Biophysical Characterization of the Unstructured Cytoplasmic Domain of the Human Neuronal Adhesion Protein Neuroligin 3

Aviv Paz,*† Tzviya Zeev-Ben-Mordehai,*† Martin Lundqvist,¶ Eilon Sherman,‡ Efstratios Mylonas,‖ Lev Weiner,§ Gilad Haran,‡ Dmitri I. Svergun,‖** Frans A. A. Mulder,¶ Joel L. Sussman,* and Israel Silman†

*Departments of Structural Biology, †Neurobiology, ‡Chemical Physics and §Chemical Research Support, Weizmann Institute of Science, Rehovot, Israel; ¶Department of Biophysical Chemistry, Groningen University, Groningen, The Netherlands; ‖European Molecular Biology Laboratory, Hamburg, Germany; and **Institute of Crystallography, Russian Academy of Sciences, Moscow, Russia

ABSTRACT Cholinesterase-like adhesion molecules (CLAMs) are a family of neuronal cell adhesion molecules with important roles in synaptogenesis, and in maintaining structural and functional integrity of the nervous system. Our earlier study on the cytoplasmic domain of one of these CLAMs, the *Drosophila* protein, gliotactin, showed that it is intrinsically unstructured in vitro. Bioinformatic analysis suggested that the cytoplasmic domains of other CLAMs are also intrinsically unstructured, even though they bear no sequence homology to each other or to any known protein. In this study, we overexpress and purify the cytoplasmic domain of human neuroligin 3, notwithstanding its high sensitivity to the *Escherichia coli* endogenous proteases that cause its rapid degradation. Using bioinformatic analysis, sensitivity to proteases, size exclusion chromatography, fluorescence correlation spectroscopy, analytical ultracentrifugation, small angle x-ray scattering, circular dichroism, electron spin resonance, and nuclear magnetic resonance, we show that the cytoplasmic domain of human neuroligin 3 is intrinsically unstructured. However, several of these techniques indicate that it is not fully extended, but becomes significantly more extended under denaturing conditions.

INTRODUCTION

The neuroligins (NLs) are a set of neuronal adhesion proteins belonging to the family of cholinesterase-like adhesion molecules (CLAMs). All CLAMs contain an extracellular domain that bears high sequence homology to acetylcholinesterase (AChE), and share the same conserved intrachain disulfides. However, they all lack one or more of the members of the catalytic triad of AChE, and are thus believed to be catalytically inactive (1). To date, three transmembrane CLAMs have been identified: the mammalian and invertebrate NLs, and the insect proteins, gliotactin (Gli) and neurotactin (Nrt) (2) (Fig. 1). All three are single-pass transmembrane proteins whose cytoplasmic domains contain ~120–320 amino-acid residues, and bear no sequence similarity to any known protein, or even to each other.

NLs are expressed in both neurons and glia in the nervous system, as well as in a variety of other tissues (3–5). At synapses, they are associated with the postsynaptic membrane, and have been shown to bind through their AChE-like domains to the extracellular domains of both α- and β-neurexin, which are also synaptic adhesion proteins (3,6,7). The cytoplasmic domains of the NLs, as well as that of Gli, that is, NL-cyt and Gli-cyt, contain PDZ recognition motifs at their C-termini (2). Indeed, the intracellular partners of the NLs that have been reported are PDZ domain proteins that are known to be associated with the postsynaptic density, such as PSD95 and S-SCAM (8,9). Both PSD95 and S-SCAM have been shown to induce clustering of potassium channels and NMDA receptors (8,10). Whereas it has been shown that NL1 is localized at the postsynaptic membranes of excitatory glutamatergic synapses (11), NL2 is exclusively localized at inhibitory synapses (12). Scheiffele et al., who engineered nonneuronal cells to express NLs, demonstrated their functional importance; the nonneuronal cells induced morphological and functional presynaptic differentiation upon contacting axons in vitro. Moreover, this activity was inhibited by addition of a soluble form of β-neurexin (13). These findings suggest that NLs are involved in formation and remodeling of CNS synapses. Recently, it was shown that a 32 amino-acid sequence in the cytoplasmic domain of NL1 is the dendritic targeting motif of this protein (14,15).

Mutations in the extracellular region of either NL3 or NL4 have been described in certain families with a high frequency of autism spectrum disorders or of X-linked mental retardation (16–18). It is plausible that the mutations lead to subtle changes in neuronal circuitry, since certain such mutations have been shown to result in partial intracellular retention of the mutant proteins (19,20). Recently the R451C mutation in NL3 was introduced into mice resulting in impaired social interactions and enhanced spatial learning abilities that, unexpectedly, were accompanied by increased inhibitory synaptic transmission (21).

We earlier showed that the cytoplasmic domain of Gli, Gli-cyt, is natively unfolded (22). Natively unfolded proteins, also known as intrinsically unstructured proteins (IUPs), are flexible, extended, and have little or no ordered secondary structure in vitro in the absence of partners (23). These proteins question the known structure-function paradigm in proteins; despite their lack of a well-defined three-dimensional structure, they are indeed functional (24,25). A com-

Submitted December 27, 2007, and accepted for publication March 18, 2008.

Address reprint requests to Israel Silman, Tel.: 972-8-934-3649; E-mail: israel.silman@weizmann.ac.il.

Editor: Doug Barrick.

0006-3495/08/08/1928/17 $2.00

doi: 10.1529/biophysj.107.126995

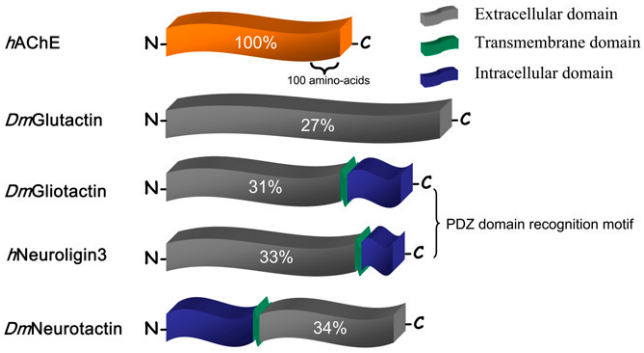

FIGURE 1 Domain architecture of the CLAMs. All CLAMs contain a ChE-domain (*shaded*), which localizes to the extracellular compartment either as a secreted protein or as the extracellular domain of a single-pass transmembrane protein. Both Gli and NL are type I membrane proteins that bear a C-terminal PDZ recognition motif, while Nrt is a type II membrane protein. The percentage identity to human AChE is shown in each extracellular domain. *Dm* stands for *Drosophila melanogaster* and *h* stands for human. The *h*AChE scale bar represents 100 amino acids.

bination of low sequence complexity, low overall hydrophobicity, and relatively high net charge are important prerequisites for a protein to be devoid of compact structure (23,26,27). Functionally, IUPs are seen to be involved in protein-protein interaction networks and protein-DNA complexes, and participate in signal transduction, transcriptional activation, nucleic acid recognition, and cell cycle regulation (25). Protein disorder has been suggested to increase the rates of protein association (28). Moreover, evidence has been presented that IUPs display the capacity to interact with multiple protein partners, and thus to function as hub proteins in interactomes (for a recent review see (29)), and this binding promiscuity has been suggested to increase network complexity (30). Recently, a number of in silico predictors for protein disorder have been developed which reveal a much higher frequency of IUPs, or of long unfolded stretches, in eukaryotes than in prokaryotes (31). Some of them appear to be associated with scaffolding elements in complex structures such as synapses, which are absent in prokaryotes.

The high flexibility of the IUPs renders them highly sensitive to proteases, a feature that has been taken advantage of in their characterization (32). Recently, we have described an assay based on the susceptibility of IUPs to 20S proteasomal degradation machinery as a simple, rapid, and reliable method of operationally defining IUPs and unstructured sequences (33).

Although IUPs and denatured proteins lack stable secondary structure elements, the left-handed poly-L-proline II helix (PPII) has been found to be prominent in both (34). The main chain dihedral angles in PPII are $\phi,\psi \approx -78°, +146°$, corresponding to three residues per turn. Its helical rise per residue is 3.1 Å, compared to 1.5 Å for the α-helix, and it is thus much more extended (35). While proline residues are common in PPII helices, their presence is not obligatory (36). Although proline possesses the highest intrinsic propensity to adopt the PPII conformation, glutamine, alanine, and glycine also display high propensities (37). Already in 1968, Tiffany and Krimm (38) noted a resemblance between the circular dichroism (CD) spectra of denatured proteins and those of PPII and collagen, and suggested that the unfolded state produced by denaturation is, in fact, composed of short stretches (4–7 residues) of PPII connected by sharp bends. Zagrovic et al. (39) have suggested that in the unfolded state, individual residues are likely to sample a broad basin around the PPII region of the Ramachandran map, but do not form regular PPII helices due to the occurrence of rapid changes in their ϕ- and ψ-angles (see also (40)).

It was already pointed out by Steinberg et al. (41) that PPII itself is considerably more compact than would be expected for a fully extended rod, and this has been borne out by more recent experimental and theoretical studies (39,42). Thus, it may be expected that IUPs, in general, will be substantially more compact than denatured proteins, as was shown to be the case for Gli-cyt (22) and in other cases (43).

In the following, we describe the overexpression, purification, and physicochemical characterization of the cytoplasmic domain of human neuroligin 3 (hNL3-cyt), and provide evidence that it belongs to the family of IUPs, with a conformation that is not, however, fully extended.

MATERIALS AND METHODS

Bioinformatic procedures

Disorder in hNL3-cyt was assessed using Foldindex (44), the VL-XT algorithm in PONDR (45,46), and IUPred (47). Secondary structure prediction made use of the program PROF (48).

The domain mean net charge $\langle R \rangle$, i.e., the net charge at pH 7.0 divided by the number of amino acids, and the domain mean hydrophobicity $\langle H \rangle$, i.e., the sum of the normalized hydrophobicities of all residues divided by the number of residues in a given polypeptide sequence, were calculated as described by Uversky et al. (26). The boundary mean hydrophobicity $\langle H \rangle_b$, namely, the mean hydrophobicity below which a polypeptide with a given mean net charge $\langle R \rangle$ is predicted to be unfolded, was calculated according to Uversky (23) and Oldfield et al. (49).

Cloning of hNL3-cyt

DNA encoding the intracellular domain of hNL3-cyt (amino acids 731–848) was amplified by PCR from a cDNA template of a full-length human NL3 clone (clone ID 5258980) provided by Dr. Peter Scheiffele (Department of Cell Biology, College of Physicians and Surgeons, Columbia University, New York) using primers with *Nde*I and *Not*I restriction sites. The PCR product was purified, restricted with *Nde*I and *Not*I, and ligated between the *Nde*I and *Not*I sites of the pET28a (Novagen, Madison, WI) expression vector. The sequence of the hNL3-cyt coding region was confirmed by nucleotide sequencing.

Protein expression

Escherichia coli BLR(DE3) cells transformed with hNL3-cyt were grown to saturation overnight at 37°C with shaking in LB medium containing kanamycin (30 μg/ml) and tetracycline (12.5 μg/ml). An aliquot from the over-

night culture was diluted 1:100 in fresh LB medium containing kanamycin (30 μg/ml) and tetracycline (12.5 μg/ml), and grown at 37°C, with shaking, to $OD_{600} = 0.6$, at which time hNL3-cyt expression was induced by addition of isopropyl-β-D-galactoside to a final concentration of 1 mM. After a further 3 h incubation the cells were harvested by centrifugation (10,000 g, 10 min, 4°C), yielding ~2.2 g/L of cell pellet, and stored at −80°C. hNL3-cyt labeled with ^{13}C and ^{15}N was produced in transformed *E. coli* BLR(DE3) cells grown to saturation overnight, at 37°C with shaking, in LB medium containing kanamycin (30 μg/ml) and tetracycline (12.5 μg/ml). An aliquot from the overnight culture was diluted 1:100 in fresh M9 medium containing kanamycin (30 μg/ml) and tetracycline (12.5 μg/ml), 6 g/l Na_2HPO_4, 3 g/l KH_2PO_4, 1 g/l ^{15}N NH_4Cl (Cambridge Isotope Laboratories, Andover, MA), 0.5 g/l NaCl, 1 ml/l 0.1M $CaCl_2$, 2ml/l 1M $MgSO_4$, 2 g/l ^{13}C D-glucose (Cambridge Isotope Laboratories), 5 mg/l thiamine-HCl, 1ml/l trace elements (1 mM Cu[OAc], $ZnSO_4$, $MnSO_4$, $CoSO_4$, Ni$[SO_4]_2$, Na_2MoO_4), and 10 ml/l RPMI 1640 vitamin cocktail (Sigma-Aldrich, St. Louis, MO). After incubation with shaking at 225 RPM for 9 h at 37°C, hNL-3cyt expression was induced with 1 mM isopropyl-β-D-galactoside, and incubation continued overnight until harvesting.

Protein purification

The bacterial cell pellet was resuspended in 30 ml 8 M urea/100 mM NaH_2PO_4/10 mM Tris, pH 8.0 (buffer A), clarified by centrifugation (12,000 g, 30 min, 4°C), and filtered through a 0.22 μm Stericup filtration device (Millipore, Billerica, MA). The filtered protein solution was loaded onto a 1 ml HisTrap column (GE Healthcare, Waukesha, WI), preequilibrated in buffer A, and eluted using a linear pH 8.0–4.5 gradient of buffer A. Eluted fractions were analyzed by SDS-PAGE, and fractions containing hNL3-cyt were pooled and loaded onto a SP FF cation exchange column (GE Healthcare). The protein was eluted using a linear gradient of 0–1 M NaCl in 100 mM MES, pH 6.0. Pure fractions were pooled and concentrated, with concomitant change of the buffer to 1 mM EDTA/5 mM dithiothreitol (DTT)/250 mM NaCl/100 mM Tris, pH 8.5, using a Vivaspin concentration device (3000 MWCO, Sartorius, Goettingen, Germany), and further purified through a HiLoad 16/60 Superdex 75 pg column (GE Healthcare) preequilibrated with the same buffer.

Protease digestion

Samples of hNL3-cyt (25 μg in a volume of 200 μl) and of purified native *Torpedo californica* AChE (*Tc*AChE) (50 μg in a volume of 200 μl), in 50 mM NaCl/100 mM Tris, pH 8.0, were digested at room-temperature by addition of proteinase K (Sigma) to a final concentration of 50 ng/ml for hNL3-cyt, and 100 ng/ml for *Tc*AChE. Aliquots were removed at appropriate times, mixed with Laemmli sample buffer containing 1 mM phenylmethylsulfonyl fluoride (Sigma), and boiled for 5 min to arrest proteolysis. Samples were loaded onto a 15% SDS-PAGE gel, electrophoresed, and stained with GelCode.

Determination of the hydrodynamic radius by size exclusion chromatography (SEC)

A 50-μl aliquot of a 10 mg/ml solution of purified hNL3-cyt, in 1 mM EDTA/5 mM DTT/250 mM NaCl/100 mM Tris, pH 8.5, was injected onto an analytical Superdex 75 HR 10/30 column using an AKTA fast-performance liquid chromatography system (GE Healthcare) under both nondenaturing (1 mM EDTA/5 mM DTT/250 mM NaCl/100 mM Tris, pH 8.5) and denaturing (8 M urea/10 mM Tris/100 mM NaH_2PO_4, pH 7.0) conditions. The column was calibrated with molecular weight standards under both nondenaturing and denaturing conditions, and a calibration curve was created for the nondenaturing run using Eq. 1 and a plot of K_{av} versus log molecular weight (see Supplementary Material, Fig. S1, in Data S1).

$$K_{av} = \frac{V_e - V_0}{V_t - V_0}, \tag{1}$$

where V_e is the protein elution volume, V_0 is the column's exclusion volume, determined with dextran blue, and V_t is the total bed volume.

Fluorescence correlation spectroscopy (FCS)

Fluorescence correlation spectroscopy (FCS) measurements were performed using 20–40 nM hNL3-cyt that had been labeled on Cys^{775}, the only cysteine residue in the cytoplasmic domain, with ATTO 488 maleimide (ATTO-TEC, Siegen, Germany) according to the manufacturer's protocol. The final step of removal of free label utilized a 1 ml HiTrap desalting column (GE Healthcare) in 50 mM NaCl/50 mM NaH_2PO_4, pH 7.0, on an AKTA FPLC system. Measurements were performed using a home-built confocal microscope, with a water immersion UplanApo 60× NA 1.2 objective (Olympus, Tokyo, Japan). The setup consisted of a 488 nm Ar^+ ion laser (35 LAP 431, Melles Griot, Carlsbad, CA), focused through the objective, at a power level of 30 μW. Emitted fluorescence was collected through the objective, filtered by a 500 DCLP dichroic mirror (Chroma, Brattleboro, VT) and an HQ 500 LP long-pass interference filter (Chroma), and then focused by the microscope tube lens onto a 50 μm pinhole. The beam was then split by a nonpolarizing beam splitter, and focused onto two identical SPCM-AQR-15 single-photon counting avalanche photodiodes (Perkin Elmer, Fremont, CA). A Flex02-12D/C hardware correlator (Correlator.com, Bridgewater, NJ) was used for data collection and for generation of the FCS curves. The correlation functions were fitted to a model of diffusion through a three-dimensional Gaussian observation volume which takes into account triplet state photodynamics (Eq. 2) (50).

$$G(T) = \frac{1}{N}\frac{(1 + Ke^{-kt})}{\left(1 + \frac{t}{\tau}\right)\sqrt{1 + \frac{t}{\tau(\omega_z/\omega_{xy})^2}}}, \tag{2}$$

where N is the average number of molecules in the observation volume; τ is the mean diffusion time of a molecule through the observation volume; ω_{xy} and ω_z are, respectively, the dimensions of the Gaussian beam waist perpendicular and parallel to the direction of light propagation, K is the amplitude of the populated triplet state, and k is the effective rate of this process. The diffusion coefficient is related to the above parameters through the relation $D = \omega_{xy}{}^2/4\tau$. The system was calibrated using rhodamine 6G before the measurements on the labeled protein (Fig. S2 in Data S1).

Analytical ultracentrifugation

Sedimentation velocity experiments were carried out on an Optima XL-A analytical ultracentrifuge (Beckman Coulter, Fullerton, CA) in an An-50Ti rotor. The hNL3-cyt sample was dialyzed in phosphate-buffered saline (PBS) adjusted to pH 8.0, and 400-μl aliquots of the hNL3-cyt sample and of PBS buffer were injected into a double-sector cell with a 12 mm Epon centerpiece and quartz windows.

Ultracentrifugation was performed at 50,000 rpm, 20°C (after a 3-h temperature equilibration period), at a protein concentration of 0.3 mg/ml. The absorbance wavelength was 230 nm. A radial step size of 0.003 was used for scanning, and >500 scans were collected, at which point sedimentation was complete.

Data from the first 500 scans were analyzed using the SEDFIT program (51), either with a diffusion deconvoluted sedimentation coefficient $c(s)$ distribution model, with a continuous molar mass distribution $c(M)$, or with the two-dimensional $c(s, f_r)$ approach, which is more suitable for dynamic ensembles of molecules for which the assumption of a continuous f_r in the $c(s)$ distribution is unclear (52). For both approaches the solvent density, the partial specific volume, and the Stokes radius (R_H) were obtained using the SEDNTERP program (53).

Synchrotron x-ray scattering (SAXS) data collection and processing

Synchrotron x-ray scattering (SAXS) data for hNL3-cyt were collected on the X33 beamline, at the EMBL Hamburg Outstation (54), using a MAR345 image plate detector. The scattering patterns were measured with a 2-min exposure time for multiple solute concentrations ranging from 3 to 1.5 mg/ml. DTT was added to a final concentration of 7 mM before measurement. To check for radiation damage, two 2-min exposures were compared, and no apparent damage was detected. Using the sample-detector distance of 2.7 m, a range of momentum transfer of $0.09 < s < 0.5$ Å^{-1} was covered ($s = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle, and $\lambda = 1.5$ Å is the x-ray wavelength). The data were processed using standard procedures, and extrapolated to infinite dilution using the PRIMUS program package (55). The forward scattering, $I(0)$, and the radius of gyration, R_g, were evaluated using the Guinier approximation (56), assuming that at very small angles ($s < 1.3/R_g$) the intensity is represented as $I(s) = I(0)\exp(-(sR_g)^2/3$. Values of $I(0)$ and R_g, as well as the maximum dimension, D_{max}, and the interatomic distance distribution functions, ($p(r)$), were computed using the GNOM indirect transform package (57).

The molecular mass of hNL3-cyt was evaluated by comparison of the forward scattering with that for a reference solution of bovine serum albumin (BSA) (66 kDa).

To quantitatively characterize the range of conformations adopted by hNL3-cyt in solution, the program EOM was used (58). Given a pool of randomly generated structures (with a R_g distribution similar to that of a random coil), EOM employs a genetic algorithm to select an ensemble of (typically 20–50) models, average scattering from which fits the experimental pattern. The properties of the selected ensembles obtained in multiple EOM runs provide information about the flexibility of the protein in solution.

Circular dichroism (CD) spectroscopy

CD spectra were recorded on a Chirascan circular dichroism spectrometer (Applied Photophysics, Surrey, UK). Far-UV CD spectra were recorded in a 0.1 cm path-length cuvette at a protein concentration of 12 μM (0.19 mg/ml). All CD spectra were recorded in 10 mM potassium phosphate, pH 7.0, except for the SDS-containing samples, which were measured in 10 mM Tris, pH 7.9. The step size was 0.5 nm, and the averaging time 3 s. Each trace displayed represents the average of three spectra, and is corrected for the contribution of the buffer. Denaturation experiments utilized high purity guanidinium hydrochloride (Gdn.HCl) (Pierce, Rockford, IL) and trifluoroethanol (TFE) (Sigma). The TFE was handled using glass tips and glass tubes.

1,2-Dioleoyl-*sn*-glycero-3-[phospho-L-serine] (DOPS) liposomes (Avanti Polar Lipids, Alabaster, AL) were prepared by evaporation, in a glass test tube, of a 100-μl aliquot of a 20 mg/ml chloroform solution, followed by resuspension in 500 μl of 10 mM KPO_4, pH 8.0, followed by vortexing for 1 min, and sonication for 3 min at 188 W in a Vibracell Sonicator (Sonics, Newtown, CT), to yield a 5 mM stock.

Electron spin resonance (ESR) measurements

The spin-labeled thiol reagent (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl)-methathiosulfinate (MTSSL, Toronto Research Chemicals, Toronto, Ontario, Canada), was used to covalently label the single free cysteine residue, Cys^{775}. MTSSL was dissolved in ethanol, and mixed at a 20:1 molar ratio with 200 μM hNL3-cyt in 10 mM Na_2HPO_4, pH 7.0. After incubation for 14 h at 4°C, the reaction mixture was loaded onto a 1 ml HiTrap desalting column (GE Healthcare) equilibrated with 150 mM NaCl/10 mM NaH_2PO_4, pH 7.0, on an AKTA FPLC system, to separate the labeled protein from the free reagent. All measurements were performed on an ELEXSYS 500 spectrometer (Bruker, Billerica, MA), equipped with an Euroterm ER 4113VT temperature unit (Bruker) with an accuracy of ±1 K, in a 60 μl flat cell, in a protein concentration range of 4–8 μM.

The rotational correlation time (τ_R) of weakly immobilized probes was estimated using the equation (59)

$$\tau_R = 6.5 \times 10^{-10} \Delta H_0 \left(\sqrt{\frac{h_C}{h_H} - 1} \right), \quad (3)$$

where h_C and h_H are the heights of the central and high field components, respectively, and ΔH_0 is the line-width of the central component of the ESR spectrum. The rotational motion of the spin probes is assumed to be isotropic.

NMR spectroscopy

NMR samples contained 1 mM of uniformly labeled (>95%) $^{15}N/^{13}C$-labeled protein in 90%/10% v/v H_2O/D_2O with 5 mM dithiothreitol (DTT), 20 mM potassium phosphate buffer, pH 6.0, and a trace amount of sodium azide. Experiments were performed at 25.0 ± 0.1°C on a Unity Inova 600 MHz spectrometer (Varian, Palo Alto, CA) equipped with a triple resonance probe-head and single-axis pulsed-field gradient capabilities. ^{15}N Nuclear Overhauser Effect (NOE) spectra were recorded as described by Farrow et al. (60). Two datasets were acquired; one without (reference), and one with proton saturation (steady-state NOE) during the recycle delay. In the reference experiment, the interscan delay was 8 s to allow for the complete recovery of proton and nitrogen magnetization. The saturation experiment consisted of a 5 s delay followed by 3 s of proton saturation by a train of nonselective 120° pulses spaced by 5 ms. Reference and saturation experiments were recorded as complete two-dimensional datasets accumulating eight scans per increment in 7 h. Five copies of the reference and saturation experiments were acquired, and were merged after data processing to increase the signal/noise ratio while, at the same time, avoiding the accumulation of systematic errors due to long-term spectrometer instabilities. Spectra were recorded with spectral widths of 8000 Hz, sampled over 512 complex points in the ^1H dimension (maximum evolution time 64 ms), and 1650 Hz, sampled over 200 complex points in the ^{15}N dimension (maximum evolution time 121.2 ms). Data sets were processed using NMRPipe (61). The time domain data were apodized using a shifted sine bell in both the ^1H and ^{15}N dimensions. The final size of each matrix was 1024×512 real points after zero filling and Fourier transformation, giving a digital resolution of 7.8 Hz (^1H) and 3.2 Hz (^{15}N), respectively. Complete chemical shift assignments for hNL3-cyt are not yet available.

Mass spectrometry

Determination of mass and peptide mapping, making use of matrix-assisted laser desorption/ionization time-of flight, were performed by the Weizmann Institute Mass Spectrometry Service.

Dynamic light scattering (DLS)

Dynamic light scattering (DLS) measurements were performed on a model No. 802 machine (Viscotek, Houston, TX), in 1 mM EDTA/5 mM DTT/250 mM NaCl/100 mM Tris, pH 8.5, and in 6 M urea dissolved in the same buffer, at a protein concentration of 1 mg/ml.

Calculation of radii under different conditions

Hydrodynamic radii to be anticipated for a 15.4 kDa protein in various states, e.g., the native globular state, the premolten globule, and under conditions of denaturation at high concentrations of either urea or Gdn.HCl, were calculated according to Uversky (23).

Calculation of the radius of gyration for an unfolded protein was performed using Flory's scaling law (62)

$$R_g = R_0 N^{\nu}, \quad (4)$$

where R_0 is a constant relating to persistence length, N is the number of residues, and ν is a scaling factor that depends on the solvent. Values of $R_0 = 2.08 \pm 0.12$ and $\nu = 0.598 \pm 0.029$ were obtained from (63).

Calculation of the radius of gyration of an unfolded protein was also performed according to the equation utilized by Moncoq et al. (64),

$$(R_g)^2 \approx b^2\left[\frac{y}{6} - \frac{1}{4} + \frac{1}{4y} - \frac{1}{8y^2}\right], \quad (5)$$

where $y = L/b$. The value of b, which is twice the persistence length, was assumed to be 20 Å (65). L, the chain contour length, is given by $L = naf$, where n is the number of residues, the characteristic length of one residue a is taken to be 3.78 Å, and $f = 0.95$ takes into account the constraints of the polypeptide chain (66).

Calculation of the radius of gyration for a freely jointed chain (67) was performed using

$$\langle R_g^2 \rangle = \frac{l^2 n}{6}, \quad (6)$$

where n is the number of links in a polymer and l is the link length, set to 4 nm according to the literature (41,42).

RESULTS

Bioinformatic analysis

hNL3-cyt has a low content of hydrophobic amino acids, ~16.5% aliphatic residues (Val, Leu, and Ile), and ~4% aromatic amino acids (Phe, Trp, and Tyr). It contains ~20% charged residues, and has a net charge of −1. Analysis of its sequence, making use of PONDR and Foldindex as predictors of intrinsic disorder, indicated that it contains substantial stretches of both ordered and disordered sequences, although the assignments made by the two programs differ substantially. Both assign the N-terminal 50 amino-acid sequence as disordered (Fig. 2, *A* and *B*). However, Foldindex predicts that residues 51–118 will be almost completely ordered, whereas PONDR predicts alternating folded and unfolded sequences. In contrast, IUPred, a predictor that estimates an energy resulting from interresidue interactions (47), predicts the entire hNL3-cyt sequence to be fully unstructured (Fig. 2 *C*). In support of this latter assignment, the secondary structure predictor, PROF, predicts that ~86% of the domain is in loop conformations, with only ~9% being assigned as two short α-helices, and ~4% as two β-strands (Fig. 2 *D*). Furthermore, three out of four of these motifs have low confidence levels. The overall charge and hydrophobicity values for hNL3-cyt are $\langle R \rangle = 0.008$ and $\langle H \rangle = 0.4295$. The boundary mean hydrophobicity $\langle H \rangle_{boundary}$, i.e., the mean hydrophobicity below which a polypeptide with a given mean net charge $\langle R \rangle$ will most probably be unfolded (26), is $\langle H \rangle_b = 0.407$ for hNL3-cyt. Consequently, hNL3-cyt, although actually on the folded side, lies very close to the boundary between the folded and unfolded regions of the charge-hydrophobicity phase space defined by Uversky and co-workers (Oldfield et al. (49)). If the two segments, 1–54 and 55–118, are similarly analyzed using the Uversky method, the first segment is predicted to be well within the unfolded phase-space ($\langle R \rangle = 0.0377$, $\langle H \rangle = 0.3634$, $\langle H \rangle_b = 0.418$), and the second within the folded region ($\langle R \rangle = 0.0153$ $\langle H \rangle = 0.4825$ $\langle H \rangle_b = 0.41$).

Expression and purification

SDS-PAGE performed on an induced cell pellet of *E. coli* transfected with hNL3-cyt revealed overexpression of a

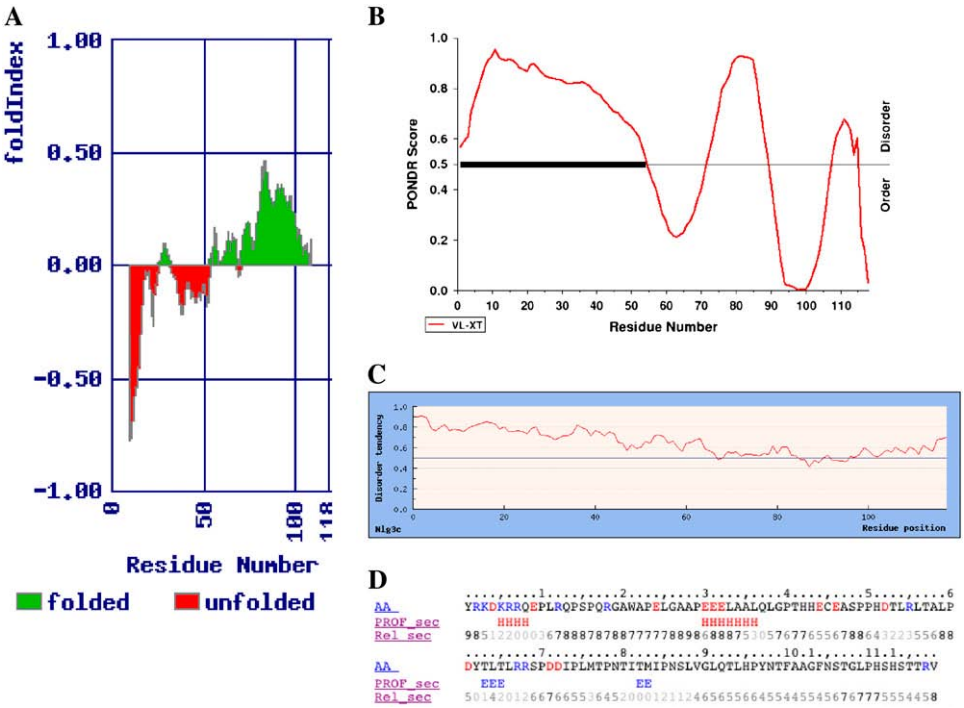

FIGURE 2 Bioinformatics analysis of hNL3-cyt. (*A*) Foldindex prediction; (*B*) PONDR prediction. Both algorithms predict that the 50 N-terminal amino acids of hNL3-cyt are disordered; (*C*) IUPred prediction for hNL3-cyt; and (*D*) PROF secondary structure prediction (in the PROF_sec line, *H* denotes helix and *E* denotes strand). Rel_sec line is the reliability index, ranging from 0 (low) to 9 (high).

~20 kDa protein. This band yielded a positive Western blot with anti-His antibodies (data not shown). MALDI/MS of the band yielded a mass of 15,367 Da, the calculated mass for the construct being 15,378 Da. Peptide mapping confirmed that the band corresponded to hNL3-cyt, with 11% coverage. Various purification protocols under nondenaturing conditions, even in the presence of a broad-spectrum protease inhibitor cocktail for the purification of His-tagged proteins (Sigma), all resulted in rapid protein degradation (Fig. 3, lane 3). Abnormal migration of the full-length hNL3-cyt domain on SDS-PAGE, as well as CD analysis of the heterogeneous hNL3-cyt solution in the far UV (Fig. S3 in Data S1), both provided early indications that hNL3-cyt is an IUP. IUPs are, in general, known to be highly sensitive to proteases (68). We therefore developed a protocol in which purification was carried out under denaturing conditions, which resulted in inactivation of the endogenous *E. coli* proteases. This approach succeeded, and allowed us to routinely obtain ~10 mg of stable purified NL3-cyt per L of medium (see Materials and Methods, and Fig. 3, lane 2). Such preparations were used in the physicochemical studies described below.

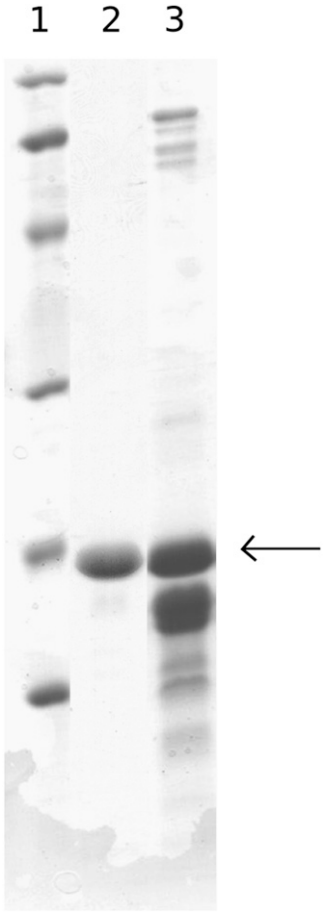

FIGURE 3 SDS-PAGE of partially purified hNL3-cyt extracted under denaturing or nondenaturing conditions. hNL3-cyt was extracted from the bacterial pellet either under denaturing conditions (8 M urea/100 mM NaH_2PO_4/10 mM Tris, pH 8.0) or nondenaturing conditions (500 mM NaCl/50 mM NaH_2PO_4/10 mM imidazole, pH 8.0, containing a broad-range protease inhibitor cocktail). In both cases, extraction was at 4°C. Both extracts underwent an initial purification step which involved absorption on a Ni-column followed by elution using a pH gradient for the protein extracted under denaturing conditions, and an imidazole gradient for the sample extracted under nondenaturing conditions (for details, see Materials and Methods). Lane 1, molecular weight markers; Lane 2, hNL3-cyt purified after extraction under denaturing conditions; and Lane 3, hNL3-cyt purified after extraction under nondenaturing conditions. Western blotting, using anti-His antibodies, reveals that many of the fast-moving species below the band of intact hNL3-cyt (*arrow*) are degradation products (not shown).

Protease sensitivity

As mentioned above, during initial attempts to purify hNL3-cyt under nondenaturing conditions, even in the presence of a broad-range protease inhibitor cocktail, the protein degraded completely over a few days at 4°C, suggesting susceptibility to trace amounts of proteases that had not been completely removed during purification. To confirm this susceptibility to proteolytic digestion, the rate of digestion of hNL3-cyt by Proteinase K was compared to that of *Tc*AChE, a globular protein that is known to be resistant to proteases in its native state (69) (Fig. 4). hNL3-cyt is, indeed, much more sensitive to Proteinase K than native *Tc*AChE.

Size exclusion chromatography

Fig. 5 shows the elution profile of hNL3-cyt under nondenaturing and denaturing conditions. Under denaturing conditions, in 8 M urea/100 mM NaH_2PO_4/10 mM Tris, pH 7.0, the 15 kDa protein eluted at a position corresponding to that to be expected for a ~80 kDa globular protein, and to a hydrodynamic radius (R_H) of ~39 Å. In the absence of urea it eluted later, at a position corresponding to that for a ~41 kDa globular protein, and to an R_H of ~28.4 Å, a radius that corresponds to a premolten globule according to the analysis of Uversky (23).

These data show that hNL3-cyt is already substantially unfolded under nondenaturing conditions, but becomes even more extended upon denaturation. Since size-exclusion chromatography (SEC) is not an ideal technique for determining the size and shape of nonglobular proteins, we carried out DLS, FCS, AUC, and SAXS measurements for hNL3-cyt. These techniques provided us with detailed information

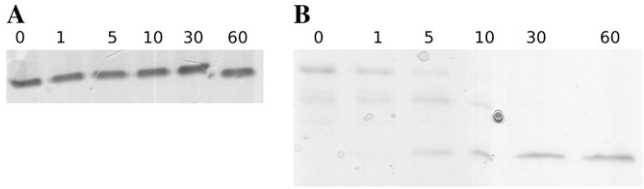

FIGURE 4 Digestion of hNL3-cyt and *Tc*AChE by proteinase K. Digestion conditions were as described under Materials and Methods. (*A* and *B*) SDS-PAGE of AChE (*A*) and hNL3-cyt (*B*) after digestion by proteinase K. Times of digestion, in minutes, are marked above each lane.

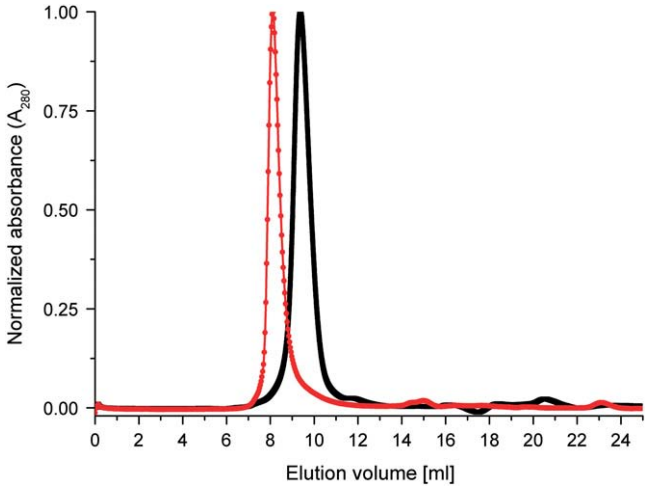

FIGURE 5 Size exclusion chromatography of hNL3-cyt. hNL3-cyt eluted at 8.1 ml under denaturing conditions (*red*), and at 9.4 ml under nondenaturing conditions (*black*). A perfect sphere with the molecular weight of hNL3-cyt would be predicted to elute at 11.85 ml.

on its dimensions, and provided further evidence for the presence of a dimer.

Fluorescence correlation spectroscopy

The correlation curve for hNL3-cyt under denaturing conditions (6 M Gdn.HCl) is shifted to longer diffusion times due to the higher viscosity of the Gdn.HCl solution and the expansion of the polypeptide chain relative to nondenaturing conditions (Fig. 6). Hydrodynamic radii of the nondenatured and denatured forms of hNL3-cyt, calculated using the Stokes-Einstein equation, were 25.0 ± 0.6 Å and 33.0 ± 0.8 Å, respectively. These values once again demonstrate that hNL3-cyt is disordered in solution, but can be further extended by denaturation.

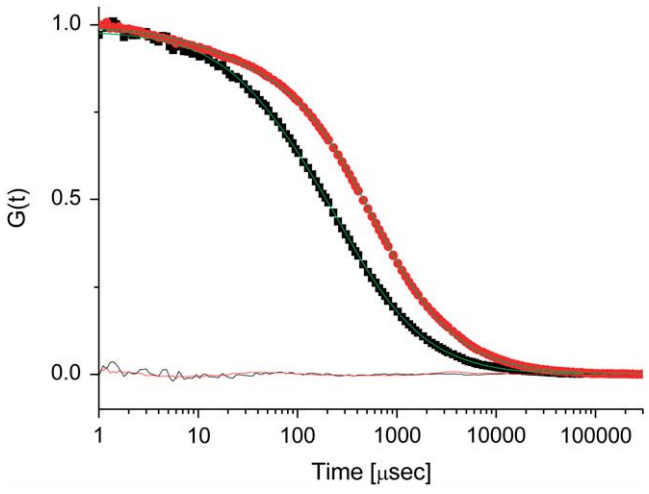

FIGURE 6 Fluorescence correlation curves for ATTO-488 labeled hNL3-cyt. Measurements in 0.01% Tween/50 mM NaCl/50 mM sodium phosphate, pH 7.0 (■). Measurements in 0.01% Tween/6M Gdn.HCl/12.5 mM NaCl/12.5 mM sodium phosphate, pH 7.0 (●). The green solid lines represent fits of the data to the model in Eq. 2. Residuals of data fits are displayed at the bottom of the figure for both the two correlation curves and for the fits.

Analytical ultracentrifugation

Analysis of hNL3-cyt fractions eluted from the preparative Superdex 75 column, using SDS-PAGE gels at high protein concentrations, and Western blots, displayed the presence of several minor oligomer bands (Fig. 7). To quantify this phenomenon, and to characterize the ensemble of different conformations that hNL3-cyt adopts in solution, we utilized analytical sedimentation velocity ultracentrifugation. The data points and fits that were used in the continuous $c(s)$ distribution are shown in Fig. 8 *A*. The $c(s)$ distribution (Fig. 8 *B*) exhibits a major peak at 1.33 S, with a shoulder at ~2 S. Using Sedfit, it can be calculated that a perfect sphere with the mass of hNL3-cyt should have an S value of 2.15. Thus, it is plausible that the 1.33 S peak corresponds to an extended monomer, and that the shoulder corresponds to a dimer, though it cannot be excluded that it may, in fact, be a collapsed monomer. To evaluate this, a $c(s,f_r)$ analysis, which uses different friction coefficients for different populations, was applied (Fig. 8 *C*). If the ~2 S population is a spherical monomer, it should have a friction coefficient that approaches a value of 1. Both the 1.33 and 2 S populations have high values of f_r, ~1.8 and 1.5, respectively, which supports the assignment of the shoulder as an extended dimer. The $c(M)$ conversion (Fig. 8 *D*) clearly shows that the principal species is a monomer with $MW = 15{,}227 \pm 300$ Da, in agreement with the calculated weight of hNL3-cyt, and $R_H = 26$ Å, calculated by use of SEDNTERP, taking the value of 1.33 S obtained in the $c(s)$ analysis.

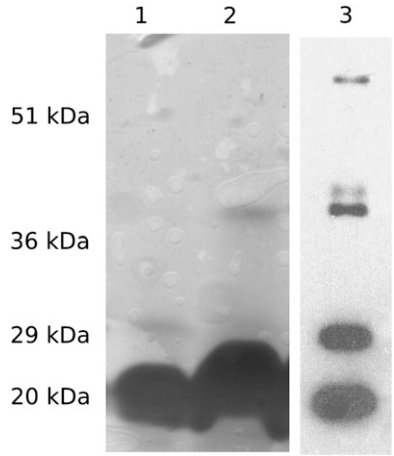

FIGURE 7 SDS-PAGE and Western-blot of hNL3-cyt after nondenaturing size exclusion chromatography. hNL3-cyt elutes in several oligomeric states, which are stable under the conditions of SDS-PAGE. Lanes 1 and 2, SDS-PAGE of hNL3-cyt loaded at 2.4 mg/ml, and 4.8 mg/ml; lane 3, Western-blot of hNL3-cyt using anti-His antibody, which displays at least three distinct oligomeric states.

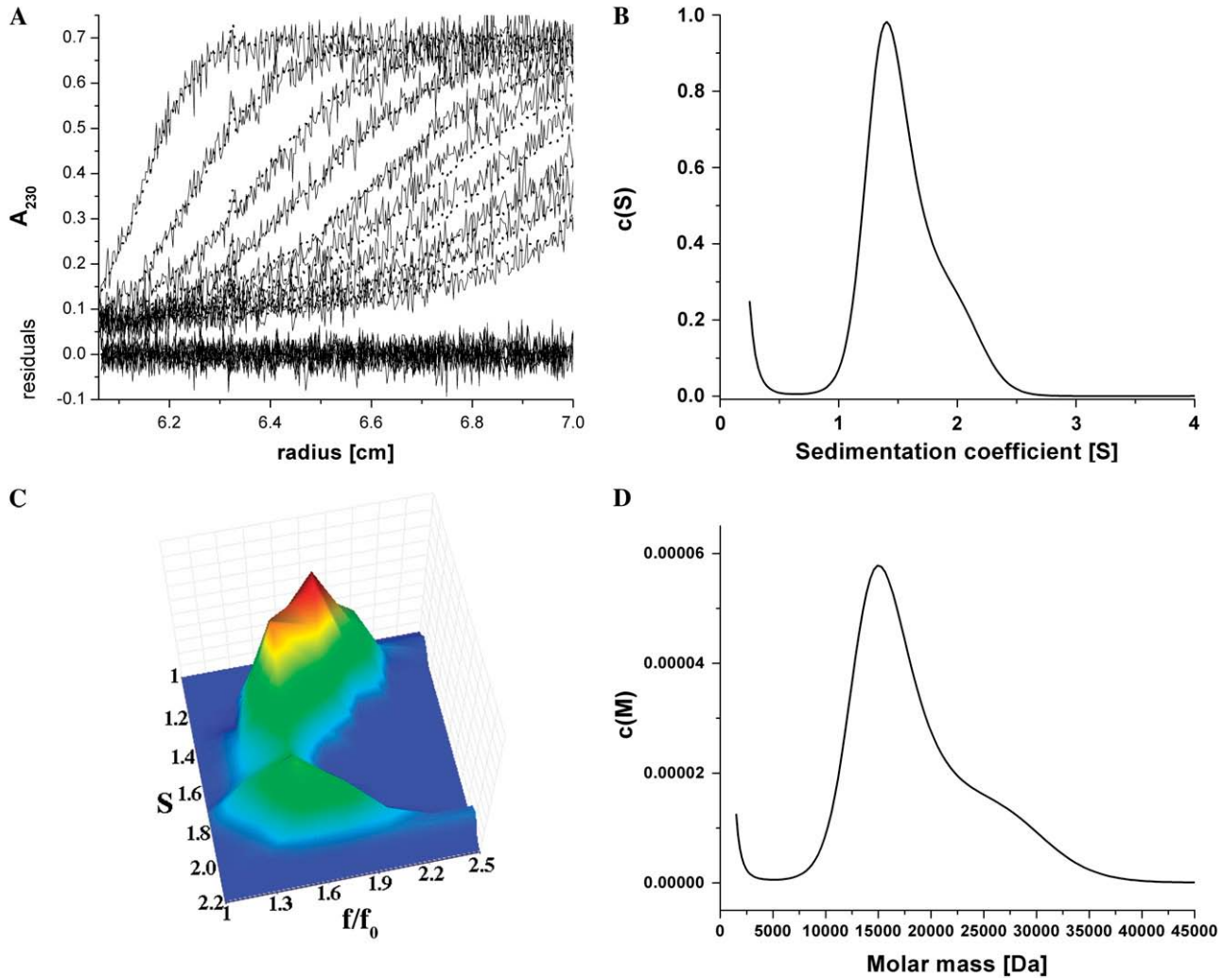

FIGURE 8 Analytical ultracentrifugation of hNL3-cyt. An AUC sedimentation velocity experiment was run at 0.3 mg/ml in PBS adjusted to pH 8.0. (*A*) Sedimentation data (*solid line*) and $c(s)$ analysis fit (*dotted line*) of every 10th scan. Residuals are shown at the bottom, with a root mean-square deviation of 0.022. (*B*) $c(s)$ distribution of hNL3-cyt. hNL3-cyt has a broad distribution with a principal peak ($S = 1.33$), and a shoulder (~2 S). (*C*) Plot of S versus the frictional ratio, f/f_o, derived from the $c(s, f_r)$ analysis. Both populations of hNL3-cyt have high f/f_o values, indicating that both are extended in solution. (*D*) $c(M)$ distribution of the same data. The first peak has a maximum at 15,227 ± 300 Da, in excellent agreement with the molecular weight predicted from the sequence.

SAXS

The scattering patterns for hNL3-cyt under nondenaturing conditions (1 mM EDTA/7 mM DTT/250 mM NaCl/100 mM Tris, pH 8.5) and under denaturing conditions (4 M urea/1 mM EDTA/7 mM DTT/250 mM NaCl/100 mM Tris, pH 8.5) are shown in Fig. 9 *A*. The processed scattering patterns, for hNL3-cyt under both denaturing and nondenaturing conditions, yield an effective molecular mass of ~14 kDa, in agreement with the value predicted from the sequence, indicating that the protein was mostly monomeric in solution. The corresponding distance distribution functions, $p(r)$, are shown in Fig. 9 *B*. Under denaturing conditions, hNL3-cyt populates the extended conformations more than under nondenaturing conditions. The values of the radius of gyration (R_g) derived from Guinier's law and the $p(r)$ functions for hNL3-cyt are 30 and 33 Å, respectively, under nondenaturing conditions, and 37 and 40 Å under denaturing conditions, using both methods. The Kratky plot ($s^2(I)$ vs. s) is commonly used to distinguish between compact and extended protein conformations. hNL3-cyt displays a shape typical of unfolded proteins, i.e., a plateau that increases monotonically at higher s values, whereas that for BSA displays a bell-shaped Kratky plot (Fig. 9 *C*). The R_g distribution of the structures selected by EOM (Fig. 9 *D*, *black curve*) for hNL3-cyt under nondenaturing conditions displays a bimodal profile. The major component adopts a more compact conformation than a modeled random coil (*blue curve*), with an average $R_g = 27$ Å, compared to 33 Å for the random coil. But a second peak, at $R_g = 53$ Å, may correspond either to a

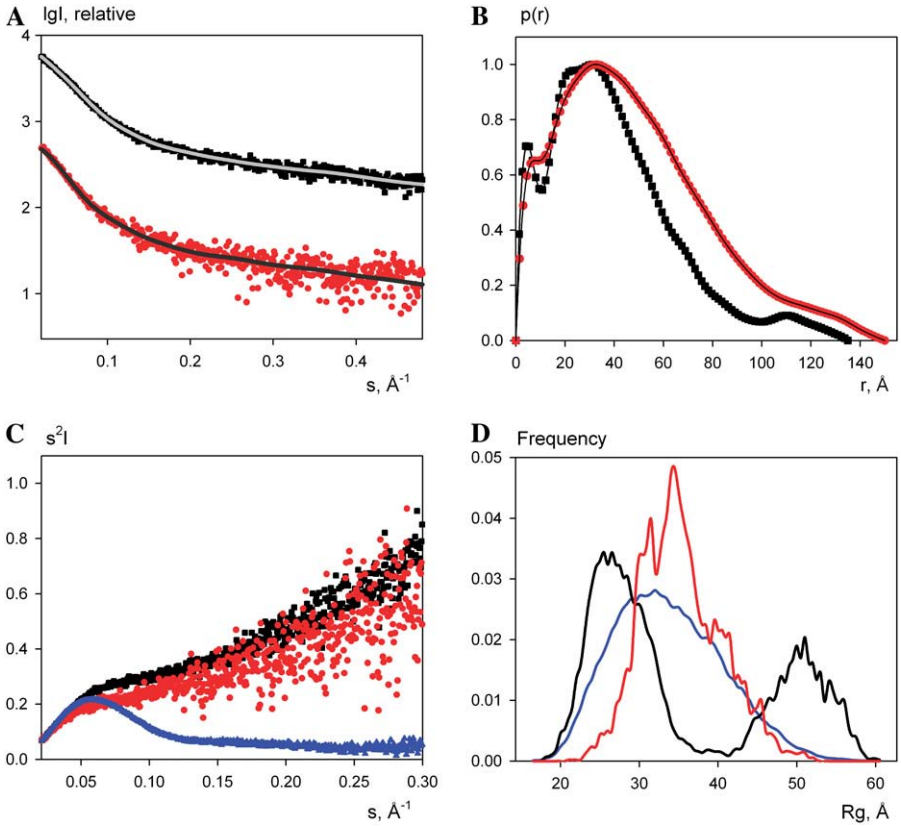

FIGURE 9 SAXS measurements on hNL3-cyt. (*A*) X-ray scattering patterns of hNL3-cyt under denaturing (●) and nondenaturing conditions (■). The fits of the ensembles selected by EOM are displayed as gray lines. (*B*) Distance distribution functions of the measurements shown in panel *A*. (*C*) Kratky plots (*red* and *black traces*) derived from the data in panel *A* and of the globular protein, BSA (*blue trace*), which served as a folded control. (*D*) Radii of gyration distributions of hNL3-cyt under denaturing (*red line*) and nondenaturing conditions (*black line*) and comparison with a pool of randomly generated structures (*blue line*).

population of highly extended molecules or the presence of dimers and higher-order oligomers. This bimodal distribution is an attribute of the native protein only; upon addition of urea the reconstructed ensembles display a monomodal R_g profile (Fig. 9 *D*, *red curve*), which more closely resembles that of the random coil distribution ($R_g = 35$ Å).

Circular dichroism

Circular dichroism (CD) in the far UV was used to assess the conformation of hNL3-cyt and the effects induced by denaturants and other reagents (Fig. 10). In 10 mM potassium phosphate, pH 8.0, the CD spectrum has a broad minimum at ~200 nm, which shifts to slightly higher wavelengths, with a concomitant decrease in negative ellipticity, as the temperature is raised (Fig. 10 *A*). There is an isodichroic point at 207 nm, above which wavelength the negative ellipticity actually increases as the temperature is raised. All three spectra lack any major contributions from α-helices, β-sheets or turns, and are characteristic of an unstructured polypeptide, with a possible PPII contribution (36,70,71).

It has been suggested that chemical denaturants increase PPII content of unfolded proteins by bonding to the polypeptide backbone, thus shifting the PPII↔disordered equilibrium to the left (72). In line with this observation, Fig. 10 *B* shows that, at 5°C, in 7 M Gdn.HCl, a maximum appears at ~224 nm, which has been ascribed to PPII (73); this peak becomes slightly negative at 25°C, and even more negative at 45°C. Unfortunately, the presence of Gdn.HCl precludes corresponding measurements at wavelengths <220 nm. Measurements with urea gave similar results (data not shown). TFE is known to promote α-helix formation in long polypeptides and proteins (74). Indeed, 50% TFE converts the CD spectrum of hNL3-cyt at 25°C to that of a more helical polypeptide, a conformation which is maintained even at 45°C (Fig. 10 *C*).

In situ, the cytoplasmic domain of NL is adjacent to the cell membrane. Possible structural changes induced by the bilayer were mimicked in vitro by addition of SDS at submicellar concentrations of 0.5 and 3 mM, or by 1 mM DOPS liposomes, also resulting in a small shift to a more helical conformation (Fig. 10 *D*). The zwitterionic detergent CHAPS, at 0.45% w/v, caused no shift in the CD spectrum of hNL3-cyt, nor did the crowding agents Ficoll 70 (30 mg/ml), sucrose (1 M), and the osmoprotectant betaine (4 M) (data not shown).

ESR

The shape of the ESR spectrum provides a measure of the mobility of the bound nitroxyl radical, which depends on the structure and flexibility of the protein in the vicinity of its site of attachment. The ESR signal of spin-labeled hNL3-cyt in 150 mM NaCl/10 mM sodium phosphate, pH 7.0, at 20°C reveals relatively sharp peaks (Fig. 11 *A*), indicating only limited restriction of the probe's movement. When a 100-fold excess of DTT is added (Fig. 11 *B*), these relatively slow

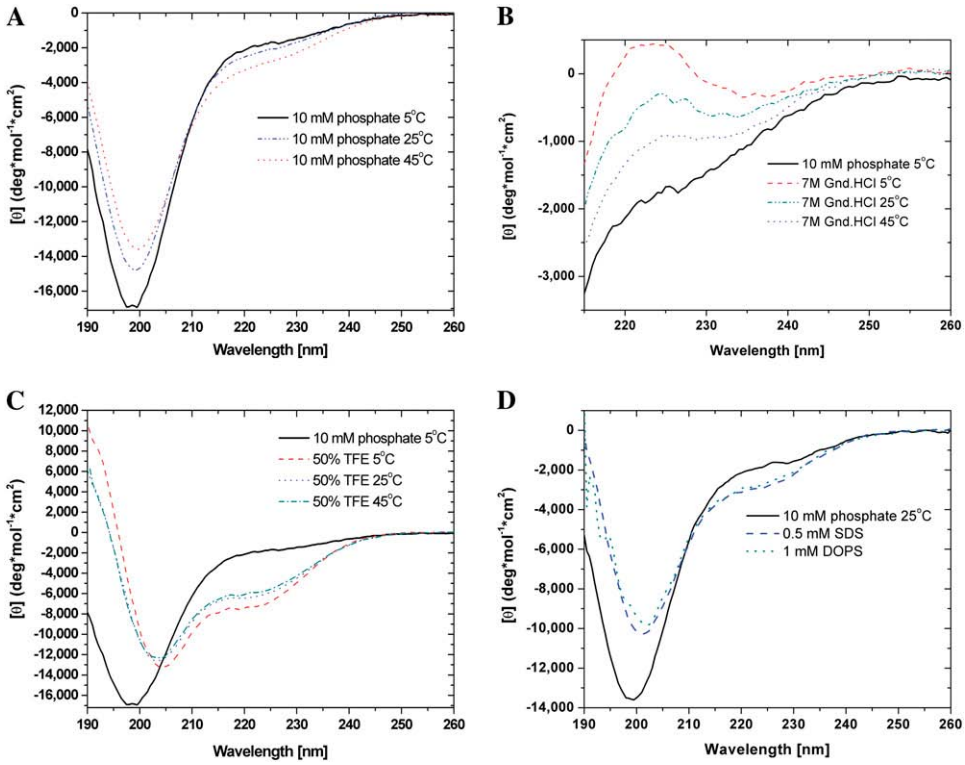

FIGURE 10 CD spectra of hNL3-cyt. All spectra were collected at an hNL3-cyt concentration of 12 μM with the appropriate buffer corrections. (*A*) CD spectra of hNL3-cyt in 10 mM potassium phosphate (KPi), pH 8.0, at various temperatures (*solid line*, 5°C; *dash-dotted line*, 25°C; and *dotted line*, 45°C). (*B*) Effect of 7M Gdn.HCl and temperature on the CD spectrum of hNL3-cyt (*solid line*, 10 mM KPi, pH 8.0, at 5°C; *long-dash*, *short-dash line*, −7 M Gdn.HCl at 5°C; *dash-dotted line*, 7 M Gdn.HCl at 25°C; and *dotted line*, 7 M Gdn.HCl at 45°C). (*C*) Effect of 50% TFE and temperature on the CD spectrum of hNL3-cyt (*solid line*, 10 mM KPi, pH 8.0, at 5°C; *long-dash*, *short-dash line*, 50% TFE at 5°C; *dotted line*, 50% TFE at 25°C; and *dash-dotted line*, 50% TFE at 45°C). (*D*) Effect of 0.5 mM SDS and 1 mM DOPS liposomes on the CD spectrum of hNL3-cyt (*solid line*, 10 mM KPi, pH 8.0, 25°C; *long-dash*, *short-dash line*, 0.5 mM SDS; and *dotted line*, 1 mM DOPS).

components disappear completely, with concomitant appearance of a sharp triplet, characteristic of unrestricted rotation of the free radical. This provides clear evidence that the probe had indeed formed a mixed disulfide with the free thiol group of the protein, which had then been reduced by DTT, with concomitant release of the spin label. Fig. 11 *C* shows the ESR spectrum obtained in the presence of 66% TFE. The peak intensity of the high-field component (h_H) of the ESR spectrum in the conjugate decreases relative to that seen in the absence of TFE, indicating a decrease in segmental mobility of the polypeptide.

Rotational correlation times for the bound nitroxyl radical were calculated according to Eq. 3, yielding τ_R values of 9.3×10^{-10} and 1.3×10^{-9} s in buffer and TFE, respectively. Comparison of these values to that for isotropic rotation of the free radical in solution, $\tau_R \sim 10^{-11}$–10^{-10} s,

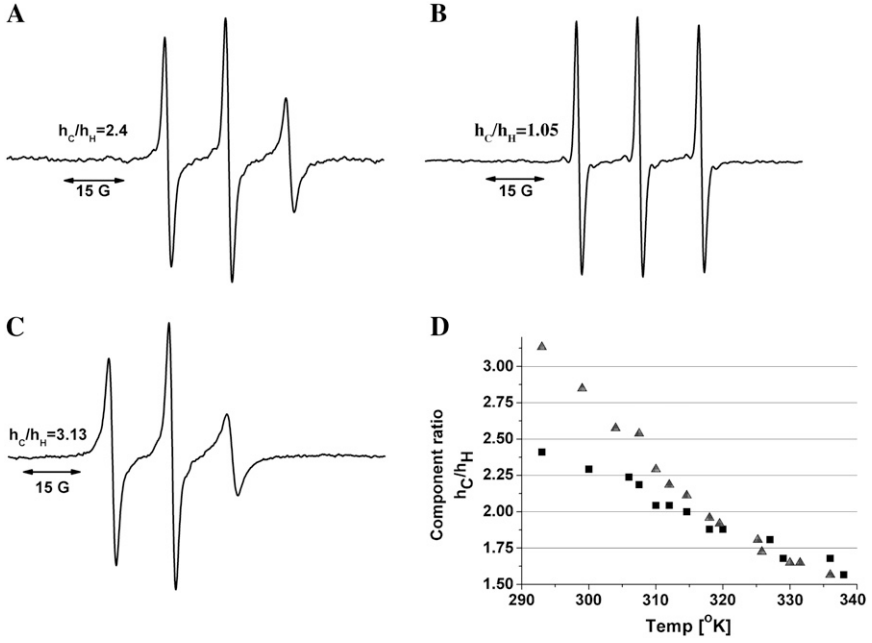

FIGURE 11 ESR measurements on spin-labeled hNL3-cyt. (*A*) The ESR spectrum of spin-labeled hNlg3-cyt (in 150 mM NaCl/10 mM sodium phosphate, pH 7.0, at 20°C), reveals relatively sharp peaks that indicate rather limited restriction of the probe's mobility. (*B*) Disappearance of the relatively slow components, with concomitant appearance of a sharp triplet, when a 100-fold molar excess of DTT was added to the spin-labeled protein. (*C*) ESR spectrum obtained in the presence of 66% TFE, at 20°C. The peak intensity of the high-field component (h_H) in the conjugate (which is very sensitive to label movement) decreases relative to that seen in the absence of TFE, suggesting a decrease in segmental mobility of the polypeptide. (*D*) Temperature-dependence of the (h_C)/(h_H) ratio. Both in the absence (■) and presence of TFE (▲), the ratio decreases with temperature, reflecting an increase in mobility.

and to values for folded proteins, $\tau_R > 10^{-7}$ s, indicates limited restriction of the motion of the bound radical in buffer, and more substantial restriction in TFE.

Fig. 11 *D* shows the temperature dependence of the ratio of the peak intensity of the central component (h_C) to that of h_H, which, as mentioned above, reflects the mobility of the covalently bound spin label. Both in the absence and presence of TFE, the ratio decreases with temperature, reflecting increased mobility. The decrease seen in the presence of TFE is more pronounced, presumably since the starting conformation is more structured. However, in both cases, no sharp transition is seen in the range of 20–65°C.

NMR

NMR spectra recorded for hNL3-cyt are typical of an unfolded protein: The signals are narrow and display very limited dispersion in the proton dimension. As can be seen in Fig. 12 *A*, the backbone amide proton chemical shifts are restricted to the coil region, i.e., 7.7–8.7 ppm. Signals observed at 112/7.0 ppm ($^{15}N/^{1}H$) originate from arginine side chains, whereas the cross peak at 130/10 ppm stems from the side chain indole of the sole tryptophan. Although the ^{15}N-^{1}H 2D NMR spectra of Fig. 12, *A* and *B*, have the appearance of an HSQC spectrum, they were recorded with the heteronuclear $^{15}N\{^{1}H\}$ NOE pulse sequence (60). The spectrum obtained with ^{1}H irradiation during the ^{15}N relaxation delay (Fig. 12 *B*) is dramatically different from the reference experiment (Fig. 12 *A*). The signal attenuation by dipolar cross relaxation (NOE) is so strong that it leads to negative crosspeaks. Such behavior is typical of unfolded proteins (75), whereas attenuation factors of 15% would be expected for the ordered regions in a folded protein with a molecular weight of 15 kDa. The extent of signal attenuation can be gleaned more clearly from Fig. 12, *C* and *D*, which were taken at 120 ppm from Fig. 12, *A* and *B*, respectively. The uniform attenuation observed for all the signals indicates that persistent structure cannot be present for any portion of the polypeptide chain on timescales longer than several nanoseconds, and that the dynamics of the amide backbone are dominated by segmental flexibility. The arginine side chains and the last residues of the sequence have greater motional freedom than backbone amides more central to the chain, and therefore give rise to even stronger negative NOEs.

DISCUSSION

Bioinformatic analysis

The cytoplasmic domains of the CLAMs, Gli-cyt, Nrt-cyt, and NL-cyt, display no sequence homology to any other polypeptide sequences or indeed to each other, except that Gli-cyt and NL-cyt possess tetrapeptide PDZ recognition motifs at their C-termini (2), and a few common eukaryotic linear motifs (76) characteristic of either phosphorylation sites or WW domain interaction sites. It is worth noting that sequence conservation among the cytoplasmic domains of the NLs is ~31%, compared to ~55% for their corresponding folded extracellular ChE-like domains.

We earlier presented physicochemical evidence, complemented by bioinformatic analysis, that Gli-cyt is an IUP, i.e., that the purified protein domain adopts an unstructured conformation in aqueous solution under physiological conditions (22). In this study, we have carried out similar characterization of hNL3-cyt.

While bioinformatic analysis clearly predicted that Gli-cyt, with a net charge of +7 and ~15% hydrophobic amino acids, should be an IUP, as was indeed confirmed by its physicochemical characterization, the situation with respect to NL3-cyt is somewhat ambiguous. As described under Results, hNL3-cyt has ~20% hydrophobic amino acids, and a net charge of −1. Furthermore, the first 54 amino acids, which have the same net charge as the whole sequence, have only ~14% hydrophobic residues; consequently, making use of the method of Uversky and co-workers (49), this segment is

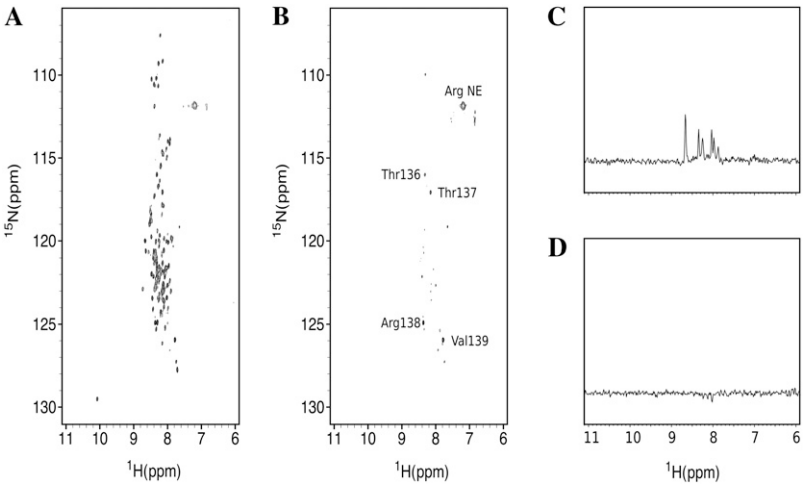

FIGURE 12 NMR measurements on hNL3-cyt. Contour plots of two-dimensional $^{15}N\{^{1}H\}$ NOE NMR spectra recorded without (*A*) and with (*B*) proton saturation. Spectra are shown at the same contour level. Contour levels in spectrum *A* are positive, and in panel *B* are negative. Arg NE denotes side-chain N^{ε}-H^{ε} correlations that are aliased in the ^{15}N dimension from ~84.7 ppm. One-dimensional traces taken from panels *A* and *B* at 120 ppm (^{15}N) are shown in panels *C* and *D*, respectively.

predicted to be unfolded ($\langle R\rangle = 0.0377$, $\langle H\rangle = 0.3634$, $\langle H\rangle_b = 0.418$). Residues 55–118 have zero net charge and ~29% hydrophobic residues; thus, making use of the same method this segment is predicted to be folded ($\langle R\rangle = 0.0153$ $\langle H\rangle = 0.4825$ $\langle H\rangle_b = 0.41$). The same analysis of the entire sequence predicts that the whole protein will fall just on the folded side of the boundary. Evidence that IUPs may contain regions that are predicted to be folded has been presented for such proteins as α-synuclein, negative factor, and helix-destabilizing protein (26). It has been suggested that the regions of these proteins predicted to be intrinsically unfolded prevent the remainder of the sequence from folding, perhaps through electrostatic interactions.

When PONDR and Foldindex are used to analyze the hNL3-cyt sequence, both clearly predict that the 50 NH_2-terminal residues are unfolded, but Foldindex predicts that residues 54–118 should be folded, whereas PONDR predicts alternating folded and unfolded stretches. Yet, secondary structure predictors predict minimal amounts of secondary structure. Furthermore, IUPred predicts the whole domain to be unstructured. Taking into account these predictions, hNL3-cyt might be assigned to the class of dynamic and collapsed IUPs, with some secondary structure propensity, as defined by Uversky and co-workers (49). These somewhat contradictory findings illustrate the pitfalls of bioinformatics analysis, and clearly show the need for experimental clarification. Although this study focused on hNL3-cyt, it is worth mentioning that the cytoplasmic domains of all the NLs have similar patterns in the various secondary structure and unstructured predictors.

Bearing in mind the ambiguities in the bioinformatics analysis we decided to clone, express, and purify hNL3-cyt, so as to permit its physicochemical characterization.

Protein expression and stability

As mentioned above, hNL3-cyt was rapidly degraded in the course of purification. This was a positive indication that hNL3-cyt is indeed an IUP, since unstructured proteins are highly sensitive to proteolysis, a feature that has been widely used as an experimental tool for characterizing IUPs (32). Purification under denaturing conditions was used to deactivate the proteases present during the initial stages, with a transition to nondenaturing buffers during the final purification steps. We suggest this method as a general procedure for purification of IUPs. It could thus serve as an alternative to other more aggressive lysis and purification techniques, such as boiling (77), which might result in protein aggregation and concomitant trapping of IUPs in the aggregates.

hNL3-cyt is relatively unstable in solution, and displays a tendency to aggregate reversibly at high concentrations (>5 mg/ml), upon contact with chemical chaperones and positive micelles, and during various mechanical manipulations, all of which hinder certain experiments. A possible explanation for this instability is the presence of hydrophobic residues in hNL3-cyt that are not part of an intramolecular hydrophobic core, and, thus being exposed, have a propensity to participate in intermolecular interactions.

Obtaining a relatively stable form of hNL3-cyt enabled us to carry out structural characterization under a variety of conditions using an extensive repertoire of methodologies.

Physicochemical characterization

Several physicochemical techniques, e.g., FCS, SDS-PAGE (78), SEC, SAXS, NMR, all provide strong evidence that hNL3-cyt has an extended conformation, as summarized in Table 1. Both CD in the far UV and NMR, that displays limited backbone amide chemical shift dispersion, rule out the presence of substantial amounts of classical secondary structure, in agreement with the prediction of PROF. Instead of retaining a fixed secondary structure, the protein rapidly interconverts between different conformations of similar free energy. The attenuation obtained in the heteronuclear $^{15}N\{^1H\}$ NOE NMR experiment shows that the chain is flexible on the (sub)nanosecond timescale, indicating that energy barriers between the visited conformations are small. Furthermore, there is no significant ellipticity in the near UV, little binding of ANS, no change in tryptophan fluorescence, and only a slight increase in mobility of the already mobile spin label upon denaturation (data not shown). These data, taken together, suggest that hNL3-cyt is an IUP in vitro, in the size range of a premolten globule (79,80) as operationally defined by Uversky (23). Indeed, FCS, SAXS, and SEC clearly show that it becomes substantially more extended under denaturing conditions under which it approaches dimensions expected for a random coil configuration (see Table 1). The more compact dimensions of hNL3-cyt under nondenaturing conditions than in the presence of denaturants may be ascribed to dynamic intrachain interactions that constrict some parts of the polypeptide chain in water. It is possible that its conformation may bear some resemblance to the collapsed conformations observed during protein folding from denaturant (81) as reported for the intrinsically unfolded NM segment of the yeast prion protein, Sup35 (82).

The ESR measurements provide local information pertaining to the vicinity of the bound radical, and clearly show that this region is highly flexible in solution. Thermal denaturation experiments monitored using both ESR and CD (data not shown) reveal a monotonic change in conformation with temperature, with no evidence for a two-state transition typical of globular proteins (83).

Near-UV CD can, in principle, serve as a tool to detect constrained conformational freedom of aromatic side chains in solution, but its utility for studying IUPs is rather limited. This is because, in general, IUPs have a low aromatic content (84), as well as a low level of nonlocal interactions which, in turn, might favor fast interconversions of those interactions which exist. Breaking of these intrachain interactions by denaturants should thus produce a more extended structure. It

TABLE 1 Calculated and measured mass and sizes of hNL3-cyt under various conditions

Molecular mass	15.34 kDa	
Calculated* folded R_H	19.58 ± 0.1 Å	
Calculated* unfolded R_H	35.67 ± 0.30 Å (Gdn.HCl), 34.21 ± 0.06 Å (Urea)	
Calculated†‡ unfolded R_g	40.7 ± 9.3 Å†, 39.6 Å‡	
	Nondenaturing conditions	Denaturing conditions
SEC R_H	28.3 ± 0.5 Å	33.1 ± 0.4 Å (urea)
DLS R_H	23.3 ± 2.0 Å	28.4 ± 3.0 Å (Gdn.HCl)
FCS R_H	25.0 ± 0.6 Å	33.0 ± 0.8 Å (Gdn.HCl)
AUC R_H	27.3 ± 0.4 Å	
SAXS R_g	33.0 ± 3.0 Å	40.0 ± 4.0 Å (urea)

*According to Uversky (23).
†According to Kohn et al. (63).
‡According to Moncoq et al. (64).

has, however, been suggested that denaturants increase the percentage of residues in the PPII conformation relative to disordered backbone conformations (73,85). Since PPII is believed to consist of short stretches of residues in an extended helical conformation, broken by sharp turns, a polypeptide sequence with a high and stable PPII content should, in principle, have more expanded dimensions than a true random coil (86). Thus, if a protein containing the same number of residues as hNL3-cyt were assumed to be entirely composed of PPII helices, with a persistence length of 4 nm (corresponding to ~14 residues) (41,42), crudely modeled for simplicity as a freely jointed chain (67), the calculated radius of gyration would be ~52 Å, as compared to ~40 Å calculated for a random coil by two different methods (Table 1). None of the experimental values that we measure, even under strongly denaturing conditions, approach the value for a PPII jointed chain, but values corresponding to a random coil are attained. This does not exclude the possibility that sequences with PPII propensity may exist, possibly with shorter persistence lengths which might, moreover, dynamically fluctuate along the chain.

As mentioned above, it has been proposed that increasing denaturant concentration produces increased PPII content in short peptides and proteins (73). In these terms of reference, protein denaturation can be considered as a transition from a globally structured state to a dynamic ensemble of locally structured states. Denaturation of IUPs cannot, however, be viewed in such terms. Rather, their denaturation can be seen as a transition from a set of globally flexible, extended, and largely unstructured states to a dynamic ensemble of even more extended locally structured states. The CD spectrum of a PPII helix consisting entirely of proline residues displays a dichroic minimum at ~200 nm, and a maximum at 228 nm, which shifts toward lower wavelengths if the polypeptide contains a substantial percentage of residues other than proline (87). Upon heating, the ~200 nm minimum becomes less pronounced, and the ellipticity at ~228 nm decreases, suggesting that the PPII ↔ unstructured equilibrium is shifted to the right due to destruction of the water shell around the polypeptide that lacks intrachain hydrogen bonds which might otherwise serve to stabilize it (38,88). The aforementioned ellipticity maximum characteristic of PPII is absent from the spectrum of hNL3-cyt under nondenaturing conditions, but small increases in ellipticity at 228 nm are observed for hNL3-cyt upon denaturation with either Gdn.HCl or urea (Fig. 10 *B*), implying that denaturation raises the propensity of hNL3-cyt to assume a PPII conformation. It may be noted that a much more pronounced effect is observed in similar measurements performed on Gli-cyt (A. Paz, J. L. Sussman, and I. Silman, unpublished).

Taken together, the hydrodynamic and spectroscopic data discussed above suggest that the PPII conformation may contribute to the ensemble of conformations that an IUP samples, but that in hNL3-cyt there are apparently no prominent PPII helices per se.

An alternative interpretation of the spectral changes observed, in particular of the reversible changes in ellipticity at 200 and 222 nm observed upon heating, is the turnout response; a gain in structural complexity due to strengthening of the hydrophobic effect, resulting in a stronger driving force for partial folding (23,89,90).

We chose to refrain from quantifying the secondary structure content on the basis of the far-UV CD spectra, since use of the various algorithms and reference data sets implemented in DichroWeb (91) yielded conflicting results. The main reason for these inconsistencies is the lack of reliable reference data sets for IUPs. Furthermore, although ellipticity at 222 nm (θ_{222}) has been widely employed as a measure of α-helical content, both in globular proteins and in IUPs (92), this same parameter has also been used to estimate PPII content (93,94). Indeed, Bienkiewicz et al. (95) used this criterion to estimate PPII content of an IUP, assuming a two-state equilibrium between PPII and unstructured conformations. However, if an IUP were to possess some α-helical content, it would affect the two-state assumption, and, in any case, it would not be possible to apply the θ_{222} measure to assessment of both PPII and α-helical content simultaneously. Thus, caution should be exercised in utilizing far-UV CD for quantification of the structural motifs present in a given IUP that might contain various residual structure elements along its polypeptide chain, such as molecular recognition features (96), that would contribute differently to θ_{222}.

Biological implications of the unordered nature of hNL3-cyt

Having established that hNL3-cyt is an IUP puts previous studies on NL into a new perspective. hNL3 is a human homolog of *Drosophila* Gli, based on sequence homology of their extracellular AChE-like domains (4). However, their intracellular domains display no detectable sequence homology, except for the short linear motifs referred to in the

Introduction. Nevertheless, both are IUPs that are expressed at cellular junctions and appear to participate in protein-protein interactions. Investigation of the evolutionary rates of IUPs revealed that, in general, they are less conserved than ordered proteins (97). If a domain's disorder is important to its interaction with multiple partners, and hence to its plastic function, it may well be that there is no evolutionary pressure to maintain a particular amino acid, as long as it is replaced by an amino acid which is also disorder-promoting, so that mutation does not bring about an increase in secondary structure in the disordered domain. In such a case, the only specific evolutionary pressure for sequence preservation in the cytoplasmic CLAM domains will apply to the four-residue PDZ recognition motif at the COOH-terminus, and other putative recognition motifs along the sequence, thus explaining the sequence divergence within the CLAM family. It has, furthermore, been noted that the cytoplasmic domains of the various NLs are less conserved than their extracellular ChE domains (31% and 55% identity, respectively); the sequence differences between the various NL-cyt domains was attributed to localization signals or to other unknown functions (98). We raise the possibility that this low degree of sequence conservation can be ascribed to the aforementioned weaker evolutionary pressure on the disordered cytoplasmic sequences.

NL has been shown to putatively bind 13 proteins via interactions through its PDZ recognition motif (9). This promiscuity was further suggested by the authors to be enhanced by the promiscuous binding of the target proteins themselves to additional multiple partners. We suggest that the complexity of the interactome of NL-cyt is further amplified due to its unstructured character; consequently, it may have the capacity to bind to even more targets in a PDZ-domain-independent manner. In this context, it may be noted that a large percentage of hub proteins in interactomes are, indeed, IUPs (29,99–103). The degree of sequence homology of the cytoplasmic domains of the human NLs argues that they will function similarly in an intracellular context.

A major issue in research on IUPs is their conformational state in vivo. For some IUPs, experimental evidence suggests that they are fully unfolded within the living cell (104). For a number of other IUPs, it has been shown that binding to a partner such as a protein or DNA is coupled to folding, and has been termed induced folding (24). hNL3-cyt has the propensity to form helices upon addition of TFE. More biologically relevant is the moderate structuring induced in hNL3-cyt in the presence of either SDS or DOPS liposomes, which may mimic the physiological environment of NLs.

The high sensitivity of IUPs to proteases offers the possibility of tight regulation of their expression pattern (33); indeed, disorder has recently been shown to be correlated with short half-lives (105). A number of proteins have been found to bind to NL, some of which colocalize with it within the postsynaptic density (9). We have previously shown that binding of PSD-95 and/or S-SCAM to hNL3-cyt can protect it from degradation by the 20S proteasome (33). Obviously, under conditions of development and plasticity, changes in expression patterns of partner proteins could render hNL3-cyt and its homologs readily susceptible to degradation.

In this study we have shown that hNL3-cyt is an IUP, like Gli-cyt (and most likely also the cytoplasmic domain of a third CLAM, neurotactin (2)). As already mentioned, IUPs break the accepted structure/function paradigm that folding to a single native conformation confers biological function and specificity (106). On the contrary, IUPs have the potential to assume multiple conformations, and to interact with multiple partners. Indeed, many IUPs have been shown to serve as hub proteins in interactomes (99). The synapse is a site at which functional flexibility, i.e., synaptic plasticity, is believed to play crucial roles in development and remodeling of the nervous system, in particular in learning and memory (107). We have shown that hNL3-cyt is an IUP, whose structural flexibility confers on it functional promiscuity that is in keeping with the roles assigned to NLs in synaptic assembly (108).

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

We thank Dr. Fred Naider (Department of Chemistry, College of Staten Island, The City University of New York, Staten Island, New York) for valuable discussions and critical reading of the manuscript, Dr. Peter Scheiffele (Department of Cell Biology, College of Physicians & Surgeons, Columbia University, New York) for the human NL3 clone, Lilly Toker and Michal Cohen for a sample of purified *Tc*AChE, and Yacov Kacher for his help in preparing the liposomes.

This research was supported by the Israel Science Foundation; Autism Speaks; the European Commission VIth Framework Research and Technological Development Program (grant No. 031220); the US-Israel Binational Science Foundation (grant No. 2002371); a grant of the Israel Ministry of Science, Culture, and Sport to the Israel Structural Proteomics Center; the Benziyo Center for Neuroscience; the Kimmelman Center for Biomolecular Structure and Assembly; the Minerva Foundation; the Divadol Foundation; the Bruce Rosen Foundation; the Jean and Julia Goldwurm Memorial Foundation; the Neuman Foundation; and by the historic generosity of the Harold Perlman Family. M.L. was supported by Training and Mobility of Research grant No. HPRN-CT-2002-00241 of the European Community. F.A.A.M. is a recipient of a VIDI Scheme grant from the Netherlands Organization for Scientific Research. J.L.S. is the Morton and Gladys Pickman Professor of Structural Biology.

REFERENCES

1. Botti, S. A., C. E. Felder, J. L. Sussman, and I. Silman. 1998. Electrotactins: a class of adhesion proteins with conserved electrostatic and structural motifs. *Protein Eng.* 11:415–420.
2. Gilbert, M. M., and V. J. Auld. 2005. Evolution of clams (cholinesterase-like adhesion molecules): structure and function during development. *Front. Biosci.* 10:2177–2192.
3. Ichtchenko, K., Y. Hata, T. Nguyen, B. Ullrich, M. Missler, C. Moomaw, and T. C. Sudhof. 1995. Neuroligin 1: a splice site-specific ligand for β-neurexins. *Cell.* 81:435–443.

4. Gilbert, M., J. Smith, A. J. Roskams, and V. J. Auld. 2001. Neuroligin 3 is a vertebrate gliotactin expressed in the olfactory ensheathing glia, a growth-promoting class of macroglia. *Glia*. 34:151–164.
5. Bolliger, M. F., K. Frei, K. H. Winterhalter, and S. M. Gloor. 2001. Identification of a novel neuroligin in humans which binds to PSD-95 and has a widespread expression. *Biochem. J*. 356:581–588.
6. Boucard, A. A., A. A. Chubykin, D. Comoletti, P. Taylor, and T. C. Sudhof. 2005. A splice code for *trans*-synaptic cell adhesion mediated by binding of neuroligin 1 to α- and β-neurexins. *Neuron*. 48:229–236.
7. Taniguchi, H., L. Gollan, F. G. Scholl, V. Mahadomrongkul, E. Dobler, N. Limthong, M. Peck, C. Aoki, and P. Scheiffele. 2007. Silencing of neuroligin function by postsynaptic neurexins. *J. Neurosci.* 27:2815–2824.
8. Hirao, K., Y. Hata, N. Ide, M. Takeuchi, M. Irie, I. Yao, M. Deguchi, A. Toyoda, T. C. Sudhof, and Y. Takai. 1998. A novel multiple PDZ domain-containing molecule interacting with *N*-methyl-D-aspartate receptors and neuronal cell adhesion proteins. *J. Biol. Chem.* 273: 21105–21110.
9. Meyer, G., F. Varoqueaux, A. Neeb, M. Oschlies, and N. Brose. 2004. The complexity of PDZ domain-mediated interactions at glutamatergic synapses: a case study on neuroligin. *Neuropharmacology*. 47:724–733.
10. Kim, E., M. Niethammer, A. Rothschild, Y. N. Jan, and M. Sheng. 1995. Clustering of *Shaker*-type K^+ channels by interaction with a family of membrane-associated guanylate kinases. *Nature*. 378:85–88.
11. Song, J. Y., K. Ichtchenko, T. C. Sudhof, and N. Brose. 1999. Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses. *Proc. Natl. Acad. Sci. USA*. 96:1100–1105.
12. Varoqueaux, F., S. Jamain, and N. Brose. 2004. Neuroligin 2 is exclusively localized to inhibitory synapses. *Eur. J. Cell Biol.* 83:449–456.
13. Scheiffele, P., J. Fan, J. Choih, R. Fetter, and T. Serafini. 2000. Neuroligin expressed in nonneuronal cells triggers presynaptic development in contacting axons. *Cell*. 101:657–669.
14. Rosales, C. R., K. D. Osborne, G. V. Zuccarino, P. Scheiffele, and M. A. Silverman. 2005. A cytoplasmic motif targets neuroligin-1 exclusively to dendrites of cultured hippocampal neurons. *Eur. J. Neurosci*. 22:2381–2386.
15. Dresbach, T., A. Neeb, G. Meyer, E. D. Gundelfinger, and N. Brose. 2004. Synaptic targeting of neuroligin is independent of neurexin and SAP90/PSD95 binding. *Mol. Cell. Neurosci*. 27:227–235.
16. Jamain, S., H. Quach, C. Betancur, M. Rastam, C. Colineaux, I. C. Gillberg, H. Soderstrom, B. Giros, M. Leboyer, C. Gillberg, and T. Bourgeron. 2003. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat. Genet.* 34:27–29.
17. Laumonnier, F., F. Bonnet-Brilhault, M. Gomot, R. Blanc, A. David, M. P. Moizard, M. Raynaud, N. Ronce, E. Lemonnier, P. Calvas, B. Laudier, J. Chelly, J. P. Fryns, H. H. Ropers, B. C. Hamel, C. Andres, C. Barthelemy, C. Moraine, and S. Briault. 2004. X-linked mental retardation and autism are associated with a mutation in the NLGN4 gene, a member of the neuroligin family. *Am. J. Hum. Genet*. 74:552–557.
18. Philibert, R. A., S. L. Winfield, H. K. Sandhu, B. M. Martin, and E. I. Ginns. 2000. The structure and expression of the human neuroligin-3 gene. *Gene*. 246:303–310.
19. Comoletti, D., A. De Jaco, L. L. Jennings, R. E. Flynn, G. Gaietta, I. Tsigelny, M. H. Ellisman, and P. Taylor. 2004. The $Arg^{451}Cys$-neuroligin-3 mutation associated with autism reveals a defect in protein processing. *J. Neurosci*. 24:4889–4893.
20. Chih, B., S. K. Afridi, L. Clark, and P. Scheiffele. 2004. Disorder-associated mutations lead to functional inactivation of neuroligins. *Hum. Mol. Genet*. 13:1471–1477.
21. Tabuchi, K., J. Blundell, M. R. Etherton, R. E. Hammer, X. Liu, C. M. Powell, and T. C. Sudhof. 2007. A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. *Science*. 318:71–76.
22. Zeev-Ben-Mordehai, T., E. H. Rydberg, A. Solomon, L. Toker, V. J. Auld, I. Silman, S. Botti, and J. L. Sussman. 2003. The intracellular domain of the *Drosophila* cholinesterase-like neural adhesion protein, gliotactin, is natively unfolded. *Proteins*. 53:758–767.
23. Uversky, V. N. 2002. What does it mean to be natively unfolded? *Eur. J. Biochem*. 269:2–12.
24. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol*. 6:197–208.
25. Fink, A. L. 2005. Natively unfolded proteins. *Curr. Opin. Struct. Biol*. 15:35–41.
26. Uversky, V. N., J. R. Gillespie, and A. L. Fink. 2000. Why are ''natively unfolded'' proteins unstructured under physiologic conditions? *Proteins*. 41:415–427.
27. Dunker, A. K., J. D. Lawson, C. J. Brown, R. M. Williams, P. Romero, J. S. Oh, C. J. Oldfield, A. M. Campen, C. M. Ratliff, K. W. Hipps, J. Ausio, M. S. Nissen, R. Reeves, C. Kang, C. R. Kissinger, R. W. Bailey, M. D. Griswold, W. Chiu, E. C. Garner, and Z. Obradovic. 2001. Intrinsically disordered protein. *J. Mol. Graph. Model*. 19:26–59.
28. Pontius, B. W. 1993. Close encounters: why unstructured, polymeric domains can increase rates of specific macromolecular association. *Trends Biochem. Sci*. 18:181–186.
29. Dunker, A. K., M. S. Cortese, P. Romero, L. M. Iakoucheva, and V. N. Uversky. 2005. Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J*. 272:5129–5148.
30. Tompa, P., C. Szasz, and L. Buday. 2005. Structural disorder throws new light on moonlighting. *Trends Biochem. Sci*. 30:484–489.
31. Dunker, A. K., Z. Obradovic, P. Romero, E. C. Garner, and C. J. Brown. 2000. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform*. 11:161–171.
32. Fontana, A., P. P. de Laureto, B. Spolaore, E. Frare, P. Picotti, and M. Zambonin. 2004. Probing protein structure by limited proteolysis. *Acta Biochim. Pol*. 51:299–321.
33. Tsvetkov, P., G. Asher, A. Paz, N. Reuven, J. L. Sussman, I. Silman, and Y. Shaul. 2007. Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins*. 70:1357–1366.
34. Shi, Z., C. A. Olson, G. D. Rose, R. L. Baldwin, and N. R. Kallenbach. 2002. Polyproline II structure in a sequence of seven alanine residues. *Proc. Natl. Acad. Sci. USA*. 99:9190–9195.
35. Arnott, S., and S. D. Dover. 1968. The structure of poly-L-proline II. *Acta Crystallogr. B*. 24:599–601.
36. Creamer, T. P. 1998. Left-handed polyproline II helix formation is (very) locally driven. *Proteins*. 33:218–226.
37. Kelly, M. A., B. W. Chellgren, A. L. Rucker, J. M. Troutman, M. G. Fried, A. F. Miller, and T. P. Creamer. 2001. Host-guest study of left-handed polyproline II helix formation. *Biochemistry*. 40:14376–14383.
38. Tiffany, M. L., and S. Krimm. 1968. Circular dichroism of poly-L-proline in an unordered conformation. *Biopolymers*. 6:1767–1770.
39. Zagrovic, B., J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, S. Doniach, and V. S. Pande. 2005. Unusual compactness of a polyproline type II structure. *Proc. Natl. Acad. Sci. USA*. 102:11698–11703.
40. Makowska, J., S. Rodziewicz-Motowidlo, K. Baginska, J. A. Vila, A. Liwo, L. Chmurzynski, and H. A. Scheraga. 2006. Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins. *Proc. Natl. Acad. Sci. USA*. 103:1744–1749.
41. Steinberg, I. Z., W. F. Harrington, A. Berger, M. Sela, and E. Katchalski. 1960. The configurational changes of poly-L-proline in solution. *J. Am. Chem. Soc*. 82:5263–5279.
42. Schuler, B., E. A. Lipman, P. J. Steinbach, M. Kumke, and W. A. Eaton. 2005. Polyproline and the ''spectroscopic ruler'' revisited with single-molecule fluorescence. *Proc. Natl. Acad. Sci. USA*. 102:2754–2759.
43. Tcherkasskaya, O., E. A. Davidson, and V. N. Uversky. 2003. Biophysical constraints for protein structure prediction. *J. Proteome Res*. 2:37–42.

44. Prilusky, J., C. E. Felder, T. Zeev-Ben-Mordehai, E. H. Rydberg, O. Man, J. S. Beckmann, I. Silman, and J. L. Sussman. 2005. FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics*. 21:3435–3438.

45. Li, X., P. Romero, M. Rani, A. K. Dunker, and Z. Obradovic. 1999. Predicting protein disorder for N-, C-, and internal regions. *Genome Inform. Ser. Workshop Genome Inform.* 10:30–40.

46. Romero, P., Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, and A. K. Dunker. 2001. Sequence complexity of disordered protein. *Proteins*. 42:38–48.

47. Dosztanyi, Z., V. Csizmok, P. Tompa, and I. Simon. 2005. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*. 21:3433–3434.

48. Rost, B., G. Yachdav, and J. Liu. 2004. The PredictProtein server. *Nucleic Acids Res.* 32:W321–W326.

49. Oldfield, C. J., Y. Cheng, M. S. Cortese, C. J. Brown, V. N. Uversky, and A. K. Dunker. 2005. Comparing and combining predictors of mostly disordered proteins. *Biochemistry*. 44:1989–2000.

50. Krichevsky, O., and G. Bonnet. 2002. Fluorescence correlation spectroscopy: the technique and its applications. *Rep. Prog. Phys.* 65:251–297.

51. Schuck, P. 2000. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* 78:1606–1619.

52. Brown, P. H., and P. Schuck. 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys. J.* 90:4651–4661.

53. Laue, T., B. Shah, T. Ridgeway, and S. Pelletier. 1992. Computer-aided interpretation of analytical sedimentation data for proteins. *In* Analytical Ultracentrifugation in Biochemistry and Polymer Science. The Royal Society of Chemistry, Cambridge, UK. 96–125.

54. Roessle, M. W., R. Klaering, U. Ristau, B. Robrahn, D. Jahn, D. Gehrmann, P. Konarev, A. Round, S. Fiedler, C. Hermes, and D. I. Svergun. 2007. Upgrade of the small-angle x-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. *J. Appl. Cryst.* 40:S190–S194.

55. Konarev, P. V., V. V. Volkov, A. V. Sokolova, M. H. J. Koch, and D. I. Svergun. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Cryst.* 36:1277–1282.

56. Guinier, A. 1939. Le diffraction de rayons X aux tres petits angles. *Ann. Phys.* (*Paris*). 12:161.

57. Svergun, D. I. 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Cryst.* 25:495–503.

58. Bernado, P., E. Mylonas, M. V. Petoukhov, M. Blackledge, and D. I. Svergun. 2007. Structural characterization of flexible proteins using small-angle x-ray scattering. *J. Am. Chem. Soc.* 129:5656–5664.

59. Knowles, P. F., D. Marsh, and H. W. E. Rattle. 1976 Magnetic Resonance of Biomolecules: An Introduction to the Theory and Practice of NMR and ESR in Biological Systems. Wiley, London, UK.

60. Farrow, N. A., R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay. 1994. Backbone dynamics of a free and phosphopeptide-complexed SRC homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry*. 33:5984–6003.

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

62. Flory, P. J. 1969. Statistical Mechanics of Chain Molecules. Wiley, New York.

63. Kohn, J. E., I. S. Millett, J. Jacob, B. Zagrovic, T. M. Dillon, N. Cingel, R. S. Dothager, S. Seifert, P. Thiyagarajan, T. R. Sosnick, M. Z. Hasan, V. S. Pande, I. Ruczinski, S. Doniach, and K. W. Plaxco. 2004. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. USA*. 101:12491–12496.

64. Moncoq, K., I. Broutin, C. T. Craescu, P. Vachette, A. Ducruix, and D. Durand. 2004. SAXS study of the PIR domain from the Grb14 molecular adaptor: a natively unfolded protein with a transient structure primer? *Biophys. J.* 87:4056–4064.

65. Rowe, G., and A. Lopez Pineiro. 1990. Influence of the solvent on the conformational-dependent properties of random-coil polypeptides. I. The mean-square of the end-to-end distance and of the dipole moment. *Biophys. Chem.* 36:57–64.

66. Perez, J., P. Vachette, D. Russo, M. Desmadril, and D. Durand. 2001. Heat-induced unfolding of neocarzinostatin, a small all-β protein investigated by small-angle x-ray scattering. *J. Mol. Biol.* 308:721–743.

67. Tanford, C. 1961. Physical Chemistry of Macromolecules. Wiley & Sons, New York.

68. Fontana, A., M. Zambonin, P. Polverino de Laureto, V. De Filippis, A. Clementi, and E. Scaramella. 1997. Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol.* 266: 223–230.

69. Dolginova, E. A., E. Roth, I. Silman, and L. M. Weiner. 1992. Chemical modification of *Torpedo* acetylcholinesterase by disulfides: appearance of a ''molten globule'' state. *Biochemistry*. 31:12248–12254.

70. Eker, F., K. Griebenow, and R. Schweitzer-Stenner. 2003. Stable conformations of tripeptides in aqueous solution studied by UV circular dichroism spectroscopy. *J. Am. Chem. Soc.* 125:8178–8185.

71. Gokce, I., R. W. Woody, G. Anderluh, and J. H. Lakey. 2005. Single peptide bonds exhibit poly(pro)II (''random coil'') circular dichroism spectra. *J. Am. Chem. Soc.* 127:9700–9701.

72. Siligardi, G., and A. F. Drake. 1995. The importance of extended conformations and, in particular, the PII conformation for the molecular recognition of peptides. *Biopolymers*. 37:281–292.

73. Whittington, S. J., B. W. Chellgren, V. M. Hermann, and T. P. Creamer. 2005. Urea promotes polyproline II helix formation: implications for protein denatured states. *Biochemistry*. 44:6269–6275.

74. Buck, M. 1998. Trifluoroethanol and colleagues: cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* 31:297–355.

75. Ohnishi, S., and D. Shortle. 2003. Effects of denaturants and substitutions of hydrophobic residues on backbone dynamics of denatured staphylococcal nuclease. *Protein Sci.* 12:1530–1537.

76. Puntervoll, P., R. Linding, C. Gemund, S. Chabanis-Davidson, M. Mattingsdal, S. Cameron, D. M. Martin, G. Ausiello, B. Brannetti, A. Costantini, F. Ferre, V. Maselli, A. Via, G. Cesareni, F. Diella, G. Superti-Furga, L. Wyrwicz, C. Ramu, C. McGuigan, R. Gudavalli, I. Letunic, P. Bork, L. Rychlewski, B. Kuster, M. Helmer-Citterich, W. N. Hunter, R. Aasland, and T. J. Gibson. 2003. ELM server: a new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res.* 31:3625–3630.

77. Cherny, I., and E. Gazit. 2004. The YefM antitoxin defines a family of natively unfolded proteins: implications as a novel antibacterial target. *J. Biol. Chem.* 279:8252–8261.

78. Creamer, L. K., and T. Richardson. 1984. Anomalous behavior of bovine α s1- and β-caseins on gel electrophoresis in sodium dodecyl sulfate buffers. *Arch. Biochem. Biophys.* 234:476–486.

79. Uversky, V. N., and O. B. Ptitsyn. 1996. Further evidence on the equilibrium ''pre-molten globule state'': four-state guanidinium chloride-induced unfolding of carbonic anhydrase B at low temperature. *J. Mol. Biol.* 255:215–228.

80. Jeng, M. F., and S. W. Englander. 1991. Stable submolecular folding units in a non-compact form of cytochrome c. *J. Mol. Biol.* 221:1045–1061.

81. Sherman, E., and G. Haran. 2006. Coil-globule transition in the denatured state of a small protein. *Proc. Natl. Acad. Sci. USA*. 103: 11539–11543.

82. Mukhopadhyay, S., R. Krishnan, E. A. Lemke, S. Lindquist, and A. A. Deniz. 2007. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. USA*. 104:2649–2654.

83. Klug, C. S., and J. B. Feix. 1998. Guanidine hydrochloride unfolding of a transmembrane β-strand in FepA using site-directed spin labeling. *Protein Sci.* 7:1469–1476.

84. Dunker, A. K., C. J. Brown, and Z. Obradovic. 2002. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* 62:25–49.

85. Tiffany, M. L., and S. Krimm. 1973. Extended conformations of polypeptides and proteins in urea and guanidine hydrochloride. *Biopolymers.* 12:575–587.

86. Wang, Z., K. W. Plaxco, and D. E. Makarov. 2007. Influence of local and residual structures on the scaling behavior and dimensions of unfolded proteins. *Biopolymers.* 86:321–328.

87. Woody, R. W. 1992. Circular dichroism and conformation of unordered polypeptides. *Adv. Biophys. Chem.* 2:37–39.

88. Makarov, A. A., V. M. Lobachov, I. A. Adzhubei, and N. G. Esipova. 1992. Natural polypeptides in left-handed helical conformation. A circular dichroism study of the linker histones' C-terminal fragments and β-endorphin. *FEBS Lett.* 306:63–65.

89. Tanford, C. 1997. How protein chemists learned about the hydrophobic factor. *Protein Sci.* 6:1358–1366.

90. Receveur-Brechot, V., J. M. Bourhis, V. N. Uversky, B. Canard, and S. Longhi. 2006. Assessing protein disorder and induced folding. *Proteins.* 62:24–45.

91. Whitmore, L., and B. A. Wallace. 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* 32:W668–W673.

92. Cantor, C. R., and P. R. Schimmel. 1980. Biophysical Chemistry. Freeman, New York.

93. Drake, A. F., G. Siligardi, and W. A. Gibbons. 1988. Reassessment of the electronic circular dichroism criteria for random coil conformations of poly(L-lysine) and the implications for protein folding and denaturation studies. *Biophys. Chem.* 31:143–146.

94. Park, S. H., W. Shalongo, and E. Stellwagen. 1997. The role of PII conformations in the calculation of peptide fractional helix content. *Protein Sci.* 6:1694–1700.

95. Bienkiewicz, E. A., A. Moon Woody, and R. W. Woody. 2000. Conformation of the RNA polymerase II C-terminal domain: circular dichroism of long and short fragments. *J. Mol. Biol.* 297:119–133.

96. Mohan, A., C. J. Oldfield, P. Radivojac, V. Vacic, M. S. Cortese, A. K. Dunker, and V. N. Uversky. 2006. Analysis of molecular recognition features (MoRFs). *J. Mol. Biol.* 362:1043–1059.

97. Brown, C. J., S. Takayama, A. M. Campen, P. Vise, T. W. Marshall, C. J. Oldfield, C. J. Williams, and A. K. Dunker. 2002. Evolutionary rate heterogeneity in proteins with long disordered regions. *J. Mol. Evol.* 55:104–110.

98. Lise, M. F., and A. El-Husseini. 2006. The neuroligin and neurexin families: from structure to function at the synapse. *Cell. Mol. Life Sci.* 63:1833–1849.

99. Dosztanyi, Z., J. Chen, A. K. Dunker, I. Simon, and P. Tompa. 2006. Disorder and sequence repeats in hub proteins and their implications for network evolution. *J. Proteome Res.* 5:2985–2995.

100. Patil, A., and H. Nakamura. 2006. Disordered domains and high surface charge confer hubs with the ability to interact with multiple proteins in interaction networks. *FEBS Lett.* 580:2041–2045.

101. Ekman, D., S. Light, A. K. Bjorklund, and A. Elofsson. 2006. What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome Biol.* 7:R45.

102. Haynes, C., C. J. Oldfield, F. Ji, N. Klitgord, M. E. Cusick, P. Radivojac, V. N. Uversky, M. Vidal, and L. M. Iakoucheva. 2006. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput Biol.* 2:e100.

103. Singh, G. P., and D. Dash. 2007. Intrinsic disorder in yeast transcriptional regulatory network. *Proteins.* 68:602–605.

104. Selenko, P. 2007. The protein disorder paradox: what do natively unfolded proteins look like inside living cells? Abstracts of EMBO Workshop on intrinsically unfolded proteins: biophysical characterization and biological significance. Budapest, May 2007. 2034.

105. Tompa, P., J. Prilusky, I. Silman, and J. L. Sussman. 2007. Structural disorder serves as a weak signal for intracellular protein degradation. *Proteins.* 71:903–909.

106. Wright, P. E., and H. J. Dyson. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293:321–331.

107. Kandel, E. R. 2001. The molecular biology of memory storage: a dialogue between genes and synapses. *Science.* 294:1030–1038.

108. Scholl, F. G., and P. Scheiffele. 2003. Making connections: cholinesterase-domain proteins in the CNS. *Trends Neurosci.* 26:618–624.