-DG(654−750) (A) alone, (B) with 450 μM Trx, and (C) with 400 μM Trx-α-DG(485−620).

recombinant fragment β-DG(654−750), is characterized by the presence of negative NOEs. These results indicate that the C-terminal region of the protein displays a higher degree of internal motion then its N-terminal region.

Titration of [^{15}N]-β-DG(654−750) with Trx-α-DG(485−620) Fusion Protein. As shown by solid-phase binding assays, soluble biotinylated β-DG(654−750) was able to bind the C-terminal region of α-DG immobilized on microtiter plates (*43*). We have monitored by a series of 2D HSQC spectra the interaction of [^{15}N]-β-DG(654−750) with the unlabeled C-terminal domain of α-DG, in its fused form with thioredoxin, Trx-α-DG(485−620), which had better solubility than α-DG(485−620) alone in the experimental conditions used. Trx alone does not interact with [^{15}N]-β-DG(654−750) as it does not induce any changes in its $^1H-^{15}N$ HSQC spectrum (Figure 5A,B). The addition of Trx-α-DG(485−620) to [^{15}N]-β-DG(654−750) produces a line broadening and a strong reduction in the peak intensities of some residues located in the A691−I719 region, which therefore could represent the α-DG binding epitope (Figure 5C). The most influenced residues are A691, F692, S693, N694, D699, F700, V707, L716, Q717, F718, and I719. One representative region of the HSQC spectrum of [^{15}N]-β-DG(654−750) alone (Figure 6A) and in the presence of increasing amounts of Trx-α-DG(485−620) (Figure 6B−D) is reported.

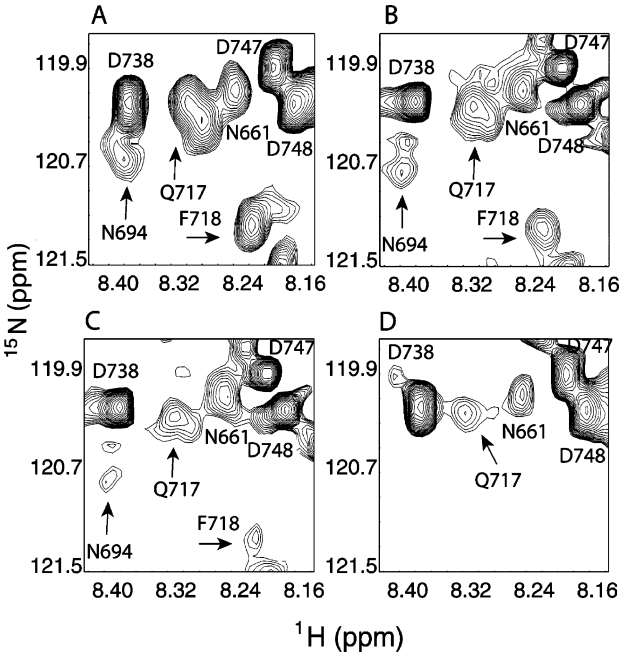

FIGURE 6: Two different regions of $^1H-^{15}N$ HSQC spectrum (700 MHz) of 320 μM [^{15}N]-β-DG(654−750) (A) alone, (B) with 160 μM unlabeled Trx-α-DG(485−620), (C) with 320 μM Trx-α-DG-(485−620), and (D) with 400 μM Trx-α-DG(485−620). Arrows indicate the most influenced residues.

DISCUSSION

Structural Analysis of β-DG(654−750). DG is an adhesion molecule that plays an important biological role since it ensures the formation of a stable molecular bridge between the cytoskeleton and the basement membrane in skeletal muscle and in a wide variety of tissues (*2*). DG is formed by two subunits, α and β, interacting tightly but in a noncovalent fashion. The interaction between DG's subunits is crucial for the formation of the DG complex and the intersubunit binding surface is likely to represent one important "hot spot", whose perturbation might influence the diverse functional activities of DG. We have already shown that the binding epitope for β-DG is located within the C-terminal region of α-DG and resides in a region between residues 550 and 565 (*23*, *24*). In this study, we carried out a series of NMR experiments in order to collect further structural information on a recombinant fragment corresponding to the extracellular N-terminal region of the β subunit, β-DG(654−750), which we had previously analyzed by fluorescence, circular dichroism, and solid-phase binding assays (*25*, *44*).

Using a set of multinuclear and multidimensional NMR experiments, we were able to obtain an almost complete assignment of the backbone nuclei of β-DG(654−750). Both the CSI analysis of $^{13}C^{\alpha}$ and ^{13}CO chemical shifts and $^3J_{HN,H\alpha}$ coupling constant values indicate a very flexible and

disordered conformation for the protein in solution. Moreover, the reduced number of sequential connectivities in the $^{1}H-^{15}N$ HSQC−NOESY spectrum confirms the absence of stable elements of secondary structure. All these data provide compelling evidence that β-DG(654−750) molecules represent an *ensemble* of different populations of disordered polypeptide chains. Interestingly, $^{1}H-^{15}N$ steady-state NOE experiments (Figure 4) show that the region between residues 659 and 722 presents weakly positive NOEs, which are indicative of some degree of motional restriction, whereas the C-terminus (≈722−750) appear to be highly flexible, as supported by the presence of negative NOEs.

Toward the Identification of the Epitope for α-DG Binding. Although the ectodomain of β-DG displays an unfolded conformation, it has been previously shown that it is able to bind recombinant fragments belonging to the C-terminal domain of α-DG in solid-phase binding assays (*43*). The binding epitope for β-DG has been identified to be located between amino acids 550 and 565 of α-DG (*23*, *24*) (Figure 1). Furthermore, it was shown that a complex between a synthetic peptide corresponding to the 550−585 sequence of α-DG and the recombinant ectodomain of β-DG is established also in solution (*24*). Similar to the C-terminal region of α-DG, the extreme conservation of the primary sequence of the β-DG ectodomain makes it very difficult to identify a *bona fide* specific sequence that would represent a unique conserved binding epitope for α-DG (*23*) (see Figure 1). In the case of the C-terminal region of α-DG, after the identification of the 550−565 binding epitope for β-DG, it was observed that some residues were conserved also in the DG-like proteins of invertebrate species such as *Drosophila melanogaster* or *Caenorhabditis elegans* (*23*). This does not emerge when the sequences corresponding to the β-DG ectodomain from different species are analyzed (data not shown). To identify the binding epitope for α-DG, further NMR experiments on the ^{15}N-labeled β-DG ectodomain in the presence of the recombinant C-terminal domain of α-DG were carried out.

The titration of [^{15}N]-β-DG(654−750) with Trx-α-DG-(485−620) shows that Trx-α-DG(485−620) does not induce any increase of the chemical shift dispersion that would be indicative of a significant conformational rearrangement of β-DG(654−750). Nevertheless, the addition of Trx-α-DG-(485−620) results in a line broadening and a reduction of the intensity of the peaks located between residues A691 and I719. In particular, the most influenced residues are A691, F692, S693, N694, D699, F700, V707, L716, Q717, F718, and I719. Differential effects on these residues can be interpreted either as a direct involvement in the interaction with Trx-α-DG(485−620) or as a consequence of a local conformational change induced by the binding. Further experiments will be necessary to identify the amino acids directly interacting with α-DG.

The region A691−I719, although basically unfolded, shows a minor degree of mobility when compared with the C-terminus. Accordingly, it was observed that the highest heteronuclear NOEs belong to the region (A691−I719) influenced by Trx-α-DG(485−620). This indeed suggests that the N-terminal portion of β-DG ectodomain is prone to form contacts with Trx-α-DG(485−620). In other words, the interaction with the C-terminal domain of α-DG is likely to drive the conformational *ensemble* toward a unique extended conformation. The mechanism of interaction would involve a strong reduction of entropy that could result in a low association rate for the complex between α- and β-DG. This hypothesis is supported by kinetic analysis of our data.

Typically, in NMR experiments three different time regimes of exchange can be distinguished: fast, intermediate, and slow (*44*). Namely, if the exchange between free and bound forms of β-DG(654−750) is fast, a progressive difference in the chemical shifts of the protein residues sensitive to the binding emerges during the titration: if the exchange is intermediate a line shape broadening of the peaks occurs upon ligand addition, whereas if the exchange is slow two distinct peaks of different heights at different chemical shifts appear. However, it should be noted that the line broadening could be due to the complexation of a small molecular weight compound (β-DG) with a large molecular weight compound (Trx-α-DG), which naturally leads to broader lines, regardless of the exchange rate.

In our system, a line shape broadening of some peaks and a strong reduction of their intensity induced by Trx-α-DG-(485−620) occur in the spectra. This allows us to rule out that the binding is in a fast exchange (Figure 6). The differences in chemical shift between bound and free β-DG are unknown and we could not discriminate whether the interaction is in a regime of intermediate or slow exchange. Nevertheless, we could expect a k_{off} in the range of 1 s^{-1} at 700 MHz. Previous studies provided an apparent K_d of about 10^{-5} M for the α/β-DG complex (*23*). From these data we can figure out a k_{on} that would be lower than the association rate expected for a process limited by diffusion. It can be assumed that the recognition between α- and β-DG requires overcoming a high activation barrier, but that the complex formed would be quite stable and it would dissociate slowly. This implies that β-DG binds α-DG with high selectivity but with moderate affinity. This is in line with the data suggesting that DG might play a role in signal transduction, as the high selectivity would correctly address the signal to a specific target molecule, whereas the moderate affinity would favor the interaction with different partners (*45*, *46*).

Model for the β-DG Ectodomain. β-DG(654−750) is able to bind the recombinant C-terminus of α-DG, alone and as fusion protein with thioredoxin, as well as the native α-DG *in vitro* (*23*, *43*), demonstrating that the glycosylation is not essential for the interaction between the two DG subunits. This important point must be taken into account in order to support a model for the structural organization of β-DG ectodomain that would be valid also in living cells or tissues.

We propose a model in which the N-terminus of the β-DG ectodomain, characterized by a restricted mobility and containing the binding epitope for α-DG, is bridged to its transmembrane peptide by a highly flexible linker domain (Figure 7) that allows it to move and to bind more components (*46*). The ectodomain of β-DG maintains its extended conformation also when bound to the α-subunit, similarly to what is observed in the complex between a short peptide spanning the C-terminal intracellular region of β-DG and dystrophin (*20*). This is in line with the observation that many intrinsically unstructured proteins preserve their extended conformation also in the presence of their target proteins (*47*).

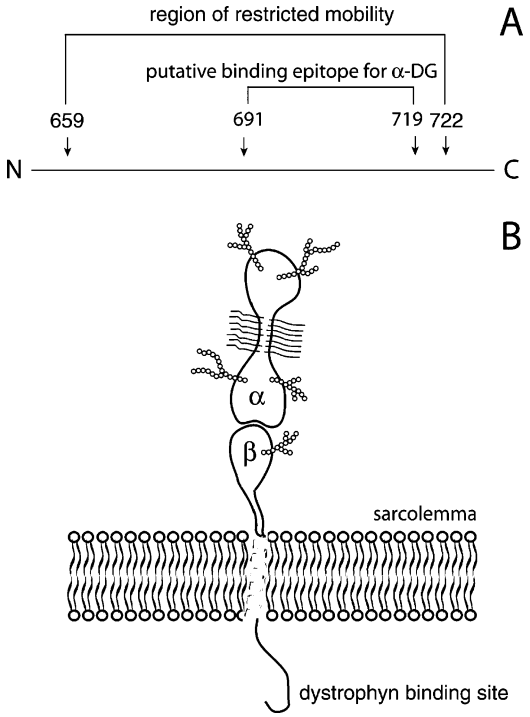

FIGURE 7: Model for β-DG(654−750). The region of reduced mobility and the putative binding epitope for α-DG are marked in panel A. The proposed simplified model for β-DG with a short flexible link between the transmembrane region and the more compact N-terminal domain, containing the binding epitope for α-DG, is shown in panel B.

The importance of the biological role of natively unfolded proteins or protein domains is becoming increasingly evident (*45*, *47*). It has been estimated that a large proportion of the proteins in living cells are natively disordered (up to 33% in bacterial and up to 66% in eukaryotic cells) (*48*). This fact raises the important question of why evolution has selected such a large fraction of proteins in a disordered conformation. It has been proposed that if the disordered proteins were more structured, cell crowding would be a severe problem leading to a dangerous increase in cell size of 15−30% (*48*). Structural characterization of natively unfolded proteins are thus of central importance in order to enlarge our understanding about the relationship between protein folding and function *in vivo*, and β-DG represents a significant example. The structural plasticity of β-DG is likely to increase the overall conformational freedom of the DG complex, helping to sustain the high mechanical forces developed upon muscle activity, and also to favor the interaction of α-DG with its multiple binding partners (*25*).

The identification of the region of β-DG in which the α-DG binding epitope is located represents a remarkable advancement in the knowledge of the molecular interface between DG subunits and it could be a crucial step to achieve a rational design either for efficient therapies for muscular dystrophies or to prevent those infections depending on DG targeting (*49*, *50*).

ACKNOWLEDGMENT

We thank Dr. Maria Giulia Bigotti (Bristol) for helpful discussions and Ernesto Pavoni (Rome) and Fabio Bertocchi (Rome) for their technical assistance.

REFERENCES

1. Henry, M. D., and Campbell, K. P. (1999) Dystroglycan inside and out, *Curr. Opin. Cell Biol. 11*, 602−607.
2. Winder, S. J. (2001) The complexities of dystroglycan, *Trends Biochem. Sci. 26*, 118−124.
3. Cartaud, A., Coutant, S., Petrucci, T. C., and Cartaud, J. (1998) Evidence for in situ and in vitro association between β-dystroglycan and the subsynaptic 43K rapsyn protein. Consequence for acetylcholine receptor clustering at the synapse, *J. Biol. Chem. 273*, 11321−11326.
4. Jung, D., Yang, B., Meyer, J., Chamberlain, J. S., and Campbell, K. P. (1995) Identification and characterization of the dystrophin anchoring site on β-dystroglycan, *J. Biol. Chem. 270*, 27305−27310.
5. Yang, B., Jung, D., Motto, D., Meyer, J., Koretzky, G., and Campbell, K. P. (1995) SH3 domain-mediated interaction of dystroglycan and Grb2, *J. Biol. Chem. 270*, 11711−11714.
6. Ervasti, J. M., and Campbell, K. P. (1993) A role for the dystrophin−glycoprotein complex as a transmembrane linker between laminin and actin, *J. Cell Biol. 122*, 809−823.
7. Durbeej, M., Henry, M. D., Ferletta, M., Campbell, K. P., and Ekblom, P. (1998) Distribution of dystroglycan in normal adult mouse tissues, *J. Histochem. Cytochem. 46*, 449−457.
8. Durbeej, M., and Campbell, K. P. (2002) Muscular dystrophies involving the dystrophin−glycoprotein complex: an overview of current mouse models, *Curr. Opin. Genet. Dev. 12*, 349−387.
9. Williamson, R. A., Henry, M. D., Daniels, K. J., Hrstka, R. F., Lee, J. C., Sunada, Y., Ibraghimov-Beskrovnaya, O., and Campbell, K. P. (1997) Dystroglycan is essential for early embryonic development: disruption of Reichert's membrane in Dag1-null mice, *Hum. Mol. Genet. 6*, 831−841.
10. James, M., Nuttall, A., Ilsley, J. L., Ottersbach, K., Tinsley, J. M., Sudol, M., and Winder, S. J. (2000) Adhesion-dependent tyrosine phosphorylation of β-dystroglycan regulates its interaction with utrophin, *J. Cell Sci. 113*, 1717−1726.
11. Sotgia, F., Lee, H., Bedford, M. T., Petrucci, T., Sudol, M., and Lisanti, M. P. (2001) Tyrosine phosphorylation of β-dystroglycan at its WW domain binding motif, PPxY, recruits SH2 domain containing proteins, *Biochemistry 40*, 14585−14592.
12. Brancaccio, A., Schulthess, T., Gesemann, M., and Engel, J. (1995) Electron microscopic evidence for a mucin-like region in chick muscle α-dystroglycan, *FEBS Lett. 368*, 139−142.
13. Brancaccio, A., Schulthess, T., Gesemann, M., and Engel, J. (1997) The N-terminal region of α-dystroglycan is an autonomous globular domain, *Eur. J. Biochem. 246*, 166−172.
14. Bozic, D., Engel, J., and Brancaccio, A. (1998) Sequence analysis suggests the presence of an IG-like domain in the N-terminal region of α-dystroglycan which was crystallized after mutation of a protease susceptible site (Arg168 → His), *Matrix Biol. 17*, 495−500.
15. Rudenko, G., Hohenester, E., and Muller, Y. A. (2001) LG/LNS domains: multiple functions−one business end? *Trends Biochem. Sci. 26*, 363−368.
16. Hohenester, E., and Engel, J. (2002) Domain structure and organisation in extracellular matrix proteins, *Matrix Biol. 21*, 115−128.
17. Talts, J. F., Andac, Z., Gohring, W., Brancaccio, A., and Timpl, R. (1999) Binding of the G domains of laminin α1 and α2 chains and perlecan to heparin, sulfatides, α-dystroglycan and several extracellular matrix proteins, *EMBO J. 18*, 863−870.
18. Tisi, D., Talts, J. F., Timpl, R., and Hohenester, E. (2000) Structure of the C-terminal laminin G-like domain pair of the laminin α2 chain harbouring binding sites for α-dystroglycan and heparin, *EMBO J. 19*, 1432−1440.
19. Rentschler, S., Linn, H., Deininger, K., Bedford, M. T., Espanel, X., and Sudol, M. (1999) The WW domain of dystrophin requires EF-hands region to interact with β-dystroglycan, *Biol. Chem. 380*, 431−442.
20. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M., and Eck, M. J. (2000) Structure of a WW domain containing fragment of dystrophin in complex with β-dystroglycan, *Nat. Struct. Biol. 7*, 634−638.
21. Ibraghimov-Beskrovnaya, O., Ervasti, J. M., Leveille, C. J., Slaughter, C. A., Sernett, S. W., and Campbell, K. P. (1992) Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix, *Nature 355*, 696−702.

22. Losasso, C., Di Tommaso, F., Sgambato, A., Ardito, R., Cittadini, A., Giardina, B., Petrucci, T. C., and Brancaccio, A. (2000) Anomalous dystroglycan in carcinoma cell lines, *FEBS Lett. 484*, 194−198.
23. Sciandra, F., Schneider, M., Giardina, B., Baumgartner, S., Petrucci, T. C., and Brancaccio, A. (2001) Identification of the β-dystroglycan binding epitope within the C-terminal region of α-dystroglycan, *Eur. J. Biochem. 268*, 4590−4597.
24. Bozzi, M., Veglia, G., Paci M., Sciandra, F., Giardina, B., and Brancaccio, A. (2001) A synthetic peptide corresponding to the 550−585 region of α-dystroglycan binds β-dystroglycan as revealed by NMR spectroscopy, *FEBS Lett. 499*, 210−214.
25. Boffi, A., Bozzi, M., Sciandra, F., Woellner, C., Bigotti, M. G., Ilari, A., and Brancaccio, A. (2001) Plasticity of secondary structure in the N-terminal region of β-dystroglycan, *Biochim. Biophys. Acta 1546*, 114−121.
26. Kammerer, R. A., Schulthess, T., Landwehr, R., Lustig, A., Fischer, D., and Engel, J. (1998) Tenascin-C hexabrachion assembly is a sequential two-step process initiated by coiled-coil α-helices, *J. Biol. Chem. 273*, 10602−10608.
27. Marley J., Lu M., and Bracken, C. (2001) A method for efficient isotopic labeling of recombinant proteins, *J. Biomol. NMR 20*, 71−75.
28. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins, *J. Magn. Reson. 89*, 496−514.
29. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein, *J. Magn. Reson. 96*, 432−440.
30. Grezsiek, S., and Bax, A. (1993) Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins, *J. Biomol. NMR 3*, 185−204.
31. Panchal, S. C., Bhavesh, N. S., and Hosur, R. V. (2001) Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins, *J. Biomol. NMR 20*, 135−147.
32. Bhavesh, N. S., Panchal, S. C., and Hosur, R. V. (2001) An efficient high-throughput resonance assignment procedure for structural genomics and protein folding research by NMR, *Biochemistry 40*, 14727−14735.
33. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1989) Overcoming the overlap problem in the assignment of proton NMR spectra of larger proteins by use of three-dimensional heteronuclear proton−nitrogen-15 Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1β, *Biochemistry 28*, 6150−6156.
34. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins, *J. Am. Chem. Soc. 111*, 1515−1517.
35. Grzesiek, S., and Bax, A. (1993) The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements, *J. Am. Chem. Soc. 115*, 12593−12594.
36. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) Measurement of HN−H alpha J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods, *J. Biomol. NMR 4*, 871−878.
37. 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.
38. 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.
39. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
40. Holt, K. H., Crosbie, R. H., Venzke, D. P., and Campbell, K. P. (2000) Biosynthesis of dystroglycan: processing of a precursor propeptide, *FEBS Lett. 468*, 79−83.
41. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a Tool for Structure Determination, *Methods Enzymol.: Nucl. Magn. Reson. C 239*, 363−392.
42. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.
43. Di Stasio, E., Sciandra, F., Maras, B., Di Tommaso, F., Petrucci, T. C., Giardina, B., and Brancaccio, A. (1999) Structural and functional analysis of the N-terminal extracellular region of β-dystroglycan, *Biochem. Biophys. Res. Commun. 266*, 274−27.
44. Lian, L.-Y., and Roberts, G. C. K. (1993) Effects of chemical exchange on NMR spectra, in *NMR of macromolecules. A practical approach*, Roberts, G. C. K., Ed., Oxford University Press, Oxford, U.K., pp153−182.
45. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure−function paradigm, *J. Mol. Biol. 293*, 321−331.
46. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradovic, Z. (2002) Intrinsic disorder and protein function, *Biochemistry 41*, 6573−6582.
47. Tompa, P. (2002) Intrinsically unstructured proteins, *Trends Biochem. Sci. 27*, 527−533.
48. Gunasekaran, K., Tsai, C. J., Kumar, S., Zanuy, D., and Nussinov, R. (2003) Extended disordered proteins: targeting function with less scaffold, *Trends Biochem. Sci. 28*, 81−85.
49. Cao, W., Henry, M. D., Borrow, P., Yamada, H., Elder, J. H., Ravkov, E. V., Nichol, S. T., Compans, R. W., Campbell, K. P., and Oldstone, M. B. (1998) Identification of α-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus, *Science 282*, 2079−2081.
50. Rambukkana, A., Yamada, H., Zanazzi, G., Mathus, T., Salzer, J. L., Yurchenco, P. D., Campbell, K. P., and Fischetti, V. A. (1998) Role of α-dystroglycan as a Schwann cell receptor for *Mycobacterium leprae*, *Science 282*, 2076−2079.

BI034867W