The Transcriptional Repressor RYBP Is a Natively Unfolded Protein Which Folds upon Binding to DNA[†]

José L. Neira,*[,‡,§] Mónica Román-Trufero,[||] Lellys M. Contreras,[⊥,#] Jesús Prieto,[▽] Gagandeep Singh,[‡,○] Francisco N. Barrera,[‡,@] M. Lourdes Renart,[‡] and Miguel Vidal*[,||]

Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, 03202 Elche (Alicante), Spain, Biocomputation and Complex Systems Physics Institute, 50009 Zaragoza, Spain, Centro de Investigaciones Biológicas, CSIC, 28040 Madrid, Spain, Departamento de Biología, Facultad Experimental de Ciencias y Tecnología, Universidad de Carabobo, 2001 Valencia, Venezuela, and Structural Biology and Biocomputing Programme, Centro Nacional de Investigaciones Oncológicas (CNIO), 28007 Madrid, Spain

Received October 15, 2008; Revised Manuscript Received December 10, 2008

ABSTRACT: RYBP (Ring1A and YY1 binding protein) is a zinc finger protein with an essential role during embryonic development, which binds transcriptional factors, Polycomb products, and mediators of apoptosis, suggesting roles in, apparently, unrelated functions. To investigate mechanisms underlying its association with functionally diverse partners, we set out to study its structural properties using a number of biophysical (fluorescence, circular dichroism, Fourier transform infrared, and NMR spectroscopies) and hydrodynamic (analytical ultracentrifugation, DOSY-NMR, and gel filtration chromatography) techniques. We find RYBP to be a noncompact protein with little residual secondary structure, lacking a well-defined tertiary structure. These observations are also supported by theoretical calculations using neural networks and pairwise energy content, suggesting that RYBP is a natively unfolded protein. In addition, structural studies on its binding to the C-terminal region of the Polycomb protein Ring1B or to DNA show conformational changes in the complexed RYBP, consistent with the acquisition of a folded structure. The data provide a structural explanation for RYBP engagement in functionally unrelated pathways by means of its assembly into various macromolecular complexes as an unstructured protein with the ability to acquire a well-structured fold due to its association with different partners.

RYBP[1] (Ring1A and YY1 binding protein) was first characterized as an interacting partner of the Polycomb group (PcG) protein Ring1A (*1*). RYBP acts as a repressor in transcriptional assays that use reporter constructs (*1*) and is found as a subunit of complexes that contain Ring1A and its paralog Ring1B (*2*), although some preparations of the canonical Polycomb repressive complex I do not contain RYBP (*3*). RYBP is a zinc finger protein that reportedly binds ubiquitylated H2A (*4*), a Ring1A/Ring1B-dependent chromatin mark associated with transcriptional repression (*5*). Despite its ability to associate with Polycomb products, there is still no evidence for a Polycomb function for RYBP. Genetic analysis in the mouse shows that the RYBP locus is essential for development, and mutagenesis through knocking-in a gene encoding a green fluorescent protein results in an embryonic lethal mutation (*6*). Mutations of the RYBP ortholog in the fly *Drosophila melanogaster*, dRYBP, lead to a number of alterations, including lethality, but not the homeotic-like phenotypes characteristic of mutations in PcG genes (*7*). Overexpression studies in *Drosophila*, however, show homeotic phenotypes that depend on the C-terminal region of dRYBP (*7*), the region expected to associate with Ring1 proteins (*1*). Additional RYBP functions in transcriptional control lie on its ability to

[†] This work was supported by Grants SAF2008-05742-C02-01 and CSD-2008-00005 (J.L.N.) and Grants SAF2004-06952-C02-01 and SAF2007-65957-C02-01 (M.V.) from the Spanish Ministerio de Innovación y Ciencia, and the Oncocycle program from the Comunidad de Madrid (M.V.). Financial support from FIPSE (Exp 36557/06) to J.L.N. is also gratefully acknowledged. G.S. was supported by a short-term EMBO fellowship. M.R.-T., F.N.B., and M.L.R. were recipients of Ph.D. fellowships from Spanish Ministerio de Educación y Ciencia.

* Corresponding authors addresses: José L. Neira, Instituto de Biología Molecular y Celular, Edificio Torregaitán, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202, Elche (Alicante), Spain. Tel: + 34 966658459. Fax: + 34 966658758. E-mail: jlneira@umh.es. Miguel Vidal, Centro de Investigaciones Biológicas, Ramiro de Maeztu, 9, 28040 Madrid, Spain. Tel: +34 918373112, Fax: +34 91 536 0432. E-mail: mvidal@cib.csic.es.

[‡] Universidad Miguel Hernández.

[§] Biocomputation and Complex Systems Physics Institute.

[||] CSIC.

[⊥] Universidad de Carabobo.

[#] On sabbatical leave at Universidad Miguel Hernández.

[▽] Centro Nacional de Investigaciones Oncológicas (CNIO).

[○] Current address: Department of Chemistry, Lyallpur Khalsa College, Jalandhar 144001, India.

[@] Current address: Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Ave., New Haven, CT 06520.

[1] Abbreviations: ANS, 8-anilinonaphthalene-1-sulfonic acid; CD, circular dichroism; C-Ring1B, C-terminal region of Ring1B (residues 227−334 of the intact protein); DED, death effector domain; DOSY, diffusion-ordered spectroscopy; *D*, translational self-diffusion coefficient; FTIR, Fourier transform infrared; GST, glutathione *S*-transferase protein; GdmCl, guanidine hydrochloride; NMR, nuclear magnetic resonance; Pc, Polycomb; PcG, Polycomb group; Ring, really interesting new gene; RYBP, Ring1 and YY1 binding protein; TSP, 3-(trimethylsilyl)propionic acid-2,2,3,3-$^2H_4$-sodium salt; UV, ultraviolet.

10.1021/bi801933c CCC: $40.75
Published on Web 01/26/2009

associate directly with DNA binding proteins, such as the members of the E2F family of transcription factors and with YY1 (*8*).

Interestingly, RYBP also interacts directly with death effector domain-containing proteins such as FADD, Procaspase 8, and Procaspase 10, which participate in signaling pathways that mediate death receptor-induced apoptosis (*9*). RYBP reportedly interacts with two other polypeptides with roles in apoptosis, the viral apoptin protein (*10*) and Hippi (*11*). Gain of function of RYBP experiments in tissue culture cells is consistent with a proapoptotic function for RYBP (*9, 11*). In addition to embryo lethality, other developmental alterations associated with either loss- or gain-of-function experiments include defects in neural tube closure and formation of anterior eye structures (*12, 13*). These pleiotropic effects are likely the consequence of the involvement of RYBP in a variety of functional pathways, as suggested from its association with proteins with roles in transcription regulation and apoptosis. In this regard, RYBP appears as another example of multitasking, or "moonlighting", proteins with an ability to participate in several unrelated functions.

We have set out to describe structurally PcG proteins and their interacting partners. Recently, we have described the structural conformational properties of the C-terminal domain of Ring1B (C-Ring1B) (*14*), showing that it is a dimer with well-formed monomers. In this work, we first describe the conformational preferences of RYBP. Spectroscopic and hydrodynamic techniques and theoretical calculations suggest that RYBP is a natively unfolded or intrinsically unstructured protein. Second, we have explored whether RYBP structure was affected upon heterotypic interactions with the Ring1B C-terminal region and with DNA; our results indicate that upon binding to either macromolecule RYBP acquires a folded conformation. Interestingly enough, as suggested by CD, the acquired folded conformations with both macromolecules seem to be different.

EXPERIMENTAL PROCEDURES

Materials. Urea and GdmCl (molecular biology grade) were from Sigma; ultrapure urea and GdmCl were from ICN Biomedicals Inc. Exact concentrations of urea and GdmCl were calculated from the refractive index of the solution (*14*). Imidazole, Trizma base, NaCl, the low-molecular protein mass marker, calf thymus DNA, ultrapure dioxan, and ANS were from Sigma. The β-mercaptoethanol was from Bio-Rad, and the Ni^{2+} resin was from Amersham Biosciences (GE Healthcare). Dialysis tubing with a molecular mass cutoff of 3500 Da was from Spectrapore. Standard suppliers were used for all other chemicals. Water was deionized and purified on a Millipore system.

Protein Expression and Purification. RYBP is 228 residues long, and the expression vector used has an additional His_6 tag at the N-terminus, which does not contain any cleavage site. The vector (an in-house modification of the pQE30 one) was a kind gift from M. Bycroft and M. Proctor (Department of Chemistry, University of Cambridge, Cambridge, U.K.). RYBP contains one tyrosine and one tryptophan residue. Recombinant protein was expressed in the *Escherichia coli* C43 strain (*15*) and purified using Ni^{2+} chromatography with a tablet of the complete EDTA-free protease inhibitor cocktail (Roche). An additional gel filtration chromatography step was carried out by using a Superdex 200 16/60 gel filtration column (GE Healthcare) on an AKTA FPLC system (GE Healthcare), following the absorbance at 280 nm; the buffer used was 25 mM phosphate (pH 7.0) with 150 mM NaCl. We observed that RYBP eluted close to the void volume of the column (~44 mL), and since in the SDS gels the protein appeared as a single band, we ran a 2.5% agarose DNA gel. The sample had DNA, and we concluded that DNA from *E. coli* was bound to RYBP. We removed the DNA by dialyzing the sample from the gel filtration step, extensively against pure water; a large amount of protein precipitated in the dialysis tubing, but the fraction in the supernatant after centrifugation was RYBP, as shown by SDS and agarose gels, and mass spectrometry (data not shown). The final yield of RYBP was 8−10 mg/L of culture. Protein concentrations were calculated from the absorbance at 280 nm, using the extinction coefficients of model compounds (*16*). This DNA-free protein was used through all the experiments, except during the DOSY-NMR measurements.

To avoid the loss of protein during the final dialysis step, we tried to remove DNA with DNase I from bovine pancreas (Roche), during the first Ni^{2+} resin purification steps, following the manufacturer's instructions. Under these conditions, the protein was degraded yielding an intense band at 6.4 kDa (data not shown).

The C-Ring1B domain was purified as previously described (*14*).

NMR Spectroscopy. The NMR experiments were acquired on a Bruker Avance DRX-500 spectrometer equipped with a triple-resonance probe and z-pulse field gradients.

(a) One-Dimensional (1D) NMR Spectroscopy. Homonuclear 1D NMR experiments were performed with RYBP concentrations ranging from 60 to 600 μM in 0.5 mL at 293 K, pH 7.0, and 50 mM phosphate buffer (uncorrected for deuterium isotope effects) in a H_2O/D_2O mixture (90%/10%, v/v), or at pH 4.5 (50 mM acetic acid). TSP was used as the external chemical shift reference. The 1D spectra were acquired with 16K data points, averaged over 512 scans with 6000 Hz of spectral width (12 ppm); water suppression was achieved with the WATERGATE sequence (*17*). Baseline correction and zero filling were applied. All spectra were processed and analyzed by using XWINNMR (Bruker GmbH, Karlsruhe, Germany) working on a personal computer.

In the exchange experiments, an RYBP aliquot to give a final concentration of 100 μM in 500 μL was taken from a stock solution, and 2 mL of D_2O was added. The resulting sample was concentrated with Amicon centrifugal devices. Four dilution steps of the sample in D_2O were carried out.

For the binding experiments, 1D NMR spectra of isolated C-Ring1B and RYBP were acquired at a concentration of 60 μM, at pH 7.0, in 50 mM phosphate buffer at 293 K. A 1D NMR spectrum of the complex formed by mixing each protein at a concentration of 60 μM was also acquired. For experiments with calf thymus DNA, a solution of the complex with a final concentration of RYPB of 1.5 μg/μL (~60 μM) and a final concentration of DNA of 1.5 μg/μL was prepared in the same buffer, and the corresponding 1D NMR spectra were recorded.

(b) Translational Diffusion Measurements (DOSY Experiments). Translational self-diffusion measurements were performed with the pulsed-gradient spin−echo. The follow-

ing relationship exists between the translational self-diffusion parameter, D, and the delays during acquisition (*14*):

$$\frac{I}{I_0} = -\exp\left[D\gamma_H{}^2\delta^2G^2\left(\Delta - \frac{\delta}{3} - \frac{\tau}{2}\right)\right] \quad (1)$$

where I is the measured peak intensity of a particular (or a group of) resonance(s), I_0 is the maximum peak intensity of the same resonance(s) at the smaller gradient strength, D is the translational self-diffusion constant (in square centimeters per second), δ is the duration (in seconds) of the gradient, G is the gradient strength (in teslas per centimeter), Δ is the time (in seconds) between the gradients, γ_H is the gyromagnetic constant of the proton, and τ is the recovery delay between the bipolar gradients (100 μs). Data are plotted as the $-\ln(I/I_0)$ versus G^2, and the slope of the resulting line is $D\gamma_H{}^2\delta^2(\Delta - \delta/3 - \tau/2)$, from which D can be easily obtained. The duration of the gradient was varied between 2.2 and 3 ms, and the time between both gradients was changed from 100 to 150 ms. The most upfield-shifted methyl groups (those between 0.2 and 1 ppm) were used to measure the changes in intensity in the spectra of RYBP and their complexes.

To determine the D and the corresponding hydrodynamic radius, we used two different approaches. In the first one, employed only for isolated RYBP, the procedure described in measuring the D of C-Ring1B was used (*14*); briefly, we conducted experiments at different RYBP concentrations, and the infinite dilution translational diffusion coefficient was obtained by the extrapolation at zero protein concentration of the straight line determined by the measurement of D at different protein concentrations (that is, the D at infinite dilution is the y-axis intercept of the straight line). The diffusion coefficients and the hydrodynamic radius in the RYBP complexes (with C-Ring1B, calf thymus DNA, or *E. coli* DNA) were calculated by comparison with these corresponding magnitudes of dioxan, as described by Dobson and co-workers (*18*); dioxan was added to the NMR samples to a final concentration of 2%. For experiments with DNA from *E. coli*, a sample from the gel filtration column step was concentrated with Amicon centrifugal devices and introduced into the magnet, without further purification.

The gradient strength was calibrated by using the D for the residual proton water line in a sample containing 100% D_2O in a 5 mm tube, as described previously (*14*).

(c) Measurements of T_2. Measurements of T_2 (transverse relaxation time) provide a convenient method for determining the molecular mass of a macromolecule, since the correlation time, τ_c, is approximately equal to $1/(5T_2)$ (*19* and references cited therein). We measured the T_2 for RYBP and its complexes with the 1−1 echo sequence (*20*), and the calculation of the τ_c was carried out as described previously (*19*).

Gel Filtration Chromatography. The standards used in column calibration, and their corresponding Stokes radii, R_S, were ovalbumin (30.5 Å), bovine serum albumin (35.5 Å), aldolase (48.1 Å), ferritin (61 Å), and thyroglobulin (85 Å) (*21*). Samples were loaded in 50 mM sodium phosphate (pH 7.0) and 150 mM NaCl (to avoid interactions with the column) in a calibrated Superdex 200 HR FPLC column (GE Healthcare). The void volume was determined from blue dextran and the bed volume from conductivity measurements. Samples were eluted at 1 mL/min and continuously monitored with an online detector at a wavelength of 280 nm. All experiments were conducted at 298 K. Gel filtration chromatography was used to determine the R_S of RYBP, as described previously (*22*). RYBP protein concentrations ranged from 10 to 60 μM; no variations were observed either in the elution volumes or in the width of the peaks.

In binding experiments with RYBP and C-Ring1B, six samples were prepared at pH 7 (50 mM, phosphate buffer) with 150 mM NaCl. The first one contained 10 μM RYBP; the second contained 10 μM (in protomer units) C-Ring1B, and the final one contained each protein at 10 μM. The other set of samples was composed of the same proteins and complexes at a concentration of 40 μM; no differences were observed at either concentration (data not shown). For the binding experiments with RYBP and DNA, we could not form complexes with equimolar stoichiometry since the exact molecular mass of the DNA is unknown; then, three samples were prepared at pH 7.0 (50 mM, phosphate buffer) with 150 mM NaCl, with the same amount in micrograms per microliter of RYBP and DNA. The first one contained 0.25 μg/μL (~10 μM) RYBP; the second one contained 0.25 μg/μL calf thymus DNA, and the third one contained each biomolecule at 0.5 μg/μL. Samples were left to incubate overnight at 278 K. Binding experiments were conducted at 298 K.

Fluorescence. Fluorescence spectra were recorded on a Cary Varian spectrofluorimeter, interfaced with a Peltier unit, at 298 K. The sample concentration was 2 μM in the pH and chemical denaturation experiments; the final concentrations of the buffers were, in all cases, 10 mM. The material for the experiments was prepared the day before and left overnight at 278 K. Control experiments were also performed after a 2 h incubation time; no differences were observed versus those prepared the day before. A 1 cm path length quartz cell (Hellma) was used.

(a) Steady-State Spectra. Protein samples were excited at 280 and 295 nm in the pH range of 2.0−12.0 to characterize a possible different behavior of tryptophan or tyrosine residues (*23*). Chemical denaturation experiments were conducted via excitation at 280 nm, because no differences were observed between the spectra obtained by excitation either at 280 nm or at 295 nm (data not shown). The slit widths were typically equal to 5 nm for the excitation and emission light. The fluorescence experiments were conducted between 300 and 400 nm. The signal was acquired for 1 s, and the increment of wavelength was set to 1 nm. Blank corrections were made in all spectra.

We conducted the chemical denaturations, followed by either fluorescence or CD, by diluting the proper amount of the 8 M urea stock solution and leaving the samples overnight at 298 K. Experiments were also conducted in 6 M GdmCl; all the denaturations were reversible. In the pH-induced unfolding curves, the pH was measured after completion of the experiments with an ultrathin Aldrich electrode in a Radiometer (Copenhagen) pH-meter. The salts and acids used were as follows: pH 2.0−3.0, phosphoric acid; pH 3.0−4.0, formic acid; pH 4.0−5.5, acetic acid; pH 6.0−7.0, NaH_2PO_4; pH 7.5−9.0, Tris acid; pH 9.5−11.0, Na_2CO_3; pH 11.5−13.0, Na_3PO_4.

The wavelength-averaged emission intensity, $\langle\lambda\rangle$, was calculated as described previously (*14*).

(b) Fluorescence Quenching. Quenching of intrinsic fluorescence by iodide was examined under different solution conditions. Excitation was at 280 and 295 nm; emission was measured from 300 to 400 nm. The ionic strength was kept constant by addition of KCl, when GdmCl was not present; also, $Na_2S_2O_3$ was added to a final concentration of 0.1 M to avoid formation of I_3^-. The presence of KCl did not modify the structure of RYBP, which one can conclude from the fact that (i) the maximal wavelengths of spectra during the quenching experiments did not change and (ii) the shape and ellipticity of the CD spectra in the presence of KCl (0−1 M) did not change (data not shown). The slit width was set at 5 nm for both excitation and emission light. The data were fitted to (*24*)

$$F_0/F = 1 + K_{sv}[X] \quad (2)$$

where K_{sv} is the Stern−Volmer constant for collisional quenching, F_0 is the fluorescence in the absence of KI, and F is the measured fluorescence at any KI concentration. The range of KI concentrations explored was 0−0.7 M.

For the experiments in the presence of acrylamide, the experimental parameters were the same as in KI, but the Stern−Volmer equation was modified to include an exponential term to account for dynamic quenching (*24*). Acrylamide quenching experiments at different protein concentrations (ranging from 2 to 10 μM) were carried out. No significant differences were observed among the measured K_{sv} values [ranging from 7.7 ± 0.5 M^{-1}, at the lower concentration explored (2 μM), to 7.5 ± 0.2 M^{-1} at 10 μM].

Circular Dichroism. Circular dichroism spectra were recorded on a Jasco J810 spectropolarimeter fitted with a thermostated cell holder and interfaced with a Peltier unit. The instrument was periodically calibrated with (+)-10-camphorsulfonic acid.

(a) Far-UV Steady-State Spectra. Isothermal wavelength spectra of RYBP at different pHs or urea concentrations were acquired at a scan speed of 50 nm/min with a response time of 2 s and averaged over four scans at 298 K. Measurements were performed with 10 μM RYBP in 10 mM buffer, in 0.1 cm path length cells. Spectra were corrected by subtracting the baseline in all cases. The chemical denaturation experiments were repeated at least three times with new samples. The material for the experiments was prepared the day before and left overnight at 278 K.

For the binding experiments, the steady-state CD spectra were acquired with the same set of parameters as described above. For the experiment testing binding between RYBP and C-Ring1B, three samples were prepared at pH 7.0 (10 mM, phosphate buffer). The first one contained 20 μM RYBP; the second contained 20 μM (in protomer units) C-Ring1B, and the final one contained each protein at 20 μM. For the binding experiments with RYBP and calf thymus DNA, three samples were prepared at pH 7.0 (10 mM, phosphate buffer) with 150 mM NaCl. The first one contained 0.5 μg/μL (~20 μM) RYBP; the second one contained 0.5 μg/μL calf thymus DNA, and the third one contained each biomolecule at 0.5 μg/μL. Samples were left to incubate overnight at 278 K, and spectra were recorded at 293 K. All spectra were corrected by subtracting the proper baseline. The mean residue ellipticity, [θ], and the helical content of RYBP were calculated as described previously (*22*).

To obtain the far-UV CD spectra of the RYBP bound to either DNA or C-Ring1B, we subtracted the spectra of the corresponding complexes, the spectra of free DNA or free C-Ring1B, respectively.

(b) Thermal Denaturation Experiments. Thermal denaturation experiments were performed at constant heating rates of 60 K/h with a response time of 8 s. Thermal scans were collected in the far-UV region at 222 nm from 298 to 363 K in 0.1 cm path length cells with a total protein concentration of 10 μM. Solution conditions were the same as those reported for the steady-state experiments. The possibility of drifting of the CD spectropolarimeter was tested by running two samples containing buffer, before and after the thermal experiments; no difference was observed between the scans. The thermal denaturations were fully reversible at any pH.

Fourier Transform Infrared. Samples of RYBP in 50 mM phosphate buffer (pH 7) were dried in a Speed Vac concentrator (Savant, Farmingdale, NY) and resuspended in a final volume of 22 μL of D_2O. The samples were placed amid a pair of CaF_2 windows separated by a 50 μm thick Mylar spacer in a Harrick (Ossining, NY) demountable cell. Experiments were performed with 1 mM RYBP. Spectra were recorded on a Bruker FTIR instrument equipped with a DTGS detector and thermostated with a Braun water bath at the desired temperature. The cell container was continuously filled with dry air. Usually, 600 scans per sample were taken, averaged, apodized with a Happ−Genzel function, and Fourier transformed to give a final resolution of 2 cm^{-1}. The contributions of the buffer spectra were subtracted, and the resulting spectra were used for analysis, as described previously (*25*). The error in the estimation of the percentage of secondary structure depends mainly on the removal of spectral noise, which was estimated to be 5% (*25*).

Analytical Ultracentrifugation. The sedimentation velocity experiments were conducted in an XL-A analytical ultracentrifuge (Beckman-Coulter Inc.) at 42000 rpm and 293 K, using an A 50Ti rotor and 3 mm charcoal-filled Epon double-sector centerpiece. Absorbance was measured at 280 nm. The protein concentration was 104 or 208 μM in 50 mM Tris buffer (pH 7.0); both concentrations yielded the same results. Data were modeled as a superposition of Lamm equation solutions with SEDFIT (available at www.analyticalultracentrifugation.com/default.htm) (*26*). The sedimentation coefficient distribution, $c(s)$, was calculated at a $p = 0.68$ confidence level. The experimental sedimentation values were determined by integration of the main peak of $c(s)$ and corrected to standard conditions to obtain the $s_{20,w}$ values with SEDNTERP (*27*), assuming a continuous sedimentation coefficient distribution. Calculation of the frictional coefficient ratio was performed with SEDFIT to yield the $c(M)$ distribution (*26*). Experiments were also conducted at pH 5.3 (50 mM acetic acid buffer) and pH 9.0 (50 mM Tris buffer). The results were similar at the three pH values (data not shown).

Analysis of the Amino Acid Sequence of RYBP. The protein sequence was submitted to the PONDR server (www.pondr.com), and analyses were performed using the neural network predictor VL-XT and CDF (*28, 29*) [access to PONDR was provided by Molecular Kinetics, 6201 La Pas Trail-Ste 160, Indianapolis, IN 46268 (main@molecularkinetics.com)]. We also used the IUPRED server (www.iupred.enzim.hu) to allow for a comparison (*30, 31*).

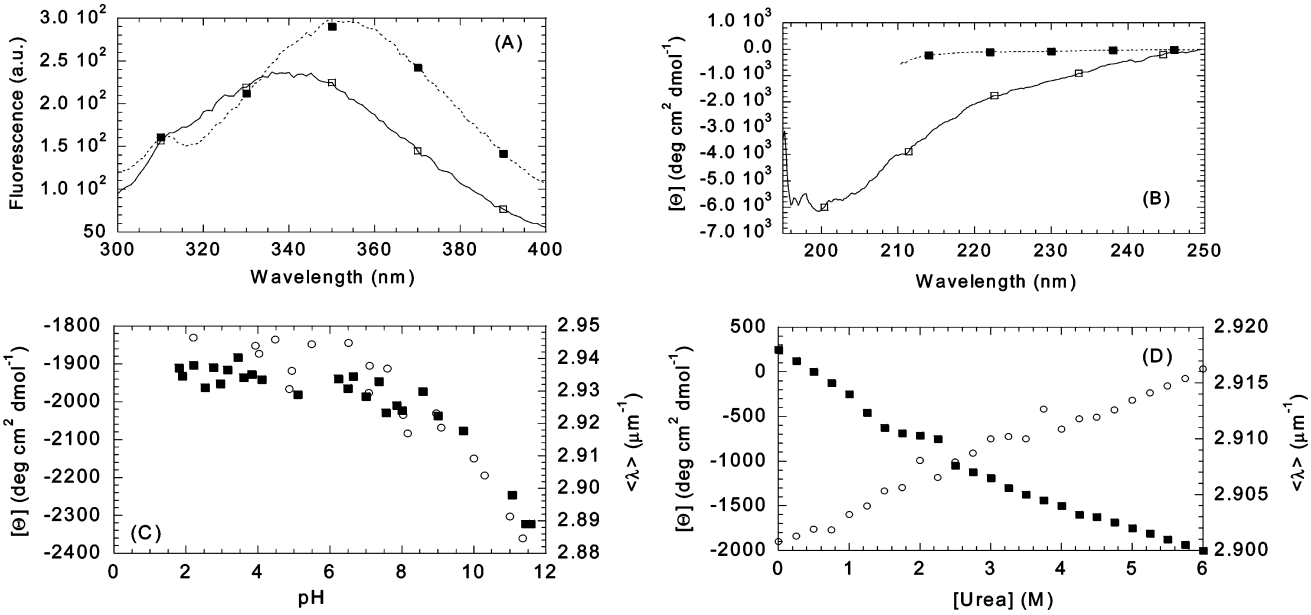

FIGURE 1: Fluorescence and far-UV CD experiments with RYBP. (A) Fluorescence spectrum of RYBP in aqueous solution, in phosphate buffer (—) and in 6 M GdmCl (•••) at pH 7.0. (B) Far-UV spectra of RYBP in aqueous solution, in phosphate buffer (—) and at 6 M GdmCl (•••) at pH 7.0. (C) Changes in fluorescence ($\langle\lambda\rangle$) (right axis, ■) and in the molar ellipticity at 222 nm (left axis, ○) upon pH. (D) Changes in fluorescence ($\langle\lambda\rangle$) (right axis, ■) and ellipticities at 222 nm (left axis, ○) with urea concentration. The conditions were as follows: 2 μM protein, 298 K, for fluorescence experiments and 10 μM for the CD spectra; the buffer concentration was 10 mM in all cases. Spectra were recorded in either 1 cm (fluorescence) or 0.1 cm (CD) path length cells.

GST−Protein Binding Assays. These were performed essentially as previously reported (*1*). Briefly, a Single Tube Protein System 3 (Novagen) was used to transcribe and translate, in the presence of [^{35}S]Met, 500 ng of supercoiled plasmid DNA corresponding to Ring1B cDNAs subcloned into a pCITE 4a vector. Ring1B deletion variants were obtained by PCR-mediated overlap extension (*32*). The integrity of PCR fragments was verified by sequencing. A description of these plasmids is available upon request. After incubation of [^{35}S]Ring1B proteins with either glutathione (GST) or GST−RYBP immobilized onto glutathione (GSH)−agarose beads, bound proteins were eluted and separated on 10% SDS−polyacrylamide gels. Dried gels were analyzed using a Fujifilm FLA-3000 scanner, and protein amounts were evaluated using ImageGauge (Fuji-Film).

RESULTS

RYBP Is a Natively Unfolded Protein. We used experimental and theoretical approaches to test the structure of RYBP. Within the experimental approach, we employed spectroscopic (namely, fluorescence, FTIR, CD, and NMR) and hydrodynamic (analytical ultracentrifugation, gel filtration, and DOSY) techniques.

(a) Spectroscopic Measurements. (1) Fluorescence. (i) Steady-State Fluorescence. We recorded fluorescence spectra to map the tertiary structure of the protein (*23*). RYBP has a sole tryptophan at position 25 and one tyrosine at position 70. The emission fluorescence spectrum of RYBP at pH 7.0, recorded by excitation at 280 or 295 nm, is red-shifted, with a maximum at 340 nm (Figure 1A, □ and solid line). This indicates that the tryptophan is partially solvent-exposed [in an aqueous environment, the maximum should be at ca. 350 nm (*23*)]. The $\langle\lambda\rangle$ did not vary as the pH changed between 2.0 and 8.0 (Figure 1C, ■, right axis), suggesting that the conformation of the protein was not modified. At high pH values, however, the $\langle\lambda\rangle$ decreased probably because of structural rearrangements due to titration of basic residues or, alternatively, basic hydrolysis (Figure 1C).

Table 1: Quenching Parameters in KI and Acrylamide[a]

	280 nm		295 nm	
conditions	K_{sv} (M^{-1}) (KI)	K_{sv} (M^{-1}) (acrylamide)	K_{sv} (M^{-1}) (KI)	K_{sv} (M^{-1}) (acrylamide)
pH 7.0	1.24 ± 0.09	7.7 ± 0.5	0.6 ± 0.3	2.1 ± 0.4
6 M urea	4.0 ± 0.2	13.0 ± 0.2	3.0 ± 0.3	8.0 ± 0.2
6 M GdmCl	4.4 ± 0.3	12.5 ± 0.2	3.1 ± 0.1	8.1 ± 0.2

[a] Errors are data fit errors to eq 2. The constants were obtained by the fitting of fluorescence intensity at 338 nm vs the concentration of quenching agent [similar constants were obtained by fitting the intensities at 335, 336, and 337 nm (data not shown)]. Experiments were conducted at 298 K, pH 7.0, and 10 mM phosphate buffer.

(ii) Examination of Tryptophan Exposure by Fluorescence Quenching. To further examine whether there is tertiary structure around Trp25 and Tyr70 at pH 7.0, we studied iodide and acrylamide quenching in the presence and absence of denaturants (Table 1). In general, the K_{sv} parameter was larger at 280 nm than at 295 nm, due to excitation of both Tyr and Trp at the former wavelength. Further, the K_{sv} values were larger in acrylamide than in KI, as observed in other proteins (*22*). The K_{sv} parameter increased in the presence of chaotropic agents. These results indicate that tryptophan was solvent-exposed (as suggested from the maximum wavelength), but it was not fully accessible in aqueous solution at physiological pH.

(iii) Chemical Denaturation. In the presence of urea or GdmCl, the fluorescence spectra showed an increase in the intensity as the concentration of denaturant varied, with small gradual changes in the maximal wavelengths (Figure 1A, dotted line, ■). Linear behavior was observed in $\langle\lambda\rangle$ (Figure 1D, ■, right axis), which is expected for solvent-exposed aromatic rings (*23*).

(2) Circular Dichroism Experiments. (i) Far-UV CD. We used far-UV CD in the analysis of the unfolding of RYBP as a spectroscopic probe sensitive to protein secondary structure (*33*). The CD spectrum of RYBP at pH 7.0 exhibited a minimum negative ellipticity at ca. 200 nm, characteristic of random-coil or denatured proteins (*33*) (Figure 1B, □). There was a shoulder at 222 nm, which suggests the presence of α-helix- or turn-like conformations, but the contribution of aromatic signals at this wavelength cannot be ruled out (*33*). The estimated population of α-helix- or turn-like structures is 5.5% (the molar ellipticity of RYBP at 222 nm is −1918.7 deg cm^2 dmol^{-1}) (*22*).

(ii) Chemical and pH Denaturation. The ellipticity at 222 nm did not change between pH 2.0 and 8.0 (Figure 1C, ○, left axis); changes were observed at very high pHs probably as a consequence of structural rearrangements due to titration of basic residues or, alternatively, to basic hydrolysis (Figure 1C). Conversely, the ellipticity at 222 nm decreased (in absolute values) gradually as the concentration of urea was increased (Figure 1D, ○, left axis), indicating that the denaturation was a noncooperative process. At the highest GdmCl concentration (Figure 1B, ■, dotted line), the shoulder at 222 nm disappeared; the same behavior was observed in 6 M urea (data not shown). These results suggest that the structure of RYBP, under native conditions, has some residual structure, which does not unfold cooperatively; we do not know, however, if this structure is local or involves short-live long-range interactions. Similar findings have been observed in some random coil polypeptide chains (*34*).

(iii) Near UV. We used near-UV CD to detect possible changes in the asymmetric environment of aromatic residues (*33*). The near UV of RYBP was very weak with no intense bands (data not shown). The absence of a near-UV signal could be due to (i) the absence of an asymmetric environment for the aromatic residues or (ii) the low content of aromatic residues (one W, one Y, and four F residues). We favor the first explanation because of the agreement with the results from the other spectroscopic probes.

(iv) Thermal Denaturations. To further investigate the possible presence of residual secondary structure in RYBP, we carried out thermal denaturation experiments followed at 222 nm. The ellipticity at this wavelength did not change in a sigmoidal fashion; rather, it showed a linear behavior (Figure 1 of the Supporting Information). This result suggests the absence of cooperativity, as expected for a noncompact structure.

(3) Fourier Transform Infrared Spectroscopy. FTIR is a powerful method for investigating protein secondary structure. The main advantage in comparison with CD and fluorescence is that FTIR is much more sensitive to the presence of β-structure or random coil. In the case of proteins, structural information can be obtained by analyzing the amide I region of the spectrum (1700−1600 cm^{-1}). The absorbance of this band is mainly due to the stretching vibration of the carbonyl peptide bond, whose frequency is highly sensitive to hydrogen bonding and thus to protein secondary structure (*35*). The band is centered near 1644 cm^{-1}, which is characteristic of nonordered conformations. Amide I band analysis showed maxima centered at 1683, 1672, 1657, 1644, and 1622 cm^{-1}. The 1622 cm^{-1} band is assigned to (π, 0) β-sheet structure (*36*). The 1644 cm^{-1} band is assigned to random coil structure (*37*). The 1657 cm^{-1} band is assigned to α-helix or disordered structure (*38, 39*). The 1672 cm^{-1} component is assigned to turns and loops (*36*), and also to (0, π) β-sheet vibration band (*38, 39*). The 1683 cm^{-1} band includes contributions from turns. The percentages of secondary structure calculated from the area of the fitted bands are listed in Table 2.

Table 2: Secondary Structural Analysis of RYBP As Determined by FTIR at 298 K and pH 7.0[a]

wavenumber (cm^{-1})	structural assignment	% of total secondary structure
1683	turns	2.9
1672	turns/loops/(0, π) β-sheet	11.5
1657	loops/disordered/α-helix	24.1
1644	random coil	48.3
1622	(π, 0) β-sheet structure	13.2

[a] Errors in the wavenumber are estimated to be ±2 cm^{-1}.

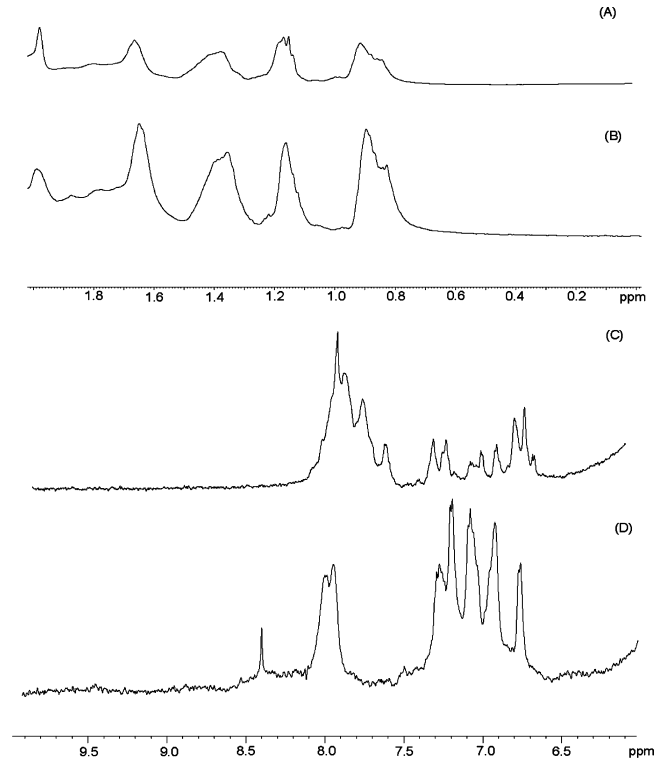

FIGURE 2: NMR spectra of RYBP at 293 K. The methyl region of RYBP is shown at pH 7 in 50 mM phosphate buffer (A) and at pH 4.5 in 50 mM acetate buffer (B). The amide region of RYBP is shown at pH 7.0 in D_2O with 50 mM phosphate buffer (C) and at pH 4.5 in D_2O with 50 mM acetate buffer (not corrected by isotope effects) (D). The peaks between 7.7 and 8.0 ppm correspond to the aromatic side chains of the His tag, and the sole histidine of RYBP.

(4) Nuclear Magnetic Resonance. NMR can give information about the general fold of a polypeptide chain in solution at the residue level. The 1D NMR spectrum of RYBP at 298 K did not show chemical shift dispersion in the amide, aromatic, or methyl regions (Figure 2). In those regions, all the resonances were clustered together as expected for random coil proteins (*40*), namely, between 8.0 and 8.7 ppm (for the amide signals), between 6.8 and 7.5 ppm (for the aromatic signals) (Figure 4A of the Supporting Information), and between 0.8 and 1.1 ppm (for the methyl protons) (Figure 2A,B). These results indicate that there is not stable tertiary structure in RYBP, but we cannot rule out the presence of

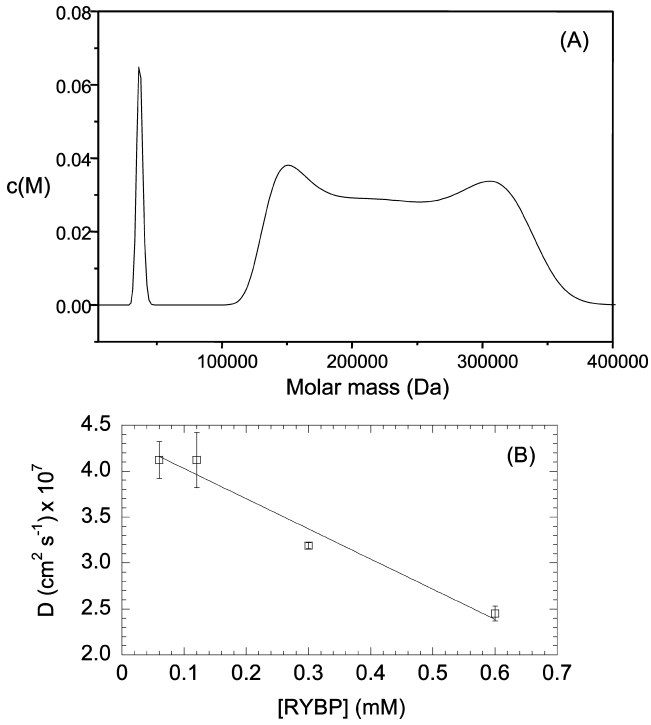

FIGURE 3: Hydrodynamic measurements of RYBP. (A) Analytical ultracentrifugation of RYBP at 108 μM protein at 293 K. (B) DOSY-NMR experiments. The NMR diffusion coefficients of the protein as a function of protein concentration are shown. The bars are fitting errors to eq 1. The solid line is the fitting to a linear equation whose *y*-axis intersection yields the diffusion coefficient at 0 M protein, used in calculation of the hydrodynamic radius.

secondary structure, as suggested by the shoulder at 222 nm in CD experiments (Figure 1B, solid line, □).

Interestingly enough, the indole proton of the sole tryptophan residue did not appear in H_2O at pH 7.0 (Figure 4A of the Supporting Information) or at pH 4.5 (data not shown). Since it does not seem to be a problem of hydrogen exchange due to the basic conditions (because it was not observed at acidic pHs, either), a possible explanation is the existence of a conformational exchange, involving the indole moiety (see Discussion). The shape and line width of the spectra did not change in the concentration range from 60 to 600 μM, suggesting that the protein had the same aggregation state in that range. However, the line widths were somewhat larger than that expected for a 228-residue long protein (Figure 2A,B), suggesting the presence of higher-order molecular species, or alternatively a conformational exchange.

In the hydrogen−deuterium exchange experiment, all the amide protons disappeared within 5 min either at pH 7.0 (Figure 2C) or at pH 4.5 (Figure 2D). Only the aromatic protons of the protein and the His tag could be observed. These results suggest that no stable hydrogen-bonding structure is present in RYBP.

(b) Hydrodynamic Measurements. (1) Gel Filtration. RYBP eluted at pH 7.0 as two peaks at 14.91 and 18.25 mL (Figure 6B). The fact that there are two peaks indicates the presence of an equilibrium (with a slow interconversion kinetic constant) between at least two RYBP conformations. The peak at 18.25 mL is very close to the bed volume (19.64 mL), suggesting that it is probably due to column−protein interactions, precluding any further interpretation. The peak eluting at the shorter volume yielded a Stokes radius of 22.7 Å (Figure 2 of the Supporting Information). Since the inherent precision of the molecular radius determinations of globular proteins by gel partitioning methods has been shown to be ±0.2 Å (*22*), it can be concluded that the measured value is larger than that expected from a 228-residue long globular protein (which is 18.7 Å, obtained from the ideal radius for an unsolvated spherical 228-residue long molecule, R_0: $R_0 = (3M\bar{V}/4N\pi)^{1/3}$, where $\bar{V}$ is the specific volume (0.6919 cm^3/g for RYBP), N is Avogadro's number, and M is the molecular mass of RYBP (24777 Da). There are several possible explanations for this result. (i) Native RYBP is an oligomer. (ii) This peak is also the result of protein−column interactions. (iii) The protein has an elongated shape relative to the globular proteins used as standards. Then, other alternative hydrodynamic techniques are needed to be sure about the shape and molecular mass of the RYBP species present.

(2) Analytical Ultracentrifugation. The results from gel filtration suggest the presence of an elongated protein and/or oligomeric species (and even a column-interacting protein), but we cannot rule out any possibility because this technique cannot distinguish contributions of mass and shape to molecular diffusion. Conversely, analytical ultracentrifugation can be used to determine directly the molar mass of the associated state of macromolecules (*41*). For RYBP, a peak was observed with a molecular mass close to that expected for monomeric RYBP (35 kDa). But also, a broad peak was detected with a molecular mass between 151 and 310 kDa, which accounts for most of the absorbance (Figure 3A). The broadness of this peak indicates sample heterogeneity due to the presence of different high-molecular mass aggregates. This conclusion is further supported by the presence of additional minor peaks at higher sedimentation coefficients (data not shown). Attempts to quantify the underlying cause of heterogeneity by using sedimentation equilibrium failed, because fitting to different equilibria involving several oligomeric species was poor (data not shown).

(3) DOSY-NMR. The hydrodynamic radius was determined from the dilution of a stock solution of RYBP from 60 to 600 μM. The obtained diffusion coefficients form a straight line (Figure 3B), and then, we can be confident that, within that concentration range, RYBP has the same oligomerization state (*14*). The diffusion coefficient and the calculated hydrodynamic radius also suggest either a self-aggregated species or, alternatively, an elongated monomeric species (Table 3).

We can confirm the presence of the high-molecular mass species by calculating the τ_c from the T_2 measurements; this calculation does not rely on a particular assumption about molecular shape, but rather it depends on only the molecular mass of the protein. Since the error in the measurements of T_2 is estimated to be 10% (*42, 43*), the τ_c of isolated RYBP (Table 3) is very large for a 228-residue long protein, which should be 18.6 ns for a monomeric protein [in fact, we should expect an almost linear relationship between the τ_c and the molecular mass of the molecule; $\tau_c = (9.18 \times 10^{-3}/T) \exp(2416/T)N^{0.93}$ (*42*)]. It could be thought that the molecular mass of RYBP (and those of its complexes; see below) might be obtained from the corresponding correlation times (*44*); however, that calculation is only valid when the global correlation time

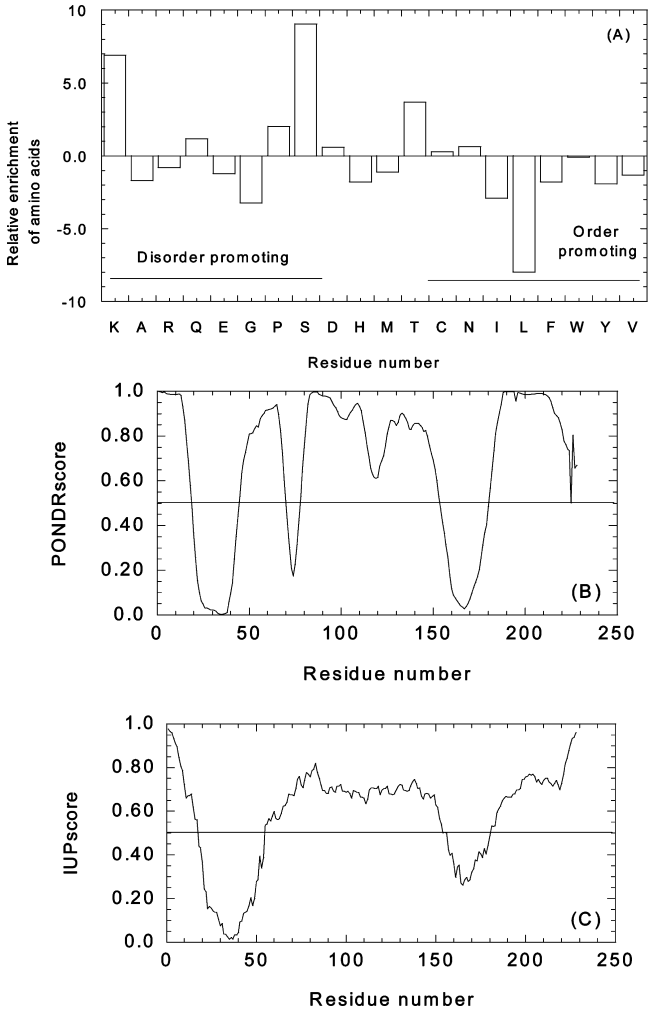

FIGURE 4: In silico analysis of the RYBP sequence. (A) Distribution of order- and disorder-promoting amino acids in the RYBP sequence. The distribution shows a deviation in the amino acid composition of RYBP from the average values in the Swiss-Prot database (as obtained at http://www.expasy.org/sprot/relnotes/relstat.html). (B) PONDR prediction of unstructured regions of the RYBP sequence with the use of VL-XT predictor. The predictor score is plotted against the residue number. The threshold is 0.5, and residues with a higher score are considered disordered. (C) IUPred prediction of unstructured regions (long disorder prediction). The threshold is also 0.5.

Table 3: Hydrodynamic Properties of RYBP and Their Complexes As Determined by NMR at 293 K

conditions	R_s (Å)	D (cm^2 s^{-1})	τ_c (ns)
RYBP	39.5 ± 0.2^a	$(4.3 \pm 0.1) \times 10^{-7}$	24.2
RYBP with DNA (*E. coli*)	47.5 ± 0.4^b	$(3.6 \pm 0.5) \times 10^{-7}$	$-^d$
RYBP with DNA (calf thymus)	47.3 ± 0.4^b	$(3.0 \pm 0.2) \times 10^{-7}$	38.7
RYBP with C-Ring1B[c]	13.3 ± 0.2^b	$(1.3 \pm 0.1) \times 10^{-6}$	30.9

[a] Determined from the Stokes−Einstein equation, taking into account the experimentally determined diffusion coefficient derived from extrapolation at infinite dilution (*14*). [b] The R_s was determined by DOSY measurements using the procedure developed by Dobson and co-workers (*18*). The diffusion coefficient was calculated by using the Stokes−Einstein equation. [c] The final concentration of each protein was 60 μM (in protomer units of each macromolecule). [d] Not determined.

dominates the overall tumbling of the molecule (or its complexes), which could not be necessarily true in RYBP (*43* and references cited therein). In partially unfolded or natively unfolded proteins, where the amide protons are highly mobile, the mobility of some of these amide protons can be higher than that of the overall molecule, and thus, it overrules the tumbling of the whole molecule (*42*).

(c) Theoretical Analysis of the RYBP Sequence. RYBP shares no significant sequence homology with previously reported proteins, as shown by the analysis carried out in the Expasy proteomics server. The protein is significantly enriched in serine, lysine, and other polar amino acids; further, it is depleted in tryptophan (one tryptophan, 0.4%), tyrosine (one tyrosine, 0.4%), phenylalanine (four residues, 1.6%), and other order-promoting residues (Figure 4A). The VL-XT program, the charge hydrophobicity phase space, and CDF (cumulative distribution function) analyses of PONDR (predictor of natural disordered regions) also indicated the unfolded nature of RYBP (Figure 4B, and Figure 3 of the Supporting Information), with ~12% of folded regions, and four main disordered polypeptide patches (residues 1−18, 45−70, 78−153, and 181−228). The region comprising residues 37−47, where the Zn finger is predicted to be, was ordered (i.e., PONDR scores are close to zero). The predictors of PONDR are neural networks that use sequence information from windows, usually comprising 21 amino acids. Particular features of the polypeptide chain, such as the fractional amino acid composition or hydropathy, are calculated over this window, and these values are used as inputs for the predictor. The neural network provides a value for the central amino acid in the window, and the predictions are smoothed over a sliding window of nine amino acids. If a residue value exceeds a threshold of 0.5 (the threshold used for training of the neural network), it is considered disordered.

Similar results were obtained by using the IUPRED web server (Figure 4C), and the agreement between both theoretical approaches is very good. IUPRED recognizes natively unfolded regions from the amino acid sequence based on the estimated pairwise energy content. Globular proteins are composed of amino acids which have the potential to form a large number of energetically favorable interactions, whereas natively unfolded proteins adopt no stable structure since their amino acid composition does not allow formation of sufficient favorable interactions. If a residue value exceeds a threshold of 0.5, it is considered disordered.

Binding and Folding Are Coupled in RYBP When It Is Bound to DNA or C-Ring1B. Many intrinsically disordered proteins undergo conformational transitions to more ordered states, and even fold in well-ordered structures, upon binding to interacting partners (*45*). To investigate whether RYBP also exhibited this property, we decided to study its interaction with PcG Ring1 proteins, since previous studies with Ring1A have shown that RYBP binding occurs through the C-terminal region of Ring1A (*1*). However, given our more extensive knowledge of the homologue structure in Ring1B (*14*), we used this protein, instead of the Ring1A fragment, in the analysis of putative changes in RYBP folding during heterotypic associations.

Therefore, we first set out to study the association of RYBP and Ring1B by using a GST pull-down assay. The analysis (Figure 5A) showed that small deletions in either of the blocks of residues conserved between Ring1A and Ring1B (the C-terminal regions) severely affected binding of RYBP. The results, summarized in Figure 5B, suggest that (i) both subdomains contribute to the overall shape of the C-terminal region of Ring1B, whose intactness is required for RYBP binding, or (ii) such a binding occurs through contacts with

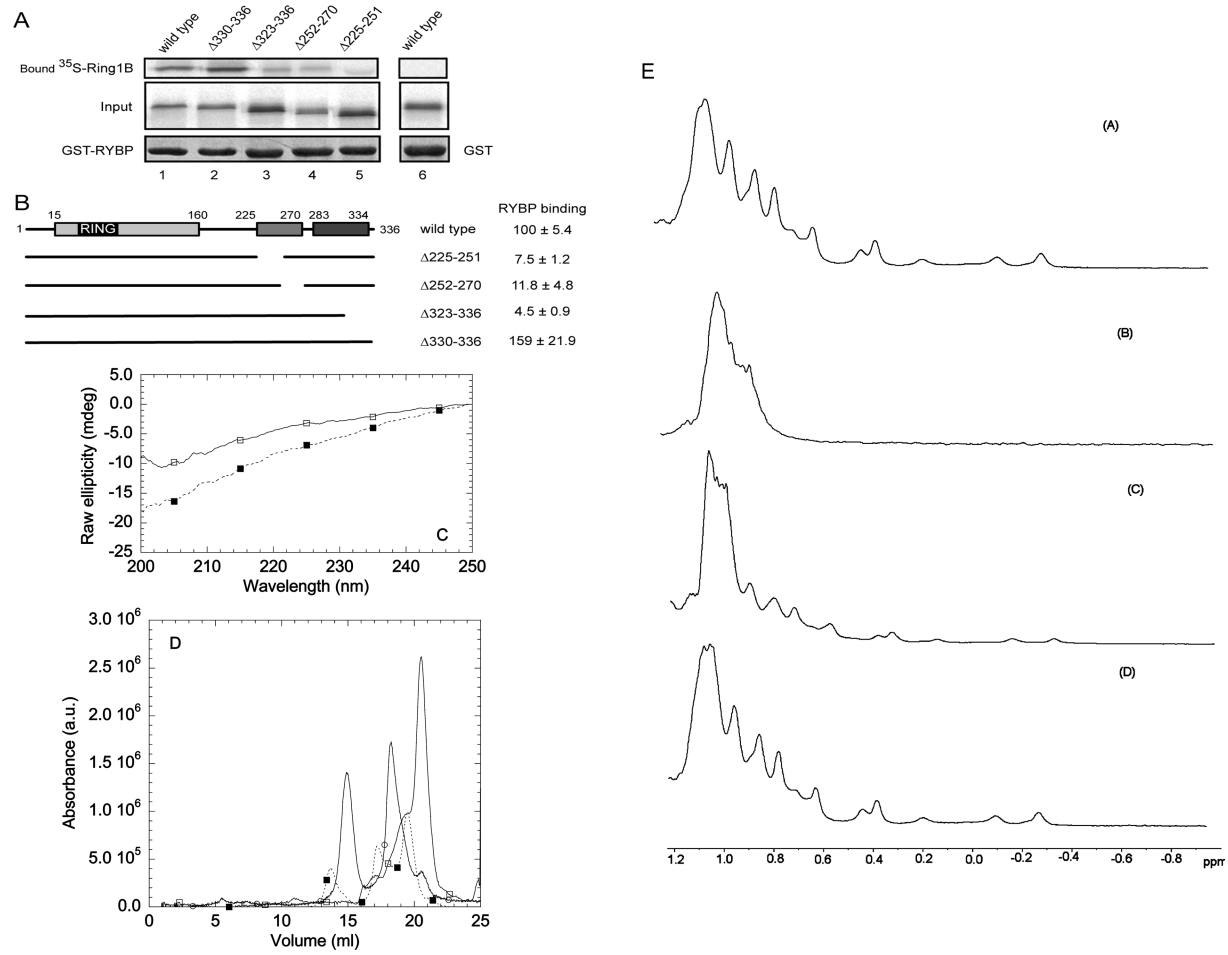

FIGURE 5: Binding of RYBP to C-Ring1B. (A) Intact Ring1B and the indicated truncated variants were synthesized in vitro, radiolabeled with [^{35}S]methionine, and mixed with bacterially produced GST−RYBP protein inmobilized on GSH-Sepharose (lanes 1−5). GST was used as a negative control (lane 6). After incubation and washes, the bound proteins were separated by SDS−PAGE (12% gels) and quantitated using a Fujifilm scanner. The input represents 10% of the total ^{35}S-labeled protein used in the binding assays. (B) Diagram of intact and truncated Ring1B proteins and summary of binding data. Conserved domains between Ring1B and Ring1A proteins are indicated by boxes, one of which contains the RING finger motif. Amino acids are numbered in the various truncated Ring1B proteins. The extent of binding between each Ring1B protein and the GST−RYBP protein is shown as the average ± the standard deviation of triplicate assays relative to that of the wild-type protein (taken as 100). (C) Far-UV CD spectra of the complex formed between equimolar amounts of C-Ring1B and RYBP. The solid line (□) is the spectrum of the macromolecular complex, and the dotted line (■) is that obtained by the addition of the CD spectra of both proteins acquired separately. The spectra are given in millidegrees to allow comparison with DNA binding in the next figure. (D) Gel filtration chromatogram of C-Ring1B with RYBP (with each protein at 10 μM) (dotted line, ■). C-Ring1B eluted closed to the bed volume of the Superdex 200 HR column (solid line, □); RYBP eluted as two peaks (solid line, ○). The units on the *y*-axis are arbitrary, and they have been scaled up for each protein for the sake of the reader. (E) Methyl region of C-Ring1B (60 μM) (A), RYBP (60 μM) (B), and the complex formed by mixing 250 μL of a solution of RYBP (120 μM) and 250 μL of a solution of C-Ring1B (120 μM) and (D) the spectrum obtained adding the spectra of RYBP (B) and C-Ring1B (A). Conditions were 50 mM phosphate buffer, pH 7.0, and 293 K.

both Ring1B subdomains. Then, having established an association between RYBP and the C-terminal region of Ring1B, we used biophysical probes to gain insight into such an association and elucidate whether RYBP acquired a folded conformation upon binding.

First, we used far-UV CD on the complex formed by mixing equimolar amounts of RYBP and C-Ring1B proteins (20 μM). The spectrum of the complex is different from that obtained by the addition of the spectra of both proteins (Figure 5C). Furthermore, assuming that upon binding to C-Ring1B only the linker between both subdomains altered its conformation, as it happens during binding to other proteins (*46*), and therefore its overall contribution to the far-UV spectrum is small, we can conclude that changes in the far-UV spectrum of the complex are mostly due to structural changes in RYBP. The gel filtration experiments not only confirm the binding (because of the small changes in the elution peaks of both proteins upon binding) but also suggest that the association kinetic constant of the RYBP−C-Ring1B complex is slow within the time of the column (Figure 5D).

In the NMR experiments, we could observe that the binding between both proteins did not induce significant changes in the spectrum of C-Ring1B, since the most upfield shifted protons of C-Ring1B (namely, those at −0.28, −0.10, and 0.20 ppm) did not vary (thus supporting our previous hypothesis that upon binding, C-Ring1B suffered only small conformational rearrangements) (Figure 5E, spectrum C). Conversely, those methyl protons appearing at ∼0.9 ppm showed significant changes (Figure 5E, spectrum C); moreover, we could also observe changes in the amide region (Figure 4 of the Supporting Information). We further explored binding of both proteins by calculating the translational diffusion coefficient and the transversal relaxation

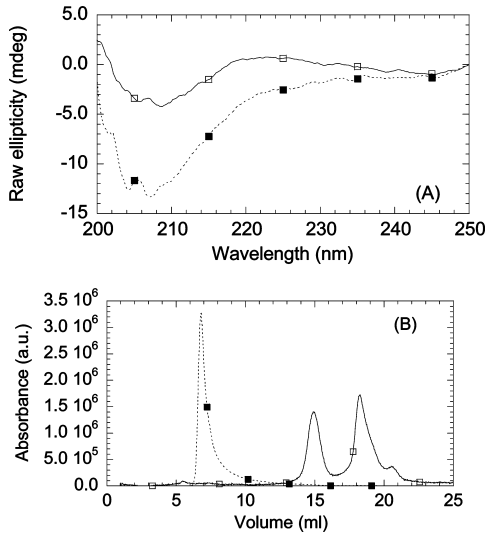

FIGURE 6: Binding of RYBP to DNA. (A) Far-UV CD spectra of the complex formed between the same amounts (in micrograms per microliter) of calf thymus DNA and RYBP, where the final concentration was 1.5 μg/μL for each macromolecule. The solid line (□) is the spectrum of the complex, and the dotted line (■) is that obtained by the addition of the CD spectra of both biomolecules acquired separately. The spectra are given in millidegrees since the molecular mass of the used DNA was unknown. (B) The gel filtration chromatogram of RYBP alone at 10 μM protein (solid line, □) is the same as in Figure 5D, in the presence of the same amount (in micrograms per microliter) of calf thymus DNA (dotted line, ■).

time (T_2) of the complex (Table 3). Since D and T_2 of the complex are different from those of RYBP (Table 3, first row) and C-Ring1B (*14*), both measurements indicate binding between the two macromolecules. Furthermore, from the small value of D, it is tempting to suggest that the binding occurs with a 1:1 stoichiometry.

During the purification of RYBP from *E. coli* cell extracts, we observed that RYBP was isolated associated with bacterial genomic DNA (see Materials and Methods). These results prompted us to study the binding directly to DNA in vitro. We used the highly nonspecific calf thymus DNA to gain further insight into a putative DNA−RYBP interaction in vitro. We tested the binding by using the same spectroscopic probes as in C-Ring1B. The far-UV spectrum of the complex between DNA-free RYBP and calf thymus DNA clearly shows that there is binding (Figure 6A, solid line, □), since this spectrum is very different from that obtained by the sum of the spectra of both macromolecules (Figure 6A, dotted line, ■). Conversely to what happens with C-Ring1B, the gel filtration experiments with DNA and RYBP showed a single elution peak close to the void volume of the column, instead of the two peaks observed when isolated RYBP was loaded; then, the kinetic association constant of the RYBP−DNA equilibrium is fast within the time of the column, and then, the different populations of RYBP (detected by gel filtration and ultracentrifugation) bind to DNA (Figure 6B). Further support for binding of both macromolecules comes from the differences between the D for the complex (either *E. coli* or calf thymus DNA) and that of free RYBP (Table 3). Moreover, the long D obtained in the presence of *E. coli* DNA (sonicated during purification) suggests the presence of either (i) a highly elongated shape of the RYBP−DNA complex or, alternatively, (ii) self-associated RYBP proteins, which are bound to DNA. However, it is important to indicate here that the 1D NMR spectrum of the complex formed between calf thymus (or *E. coli*) DNA and RYBP showed a slight broadening of the signals, when compared to the spectrum of isolated free RYBP, but there were not displacements in the chemical shifts of any proton. This suggests that the structure formed upon binding to DNA is fluctuating, and it is not detected within the time scale of the NMR. Conversely, the far-UV CD spectrum of the complex shows that RYBP acquires a folded structure because the interconversion between the different structured conformers is fast enough. Similar effects on different spectroscopic time scales have been observed in other highly unstructured proteins (*43, 47*).

We can obtain an estimate of the structure acquired by RYBP upon binding to both different targets by using difference spectroscopy. That is, assuming that the far-UV spectra of either DNA or C-RingB do not change upon binding, the spectra of bound RYBP can be obtained by finding the difference between the spectrum of the corresponding complex and that of DNA or C-Ring1B. The results indicate that the DNA-bound RYBP (Figure 5 of the Supporting Information) is more structured than the C-Ring1B-bound RYBP; the latter seems to be more disordered, and its spectrum is similar to that of the isolated RYBP (Figure 1B). Then, the structures acquired by RYBP upon binding to the different macromolecular targets are completely different.

DISCUSSION

RYBP Is a Natively Unfolded Self-Associating Protein. Both biophysical studies and the theoretical predictions in this work show that the isolated RYBP lacks a well-defined tertiary structure and is mainly a disordered protein, with neither stable hydrogen bonds nor a well-formed core. The lack of dispersion of the amide signals in the NMR spectrum and the absence of hydrogen bonds, as shown by hydrogen exchange experiments, also point to the absence of a stable tertiary structure. In agreement with the data given above, the absence of cooperativity during the thermal and chemical denaturation experiments (followed by CD and fluorescence) also indicates that the tertiary structure of RYBP, if any, is very weak and no hydrophobic core is formed; such noncooperative transitions are typically observed in chemical denaturations of partially folded states devoid of persistent long-range tertiary contacts (*34*). The theoretical predictions of the programs used reinforce the experimental findings (Figure 4), since both predict RYBP to be largely unfolded.

However, there is evidence of residual secondary structure as concluded from FTIR (Table 2) and CD experiments (Figure 1B). Far-UV CD spectra show the presence of secondary structure as supported by the shoulder at 222 nm, which disappears at high denaturant concentrations (Figure 1B). This shoulder could be due to (i) the presence of residual helix- or turn-like structure (5.5%) or (ii) aromatic residues (*33*). Moreover, FTIR also indicates the existence of turn-like conformations, although the percentage (∼25%) is higher than that observed by CD (Table 2). The differences between the values reported by both techniques could be due to (i) the deconvolution procedures (in the case of FTIR), (ii) the presence of aromatic residues (CD), which absorb at 222 nm, or (iii) the empirical character of the used equation in

determining the percentage of secondary structure (CD) (*22*). Whatever the reason, the small amount of secondary structure could be due to transient turn- or helix-like conformations around the zinc finger region, or the other ordered regions. Finally, it is worth noting that FTIR also predicts ~20% β-sheet (Table 2), which could not be confirmed by CD.

All the hydrodynamic techniques indicate that RYBP self-associates. The association seems to be very strong, since aggregated species are observed at concentrations as low as 2 μM, as concluded from the absence of protein concentration dependence in the quenching experiments, and from the elution volume (within a protein concentration range of 10−60 μM). Moreover, cross-linking experiments in vitro in the presence of glutaraldehyde showed high-molecular mass species at concentrations as low as 0.5 μM (data not shown). This self-association hampers ANS binding at any pH (data not shown), since oligomerization would hinder the hydrophobic patches. It is important to indicate that the formation of such self-associated species does not induce significant stable hydrogen-bonded secondary or tertiary structures in RYBP (Figure 2B,C). There are, however, several species with different levels of compactness (and probably with different structures) as shown by the two peaks in the gel filtration experiments, which possibly correspond to some of the high-molecular mass species detected by ultracentrifugation (Figure 3A). These two peaks disappeared in the presence of 6 M GdmCl, suggesting that the different level of compactness and, then, the different flickering structures are disrupted (data not shown). The presence of species with different levels of compactness has been also observed in other natively unfolded proteins (*48, 49*).

RYBP is one of the rare natively unfolded proteins which self-associates (*50, 51*) showing several oligomeric species. This self-association could also hamper the detection of the indole proton, since if the tryptophan is involved in a high-molecular mass complex, its tumbling would be very slow; then, we should observe in the 1D NMR spectra those regions of the protein which are more flexible, which has been described in GST-fused proteins monitored by NMR (*52*), and it would explain the broadness of the signals in the 1D NMR spectra (Figure 2). The self-association of RYBP might hide regions otherwise accessible to the cellular milieu (and also to the proteases, as when the protein binds to DNA; see below), resolving for RYBP one of the most puzzling results in natively unfolded proteins: how to avoid degradation. In addition, intramolecular interactions in RYBP could also hide regions involved in other protein interactions, which happens in the WASP (Wiskott−Aldrich syndrome protein) family (*53*).

Binding and Acquisition of Different Folded Structures Are Coupled in RYBP When Bound to DNA or C-Ring1B. In this work, we have shown that RYBP interacts with the C-terminal region of Ring1B, as expected from previous observations on the association of RYBP with Ring1A, the paralog of Ring1B (*1*). Our biophysical characterization of the interactions shows that RYBP binds to a monomeric species of C-Ring1B. We think that the binding between both species is 1:1 on the basis of the following reasoning. The theoretical Stokes radius of a protein or complex, R_0, is given by $R_0 = (3M\bar{V}/4N\pi)^{1/3}$, where $\bar{V}$ is the specific volume, N is Avogadro's number, and M is the molecular mass of the protein (or complex). Introducing the molecular mass values of RYBP and C-Ring1B into that expression, we obtain a Stokes radius of 21.5 Å, which is larger than that measured by diffusion measurements (Table 3) but smaller than those measured for the RYBP−DNA complexes. The use of the molecular mass corresponding to a dimeric C-Ring1B yields 25 Å, a value still further from the one experimentally measured. Thus, even considering the effect that experimental errors may have introduced into the measured diffusion coefficient, the obtained value is close to that theoretically estimated for a 1:1 complex. Attempts to obtain the translational diffusion parameters and the T_2 values of the complex between RYBP and C-Ring1B at concentrations above 600 μM, where the population of the dimeric species of C-Ring1B would be larger, were hampered due to precipitation during the experimental time and precluded any reliable conclusion. Therefore, although binding of RYPB to dimeric forms of Ring1B cannot be excluded, its association with the monomeric form suggests that this heterotypic association may be subject to regulation through control of Ring1B dimerization via its C-terminal region (*14*).

Previous studies have shown that upon binding of a small fragment of Cbx7, one of the Polycomb paralogs in mammals, the C-Ring1B fragment undergoes a conformational tightening (*46*) mediated by the linker sequence between the two Ring1B subdomains. We do not know if a similar tightening occurs upon binding to the RYBP, although the presence of unmodified upfield protons in the NMR spectra of the RYBP-C-Ring1B complex suggests that the core (conserved subdomains) of C-Ring1B remains unaltered. These findings would be consistent with RYBP and Polycomb proteins sharing a binding interface on C-Ring1B, as shown by in vitro pull-down assays (data not shown). More importantly, changes in the CD spectrum (Figure 5 of the Supporting Information) show that, in contrast to moderate C-Ring1B alterations (Figure 5E), RYBP must undergo large conformational rearrangements, which indicate that RYPB belongs to the category of unstructured proteins which upon binding to interacting partners acquire a folded structure.

Although serendipitously, during purification of recombinant RYBP, we have been able to detect, for the first time, an in vitro interaction between RYBP and DNA, which we have further characterized using biophysical probes. It is likely that such an association is of a nonspecific nature, since it has been observed with *E. coli* and calf thymus DNA. It is conceivable to imagine that binding to DNA occurs through electrostatic interactions between negatively charged phosphate residues in the DNA backbone and positively charged residues of the highly basic protein RYBP. Again, alterations in the far-UV CD spectra of the RYBP−DNA complex, given the low likelihood of DNA shape modification upon RYBP binding, strongly suggest that complexed RYBP has acquired a folded structure. Interestingly enough, the conformation acquired by RYBP upon binding to DNA seems to be more structured than that acquired in the presence of C-Ring1B (Figure 5 of the Supporting Information). Furthermore, indirect evidence of such RYBP folding upon DNA binding is its relative resistance to proteolysis compared to the susceptibility to proteases in the absence of DNA.

The appreciation of the coupling between folding of unstructured RYBP and binding to interacting partners is a

property often found in proteins that like RYBP, associate with a variety of interactors [that is, RYBP is a "hub" protein (*45, 54*)]. Finally, our observation of RYBP binding to DNA opens up important regulatory possibilities in the association of RYBP-containing complexes with chromatin, whereby the RYBP may stabilize association of the complexes to nucleosomes through binding both to monoubiquitylated histone H2A (*4*) and to DNA. Therefore, a putative binding of RYBP to DNA in vivo merits further examination to establish whether it actually plays a role, at least, in the transcriptional functions of the moonlighting protein RYBP (*55*).

ACKNOWLEDGMENT

We thank both reviewers for helpful comments and suggestions. We deeply thank May García, María del Carmen Fuster, and Javier Casanova for excellent technical assistance. We thank Dr. M. Bycroft and Mr. M. Proctor (Department of Chemistry, University of Cambridge) for the kind gift of the vector.

SUPPORTING INFORMATION AVAILABLE

Thermal denaturation of RYBP at pH 7, followed by the changes in ellipticity at 222 nm (Figure 1); calculation of the Stokes radius of RYBP from gel filtration (Figure 2); charge hydrophobicity phase space separation between natively folded (■) and natively unfolded (□) proteins (Figure 3), where RYBP (green diamond) falls within the natively unfolded space, as calculated by PONDR; the amide region at pH 7.0 of RYBP (60 μM) (A), C-Ring1B (60 μM) (B), and the complex formed by mixing 250 μL of a solution of RYBP (60 μM) and 250 μL of a solution of C-Ring1B (60 μM) (C) and the spectrum (D) obtained by adding the spectra of RYBP (A) and C-Ring1B (B) (Figure 4); and the difference spectra of DNA-bound RYBP and C-Ring1B-bound RYBP (Figure 5). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. García, E., Marcos-Gutiérrez, C., Lorente, M. M., Moreno, J. C., and Vidal, M. (1999) RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex and with the transcription factor YY1. *EMBO J. 18*, 3404–3418.
2. Sánchez, C., Sánchez, I., Demmers, J. A. A., Rodríguez, P., Strouboulis, J., and Vidal, M. (2007) Proteomics analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/Jhdm1B histone demethylase and the Bcl6 interacting corepressor. *Mol. Cell. Proteomics 6*, 820–834.
3. Levine, S. S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R. E. (2002) The core of the Polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol. Cell. Biol. 22*, 6070–6078.
4. Arrigoni, R., Alam, S. L., Wamstad, J. A., Bardwell, V. J., Sundquist, W. I., and Schreiber-Agus, N. (2006) The Poycomb-associated protein RYBP is a ubiquitin binding protein. *FEBS Lett. 580*, 6233–6241.
5. Stock, J. K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A. G., and Pombo, A. (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nat. Cell Biol. 9*, 1428–1435.
6. Stanton, S. E., Blanck, J. K., Locker, J., and Schreiber-Agus, N. (2007) RYBP interacts with Hippi and enhances Hippi-mediated apoptosis. *Apoptosis 12*, 2197–2206.
7. González, I., Aparicio, R., and Busturia, A. (2008) Functional characterization of the *dRYBP* gene in *Drosophila*. *Genetics 179*, 1373–1388.
8. Schilisio, S., Halperin, T., Vidal, M., and Nevins, J. R. (2002) Interaction of YY1 with E2Fs, mediated by RYBP, provides a mechanism for specificity of E2F function. *EMBO J. 21*, 5775–5786.
9. Zheng, L., Schickling, O., Peter, M. E., and Lenardo, M. J. (2001) The death effector domain-associated factor plays distinct regulatory roles in the nucleus and cytoplasm. *J. Biol. Chem. 276*, 31945–31952.
10. Danen-van Oorschot, A. A., Voskamp, P., Seelen, M. C., van Miltenburg, M. H., Bolk, M. W., Tait, S. W., Boesen-de Cock, J. G., Rohn, J. L., Borst, J., and Noteborn, M. H. (2004) Human death effector domain-associated factor interacts with the viral apoptosis agonist Apoptin and exerts tumor-preferential cell killing. *Cell Growth Differ. 11*, 564–573.
11. Gervais, F. G., Singaraja, R., Xanthoudakis, S., Guterkunts, C. A., Leavitt, B. R., Metzler, M., Hackam, A. S., Tam, J., Vaillancourt, J. P., Houtzager, V., Rasper, D. M., Roy, S., Hayden, M. R., and Nicholson, D. W. (2002) Recruitment and activation of caspase-8 by the Huntingtin-interacting protein Hip-1 and a novel partner Hippi. *Nat. Cell Biol. 4*, 95–105.
12. Pirity, M. K., Locker, J., and Schreiber-Agus, N. (2005) RYBP/DEDAF is required for early postimplantation and for central nervous system development. *Mol. Cell. Biol. 25*, 7193–7202.
13. Pirity, M. K., Wang, W.-L., Wolf, L. V., Tamm, E. R., Schreiber-Agus, N., and Cvekl, A. (2008) RYBP, a polycomb complex associated protein, is required for mouse eye development. *BMC Dev. Biol.* (doi: 10.1186/1471-213X-7-39).
14. Czypionka, A., de los Paños, O. R., Mateu, M. G., Barrera, F. N., Hurtado-Gómez, E., Gómez, J., Vidal, M., and Neira, J. L. (2007) The isolated C-terminal domain of Ring 1B is a dimer made of stable, well-structured monomers. *Biochemistry 46*, 12764–12776.
15. Miroux, B., and Walker, J. E. (1996) Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol. 260*, 289–298.
16. Pace, C. N., and Scholtz, J. M. (1997) Measuring the conformational stability of a protein. In *Protein Structure* (Creighton, T. E., Ed.) 2nd ed., pp 253−259, Oxford University Press, Oxford, U.K.
17. Piotto, M., Saudek, V., and Sklenar, V. (1993) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661–665.
18. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques. *Biochemistry 38*, 16424–16431.
19. Anglister, J., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1993) Isotope-edited multidimensional NMR of calcineurin B in the presence of the non-deuterated detergent CHAPS. *J. Biomol. NMR 3*, 121–126.
20. Sklenar, V., and Bax, A. (1987) Spin echo water suppression for the generation of pure-phase two-dimensional NMR spectra. *J. Magn. Reson. 74*, 469–479.
21. Carvalho, A. F., Costa-Rodrigues, J., Correia, I., Costa-Pessoa, J. C., Faria, T. Q., Martins, C. L., Fransen, M., Sá-Miranda, C., and Azevedo, J. E. (2006) The N-terminal half of the peroxisomal cycling receptor Pex5p is a natively unfolded domain. *J. Mol. Biol. 356*, 864–875.
22. Muro-Pastor, M. I., Barrera, F. N., Reyes, J. C., Florencio, F. J., and Neira, J. L. (2003) The inactivating factor of glutamine synthetase, IF7, is a "natively unfolded" protein. *Protein Sci. 12*, 1443–1454.
23. Schmid, F. X. (1997) Optical spectrosocopy to characterize protein conformation and conformational changes. In *Protein Structure* (Creighton, T. E., Ed.) 2nd ed., pp 261−297, Oxford University Press, Oxford, U.K.
24. Lakowicz, J. R. (1999) *Principles of fluorescence spectroscopy*, 2nd ed., Plenum Press, New York.
25. Echabe, I., Encinar, J. A., and Arrondo, J. L. R. (1997) Removal of spectral noise in the quantitation of protein structure through infrared band decomposition. *Biospectroscopy 3*, 469–475.
26. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophys. J. 78*, 1606–1619.
27. Laue, T. M. S., Ridgeway, T. M., and Pelletier, S. L. (1992) *Computer-aided interpretation of analytical sedimentation data for proteins*, Royal Society of Chemistry, Cambridge, U.K.
28. Li, X., Romero, P., Rani, M., Dunker, A. K., and Obradovic, Z. (1999) Predicting protein disorder for N-, C- and internal regions. *Genome Inf. 10*, 30–40.

29. Romero, P., Obadovic, Z., Li, X., Garner, E., Brown, C., and Dunker, A. K. (2001) Sequence complexity of disordered proteins. *Proteins: Struct., Funct., Genet. 42*, 38–48.
30. Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J. Mol. Biol. 347*, 827–839.
31. Dosztányi, Z., Csizmók, V., Tompa, P., and Simon, I. (2005) IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics 21*, 3433–3444.
32. Heckman, K. L., and Pease, L. R. (2007) Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc. 2*, 924–932.
33. Kelly, S. M., and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci. 1*, 349–384.
34. Calloni, G., Lendel, C., Campioni, S., Giannini, S., Gliozzi, A., Reini, A., Vendruscolo, M., Dobson, C. M., Salvatella, X., and Chitti, F. (2008) Structure and dynamics of a partially folded protein are decoupled from its mechanism of aggregation. *J. Am. Chem. Soc. 130*, 13040–13050.
35. Surewicz, W. K., and Mantsch, H. H. (1988) New insight into protein secondary structure from resolution-enhanced infrared spectra. *Biochim. Biophys. Acta 952*, 115–130.
36. Byler, D. M., and Susi, H. (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers 25*, 469–487.
37. Arrondo, J. L. R., and Goñi, F. M. (1999) Structure and dynamics of membrane proteins as studied by infrared spectroscopy. *Prog. Biophys. Mol. Biol. 72*, 367–405.
38. Denning, D. P., Uversky, V., Patel, S. S., Fink, A. L., and Rexach, M. (2002) The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. *J. Biol. Chem. 277*, 33447–33455.
39. Braiman, M. S., and Rothschild, K. J. (1988) Fourier transform infrared techniques for probing membrane protein structure. *Annu. Rev. Biophys. Biophys. Chem. 17*, 541–570.
40. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley & Sons, New York.
41. Tziatzios, C., Schubert, D., Lotz, M., Gundogan, B., Betz, H., Schagger, H., Haase, W., Duong, F., and Collison, I. (2004) The bacterial protein-translocation complex: SecYEG dimers associate with one or two SecA molecules. *J. Mol. Biol. 340*, 513–524.
42. Daragan, V. A., and Mayo, K. H. (1997) Motional model analysis of protein and peptide dynamics using ^{13}C and ^{15}N NMR relaxation. *Prog. NMR Spectrosc. 31*, 63–105.
43. Alcaraz, L. A., del Alamo, M., Mateu, M. G., and Neira, J. L. (2008) Structural mobility of the monomeric C-terminal domain of the HIV-1 capsid protein. *FEBS J. 275*, 3299–3311.
44. Evans, J. N. S. (1996) *Biomolecular NMR*, Oxford University Press, Oxford, U.K.
45. Dyson, H. J., and Wright, P. E. (2005) Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol. 6*, 197–208.
46. Wang, R., Ilangovan, U., Robsinson, A. K., Schirg, V., Schwarz, P. M., Lafer, E. M., Demeler, B., Hinck, A. P., and Kim, C. A. (2008) Structural transitions of the Ring 1B C-terminal region upon binding the Polycomb cbox domain. *Biochemistry 47*, 8007–8015.
47. Mayor, U., Guydosh, N. R., Johnson, C. M., Grossmann, G., Sato, S., Jas, G. S., Freund, S. M. V., Alonso, D. O. V., Daggett, V., and Fersht, A. R. (2003) The complete folding pathway of a protein from nanosecond to microsecond. *Nature 42*, 863–867.
48. Raychaudhuri, S., Majumder, P., Sarkar, S., Giri, K., Mukhopadhyay, D., and Bhattachaaryya, N. P. (2008) Huntingtin interacting protein HYPK is intrinsically unstructured. *Proteins: Struct., Funct., Genet. 71*, 1686–1698.
49. Uversky, V. N., Yamin, G., Munishkin, L. A., Karymovc, M. A., Millettd, I. S., Doniachd, S., Lyubchenkoc, Y. L., and Fink, A. L. (2005) Effects of nitration on the structure and aggregation of α-synuclein. *Mol. Brain Res. 134*, 84–102.
50. Batra-Safferiling, R., Abarca-Heidermann, K., Körschen, H. G., Tziatzios, C., Stoldt, M., Budyak, I., Willbold, D., Schwalbe, H., Klien-Seetharaman, J., and Benjamin Kaupp, U. (2006) Glutamic acid rich proteins of rod photoreceptors are natively unfolded. *J. Biol. Chem. 281*, 1449–1460.
51. Adams, V. H., McBryant, S. J., Wade, P. A., Woodcock, C. L., and Hanse, J. C. (2007) Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *J. Biol. Chem. 282*, 15057–15064.
52. Liew, C. K., Gamsjaeger, R., Mansfield, R. E., and Mackay, J. P. (2008) NMR spectroscopy as a tool for the rapid assessment of the conformation of GST-fusion proteins. *Protein Sci. 17*, 1630–1635.
53. Millard, T. H., Sharp, S. J., and Machensky, L. M. (2004) Signalling to actin assembly via the WASP (Wiskott-Aldrich syndrome protein) family proteins and the Arp2/3 complex. *Biochem. J. 380*, 1–17.
54. Haynes, C., Oldfield, C. J., Ji, F., Klitgord, N., Cusik, M. E., Radivojac, P., Uversky, V. N., Vidal, M., and Iakoucheva, L. M. (2006) Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput. Biol. 2*, e100.
55. Tompa, P., Szasz, C., and Buday, L. (2005) Structural disorder throws new light on moonlighting. *Trends Biochem. Sci. 30*, 484–489.

BI801933C