Biochemistry

 Volume 26, Number 4 February 24, 1987

Articles

Structural Characterization of the Trypsin-Resistant Core in the Nuclear Sperm-Specific Protein from *Spisula solidissima*†

Juan Ausio,* Arazdordi Toumadje, Reginald McParland, Robert R. Becker, W. Curtis Johnson, Jr., and Kensal E. van Holde

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331

Received August 27, 1986; Revised Manuscript Received October 23, 1986

ABSTRACT: Trypsin digestion of the protamine-like protein from *Spisula solidissima* has revealed the existence of an internal resistant core. The peptide contains 75 amino acid residues, and its primary structure shows some conserved sequences that are common to those found in the core of the somatic histone H5 from chicken erythrocytes. The secondary structure of this core exhibits 33% antiparallel β-sheet, 18% β-turns, 37% random coil, and only 10% α-helix, in contrast to histone H5. Hydrodynamic measurements indicate a compact globular assembly for the tertiary structure of this peptide, when compared to the more extended shape observed for the whole protein. The possible relatedness of this protein to the histone H1 family is discussed.

Quite recently, Ausio and Subirana (1982a) described an interesting protein which is found associated with DNA in the nuclei of the sperm of the surf clam *Spisula solidissima*. Since this substance shares some common features with both protamines and histones, it has been termed a "protamine-like protein" (PLP). The first amino acid analysis carried out on this protein (Subirana et al., 1973) showed a very simple composition, with four major amino acids (lysine, arginine, serine, and alanine) accounting for 86% of the whole amino acid content. Such a simple compositional pattern seems to be a common feature shared by most of the protamine-like proteins from the sperm of the bivalve molluscs (Subirana et al., 1973; Ausio, 1986). The molecular weights of such proteins exhibit a strong variability from one species to another (Ausio & Subirana, 1982b; Ausio, 1986), but this does not preclude the possibility of an evolutionary relatedness within the group (Subirana et al., 1973). The high molecular weight of the protamine-like protein from *Spisula*, together with its high lysine (25%) and alanine (14%) content, could suggest some evolutionary relatedness of this protein to the histone H1 family (Ausio & Subirana, 1982a). A further suggestion of some such relationship arises from the observation of a trypsin-resistant globular core as will be detailed in this paper. Nevertheless, as has been recently shown (Ausio & van Holde, 1987), this protein accounts for 70–80% of the nuclear protein which is found complexing the DNA in the sperm. The protamine-like protein coexists with a complete set of histones, including a sperm-specific H1 histone. However, these histones constitute only about 20% of the sperm chromatin proteins; the protamine-like component is 4 times more abundant. Thus, the protamine-like protein must play a quite different role than do most members of the H1 family. These questions are discussed in this paper.

Materials and Methods

Living Samples. The surf clam (*Spisula solidissima*) specimens were purchased from the Department of Animal Resources at the Marine Biological Laboratory in Woods Hole, MA.

Isolation and Purification of the Protamine-like Component. The major protein component from the sperm of *S. solidissima* was prepared as described elsewhere (Ausio & van Holde, 1987).

Trypsin Digestion of the Protamine-like Component. The protamine-like protein was digested with trypsin essentially as described by Hartman et al. (1977). The protamine-like protein at 4–6 mg/mL in double-distilled water was brought to 2 M NaCl by addition of 1 volume of 4 M NaCl and 20

† This research was supported by U.S. Public Health Service Research Grants GM-22916 (to K.E.v.H.) and GM-21479 (to W.C.J.) from the Institute of General Medical Sciences and by the Central Service Facility of the Center for Gene Research and Biotechnology of Oregon State University. K.E.v.H. was the recipient of a Research Professorship from the American Cancer Society.

0006-2960/87/0426-0975$01.50/0

mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), pH 7.5, buffer and was digested with trypsin at E:S = 1:1000 (w/w) for different periods of time at 22 °C. For the determination of the time dependence of the digestion, 20-μL aliquots were withdrawn at the specified times and added to 40 μL containing a 10-fold excess by weight of *N*-tosyl-L-lysyl chloromethyl ketone (TLCK) with respect to the enzyme. The samples were finally suspended in 60 μL of 10 M urea, 10% acetic acid, 3% clupeine sulfate (protamine), 5% β-mercaptoethanol, and 2% pyronine and were directly loaded on 15% polyacrylamide–urea–acetic acid gels.

In order to obtain the trypsin-resistant core (limit of digestion peptide) in quantities amenable for further purification, ~50–100 mg of original protamine-like component was digested under the same conditions as described above. The enzymatic digestion was carried out for 45 min at 22 °C, and the reaction was stopped by a 10-fold excess of TLCK and by precipitation of the protein with 18% trichloroacetic acid (TCA) at 4 °C for 5 min. The precipitate so obtained was centrifuged at 12000*g* for 10 min at 4 °C and washed twice under the same conditions with 1 N HCl–acetone at 4 °C. The pellet obtained after the last acetone wash was dried under low vacuum. The powder prepared in this way was finally resuspended in a small volume of 8 M urea, 50 mM NaCl, 20 mM HCl, and 0.1% dithiothreitol (DTT) and subjected to gel filtration chromatography to achieve its final purification.

Column Chromatography. Gel filtration chromatography was performed on Bio-Gel P-60 in a 1.5 × 160 column with 1.2-mL fractions (6 per hour) under the conditions described by Bohm et al. (1973).

Gel Electrophoresis. Polyacrylamide gel electrophoresis either in sodium dodecyl sulfate (SDS) or in 2.5 M urea–5% acetic acid gels was performed as described elsewhere (Ausio et al., 1986).

Fluorescence Analysis. Fluorescence measurements were made on a computer-interfaced fluorescence spectrometer (Ayres et al., 1974) with the modifications described by Libertini and Small (1985).

Peptide Sequencing. Peptide sequencing was performed on an Applied Biosystems Model 470A gas-phase protein sequencer. The standard ABI 02CSer program was used for coupling and cleavage with the cartridge set at 40 °C. Anilinothiazolinone (ATZ) amino acid derivatives were converted to phenylthiohydantoin (PTH) derivatives with trifluoroacetic acid (TFA) at a conversion temperature of 55 °C. The repetitive yield over the entire 75-residue sequence averages 95%.

PTH-amino acids were analyzed on a Beckman μ-flow high-performance liquid chromatography (HPLC) system equipped with an IBM cyano column. *N*-Bromosuccinimide (NBS) cleavage of protamine-like protein (PLP) was carried out as described by Ramachandran and Witkop (1967). Under these conditions, the protein is mainly cleaved at the tryptophanyl bond. The peptides obtained after the cleavage were fractionated in a Bio-Gel P-60 column under the conditions described above.

Analysis of Protein Concentrations. Protein concentrations were determined by ninhydrin assay (Hennessey & Johnson, 1981) and by the guanidine hydrochloride method of Mulvey et al. (1974) and Elwell and Schellman (1977). Lyophilized *Spisula* protamine and protamine core were dissolved in double-distilled water and brought to 0.01 M sodium phosphate buffer, pH 7.0. To determine the protein concentration by the ninhydrin method, triplicate 60-μL samples were hydrolyzed by adding 2 mL of constant-boiling HCl to a 5-mL nitrogen-flushed, evacuated, drying ampule and heating to 110 °C for 22 h. Sample ampules were then rotoevaporated to dryness, and the residues were dissolved in 1 mL of 0.2 N sodium citrate buffer containing 0.5% thiodiglycol solution, 0.01% octanoic acid, and 0.06% BRIJ solution, pH 2.20. From each ampule, 0.5 mL of solution was extracted and rediluted with an additional 2 mL of sodium citrate buffer. To the resulting solution, 1.25 mL of a ninhydrin solution was added and heated to 100 °C in a water bath for 20 min and then cooled to room temperature. The optical densities of processed samples and standards were measured at 570 nm on a Cary 15 spectrophotometer. From these readings, the concentration can be computed. The extinction coefficient determined at 190 nm for *Spisula* protamine is 1.13×10^4 M^{-1} cm^{-1}.

The second method used to determine the concentration relies on guanidine hydrochloride. The protein is unfolded by 6 M guanidine hydrochloride (Schwarz/Mann). The concentration is computed from the absorption of Trp, Tyr, and Cys at 280 nm since the number of these residues and their extinction coefficients at 280 are known. The extinction coefficient at 190 nm determined for *Spisula* protamine by this method is 1.18×10^4 M^{-1} cm^{-1}, differing only within the standard deviation from the ninhydrin method.

The extinction coefficient at 190 nm for protamine core was determined only by the guanidine hydrochloride method and found to be 7.07×10^3 M^{-1} cm^{-1}.

Spectra. CD spectra were measured over the 178–260-nm range by using a vacuum ultraviolet (VUV) CD spectrophotometer (Johnson, 1971), with 50-μm path-length cells, and protein solutions 0.65–0.75 mg/mL in 0.01 M sodium phosphate, pH 7.0. The instrument was calibrated with (+)-10-camphorsulfonic acid, assuming $\Delta\epsilon(290.5) = 2.36$. All spectra were measured at 20 °C with a 20-s time constant and a scan rate of 1 nm/min. Transmission spectra were measured before and after CD for the same sample to assure that no air bubbles formed in the solution during the CD scan. In all measurements, the total OD did not exceed 1.0 within the range of the CD spectrum.

Prediction of Secondary Structure from Analysis of CD Data. The CD data were analyzed for secondary structure by using "variable selection" (Mosteller et al., 1977) to modify the original method of Hennessey and Johnson (1981). This method has recently been applied to predict the secondary structure of proteins by Manavalan et al. (1986). The idea is to remove one or more reference proteins that contain extraneous contributions to the CD from the basis set and recalculate the secondary structure. Since we do not know which proteins should be removed from the basis set, calculations were performed for all possible subsets. If there are N proteins in the basis set and R proteins to be removed, there are $N!/R!(N - R)!$ combinations.

The following criteria have been applied to select the best solution(s) among the subsets: (1) The sum of the predicted secondary structure features should be in the range of 0.90–1.10. (2) No structure should be large and negative. (3) The sum of squares of residuals (SSR) should be less than 2.0.

Secondary Structure Prediction from Amino Acid Sequence. The secondary structure was predicted by using each of the following methods: Chou & Fasman (1978), Burgess et al. (1974), Lim (1974), and Garnier et al. (1978). A residue was assigned α-helix or β-sheet if it was predicted by three or more of these methods.

Diffusion Coefficients and Shape Parameters. The diffusion coefficients both for the whole protamine-like component and for its trypsin-resistant core were determined by free diffusion in a double-sector synthetic boundary cell in a Model E

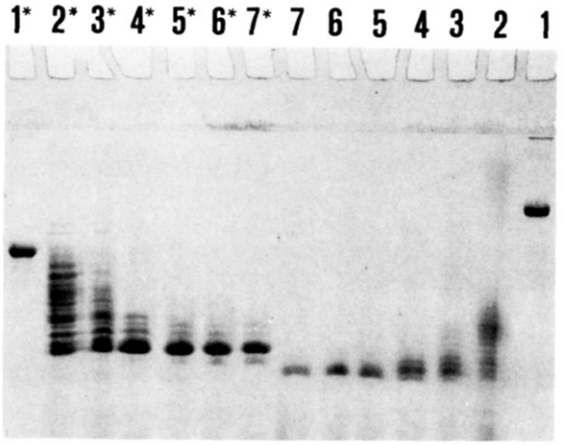

FIGURE 1: Time course of trypsin digestion of the protamine-like protein from *S. solidissima* (lanes 1–7) in comparison to histone H1 from half-thymus (lanes 1*–7*) carried out under the same conditions: 2 M NaCl–10 mM Tris-HCl, pH 7.5, at trypsin:protein ratio = 1:1000 and at 22 °C. The gel electrophoresis shown was in a urea–acetic acid gel as described under Materials and Methods. The time for each lane is as follows: (1, 1*) 0 min; (2, 2*) 4 min; (3, 3*) 8 min; (4, 4*) 16 min; (5, 5*) 32 min; (6, 6*) 64 min; (7, 7*) 128 min.

(Beckman) ultracentrifuge. Interference optics were employed. The experiments were performed at 10000 rpm in 0.15 M NaCl–20 mM Tris-HCl at pH 8.0 and 20 °C. Analysis of the photographic plates was carried out as described by Chervenka (1969).

RESULTS

Existence of a Trypsin-Resistant Core in the Protamine-like Component of Spisula Sperm. The sperm of the surf clam *Spisula solidissima* contains a major nuclear protein with characteristics intermediate between those of protamines and histones (Ausio & Subirana, 1982a,b). This peculiar protein (PLP) coexists in the sperm nuclei with a set of five histones in an average weight ratio of 4 parts PLP to 1 part histones (Ausio & van Holde, 1987). Proteins with a very similar composition to that found in *Spisula* have been described in other species of bivalve molluscs (Ausio, 1986), and they seem to be a common feature of this taxonomic group. Recently, one such protein in the sperm of the razor clam *Ensis minor* has been shown to have a trypsin-resistant core (Giancotti et al., 1983) similar to that found in histone H1. Indeed this protein, which in turn is very similar in terms of its amino acid composition to that present in *Spisula* sperm, has been ascribed to belong to the histone H1 family (Giancotti et al., 1983). On this basis, we decided to investigate whether or not such a protein core was also present in the protamine-like component of *S. solidissima*.

Figure 1 shows the time course of the digestion of the protamine-like component in comparison to that exhibited by calf thymus histone H1 digested under the same conditions. It is clear from this result that indeed the major protein component in the sperm of the surf clam *S. solidissima* has a resistant core. In urea–acetic acid gels, this core runs slightly faster than that of histone H1. The use of urea–acetic acid gels (Panyim & Chalkley, 1969) in this experiment is dictated by the fact that the *Spisula* PLP is completely insoluble in SDS, probably because of its high positive charge. Yet the limit peptide obtained by trypsin digestion is soluble in SDS. When electrophoresed on SDS gels, this PLP core runs slightly faster than the H1 core (Figure 2). Since the latter has a molecular weight of approximately 8500, we *estimate* a value of 8300 for the molecular weight of the PLP core. This value is uncertain, however, since it is difficult to find standards

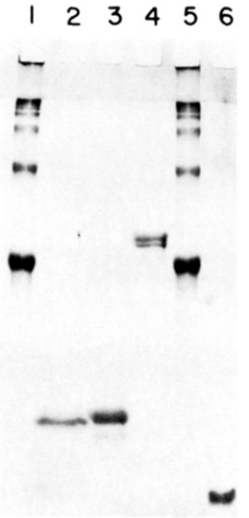

FIGURE 2: SDS gel electrophoresis (Laemmli, 1970) of the protamine-like protein core and some comparable proteins. (Lanes 1 and 5) Histones H5 and its oligomers (Russo et al., 1983); (lane 2) trypsin-resistance core of *Spisula* protamine-like protein; (lane 3) trypsin-resistant core of calf H1; (lane 4) whole calf H1; (lane 5) insulin.

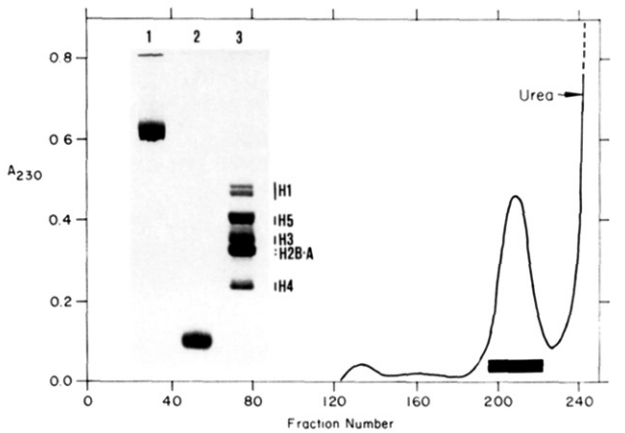

FIGURE 3: Purification of the trypsin-resistant core of the protamine-like protein from *Spisula* in Bio-Gel P-60 in a column of 1.5 × 160 cm eluted with 50 mM NaCl, 20 mM HCl, and 0.1 mM DTT (Bohm et al., 1973). The inset corresponds to an urea–acetic acid gel to show the purity of the fraction achieved by this purification method. (1) Whole protamine-like protein; (2) purified trypsin core; and (3) chicken erythrocyte histones used as a marker.

appropriate for such molecules.

Once the existence of a tryptic-resistant core for the protamine-like protein of *Spisula solidissima* was established, we decided to isolate and purify this peptide so as to be able to further analyze its structural features. Figure 3 shows the method of purification employed and the high purity of the peptide obtained.

Since the PLP of *S. solidissima* contains only one residue of tryptophan (Ausio & Subirana, 1982a), we first determined whether this amino acid was included in the core region. Figure 4 shows the fluorescence emission both for the whole protamine-like component and for its tryptic peptide. Indeed, from this figure, it is clear that tryptophan not only is present but also is in the same stoichiometry which should be expected on the basis of the molecular mass for the whole protamine (Ausio & Subirana, 1982b) and the value determined above for the core.

Primary Structure of the Protamine-like Core. The high degree of purity achieved on the purification of the peptide core (Figure 3) allowed us to obtain its complete sequence by automatic sequence analysis without any need of cleaving the

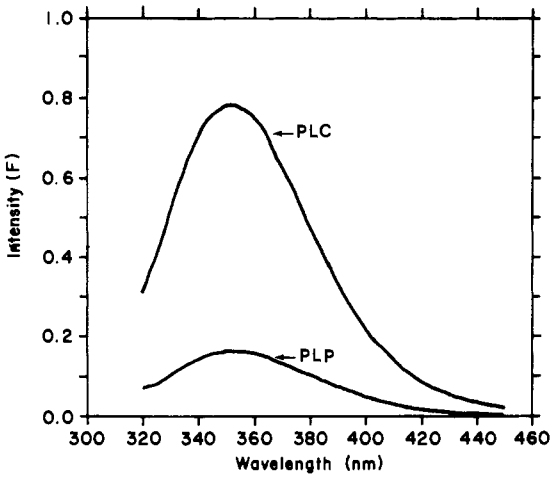

FIGURE 4: Fluorescence emission spectra with excitation at 280 nm for the whole protamine-like protein (PLP) and its core (PLC). The spectra are uncorrected for the wavelength dependence of the emission sensitivity.

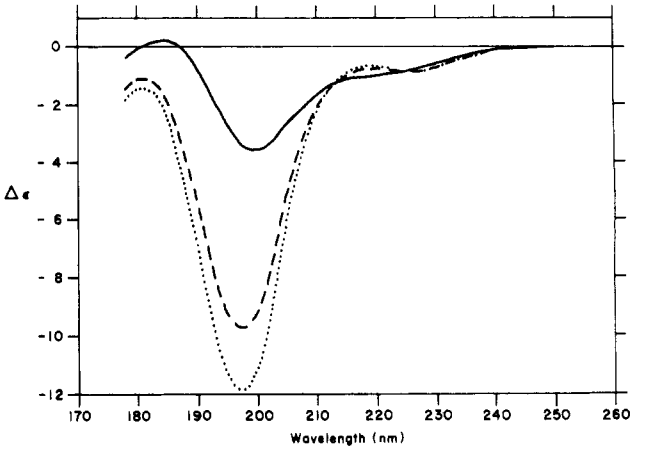

FIGURE 5: Circular dichroism spectra for the protamine-like protein (PLP) (---) and its trypsin-resistant core (PLC) (—) and for the contribution of the N- and C-terminal region tails (whole protein minus core) (PLT) (···). All the spectra were obtained in 10 mM phosphate buffer at pH 7.0.

peptide into smaller pieces. We noted ambiguity only in the last three amino acids. These uncertainties were easily resolved by sequencing one overlapping peptide in this region obtained by *N*-bromosuccinimide cleavage of the whole protamine in the presence of 8 M urea and 5% acetic acid. Under these conditions, NBS cleaves mainly at tryptophan residues. The sequence found for the PLP core was

1 10
Lys-Gly-Ser-Ser-Gly-Met-Met-Ser-Met-Val-Ala-Ala-Ala-Ile-Ala-
20 30
Ala-Asn-Arg-Thr-Lys-Lys-Gly-Ala-Ser-Ala-Gln-Ala-Ile-Arg-Lys-
40
Tyr-Val-Ala-Ala-His-Ser-Ser-Leu-Lys-Gly-Ala-Val-Leu-Asn-Phe-
50 60
Arg-Leu-Arg-Arg-Ala-Leu-Ala-Ala-Gly-Leu-Lys-Ser-Gly-Ala-Leu-
70
Ala-His-Pro-Lys-Gly-Ser-Ala-Gly-Trp-Val-Leu-Val-Pro-Lys-Lys
←NBS fragment→

The core, therefore, contains 75 amino acids, with a total mass of 7615 daltons, values slightly smaller than expected from the SDS gel electrophoretic analysis. From further partial sequencing studies, we find that this core region is preceded by an N-terminal tail of about 150 residues and followed by a C-terminal tail of about 75 residues (Ausio et al., unpublished results).

The core sequence has an amino acid composition which is completely different from the average amino acid composition for the whole PLP (Ausio & Subirana, 1982a) and contains all the aromatic residues (Phe, Tyr, Trp). Such a difference in composition is clearly reflected in the differences found in the higher levels of structural organization as will be shown in the following sections.

Secondary Structure. Figure 5 shows the CD spectra on a per amide basis for the whole PLP (~300 amino acids) (Ausio & Subirana, 1982a), the PLP core (75 amino acids), and the N- and C-terminal tails (~225 amino acids). The CD for the tails is calculated by subtracting the total $\Delta\epsilon$ of the core from the total $\Delta\epsilon$ of PLP according to

$$\Delta\epsilon_{PLP}n_{PLP} - \Delta\epsilon_{PLC}n_{PLC} = \Delta\epsilon_{PLT}n_{PLT} \quad (1)$$

where $\Delta\epsilon$ = CD per amide, n = the number of amino acid residues, PLP = protamine-like protein, PLC = PLP trypsin-resistant core, and PLT = PLP − PLC = PLP "tails". Although all three CD spectra have the shape traditionally associated with random-coiled polypeptides, their magnitudes vary by more than a factor of 3. Clearly, there are real differences in secondary structure here, and our analysis reveals these.

Table I: Percent Secondary Structure from CD Analysis of *Spisula* Protamine, Core, and Tail with Variation of Runs from Variable Selection

structure[a]	*Spisula* protamine	core	tail
H	0.05 ± 0.02	0.10 ± 0.01	0.05 ± 0.01
A	0.15 ± 0.03	0.33 ± 0.01	0.11 ± 0.02
P	−0.03[b]	0	−0.05[b]
T	0.46 ± 0.03	0.18 ± 0.01	0.52 ± 0.03
O	0.39 ± 0.02	0.37 ± 0.01	0.42 ± 0.01
total	1.04 ± 0.03	0.99 ± 0.01	1.10 ± 0.02

[a] H, α-helix; A, antiparallel β-sheet; P, parallel β-sheet; T, β-turn; O, other structure. [b] Negative numbers are assumed to be zero.

The PLP spectrum has no positive peak and is characterized by one negative peak at 198 nm. The secondary structure analysis for the CD spectrum of the whole protein yields 5% α-helix (H), 15% antiparallel β-sheet (A), −3% parallel β-sheet (P), 46% β-turns (T), and 39% random coil or "other" (O). Negative numbers are assumed to be zero; therefore, this protein does not contain any parallel β-sheet (Table I). This table is produced by using the variable selection method and averaging all of the combinations that meet the criteria described under Materials and Methods. The number of proteins in the original basis set was 22. After removal of 6 or 8 proteins by using variable selection, the remaining proteins in the basis set (16 proteins for the whole PLP and the PLC and 14 proteins for PLT) give a good analysis. However, the basis sets used to analyze the whole PLP and the PLC do not contain all of the same proteins. Table I shows the percent secondary structure from an average of 14 runs with 16 proteins that satisfy our criteria for *Spisula* PLP, 16 such runs for the PLP core, and 6 runs with 14 proteins for the tail regions. The PLP, its core, and its tails all contain a small percent of α-helix, even though the core has twice as large a fraction as the tail or whole protein. The fraction of antiparallel β-sheet in the core (0.33) is twice as much as in whole protamine (0.15) and 3-fold as much as in the tails (0.11).

Table II is produced to show the percent contribution from core or tails to the total of each secondary structure for the whole PLP. The percent of each secondary structure for whole protein, core, and tail is multiplied by their number of amino acids to calculate the number of amino acids in the α-helix, β-sheet, β-turn, and random coil. The difference between the

Table II: Amino Acid Content of *Spisula* Protamine, Core, and Tail in Each Secondary Structure and Percent Contribution of Core and Tail in Total α-Helix, β-Sheet, β-Turn, and Other Structure

structure[a]	*Spisula* protamine	core	tails	% contribution from core	% contribution from tails
H	15	8	11	42	58
A	45	24	24	51	49
T	137	14	115	11	89
O	116	28	93	23	77

[a] Abbreviations as in Table I.

Table III: Percent Secondary Structure Analysis of Protamine Core from Amino Acid Sequence

method	α-helix	β-sheet (A + P)[a]	remaining (T + O)[b]
Chou & Fasman (per amide) (1978)	0.36	0.12	0.52
Burgess et al. (1974)	0.19	0.23	0.58
Garnier et al. (1978)	0.57	0.12	0.31
Lim (1974)	0.35	0.16	0.49
joint method	0.27	0.11	0.62

[a] A, antiparallel; P, parallel. [b] T, β-turns; O, other structures.

number of amino acids in each secondary structure for the whole PLP and the sum of the core and the tails is within the variation of runs from variable selection. The antiparallel β-sheet in whole PLP is contributed equally by the core and the tails; however, 89% of the β-turn and 77% of the random coil are contributed by the tail regions.

Prediction of Secondary Structure of Spisula PLP Core from Amino Acid Sequences. The amount of α-helix, β-sheet (parallel and antiparallel), and remaining structure (β-turns and random coil) predicted from the primary structure of the core is shown in Table III. Although all the methods predict most of the α-helices and β-sheets in similar regions along the sequences, there is a significant difference in their length and hence in their total in each category. Robson's method (Garnier et al., 1978) predicted 57% α-helix and Burgess' method (Burgess et al., 1974) 19%. Because of these variations, a joint method is used which assigns a specific region as α-helix or β-sheet where at least three out of four methods have a common prediction. From this method, the regions around the segments 7–17, 23–26, and 46–53 are predicted as α-helices, which corresponds to 27% for the core (Table III). Similarly, the predicted regions for β-sheet are 41–45 and 68–72, which corresponds to 11% for the core. The remaining segments, which correspond to 62% for the core, are neither α-helix nor β-sheet. However, the segments around 17–22, 35–38, 55–57, and 62–67 are predicted as β-turn by the Chou and Fasman method (Chou & Fasman, 1978), and this corresponds to 20% for the core. If one subtracts 20% β-turn from the 62% remaining segments, the percent random coil in the core would be 42%.

Comparison of the predictions from amino acid sequences and CD analysis of the core shows a slight difference in the amount of random coil (42% vs. 37%) and β-turn (18% vs. 20%, respectively). However, the CD analysis indicated 10% α-helix and 33% β-sheet as compared to 27% α-helix and 11% β-sheet from the amino acid sequence method.

Spisula PLP is complexed with DNA in the cell (Ausio & van Holde, 1986). It is reported that some proteins have a different secondary structure and conformation when they are complexed with DNA (Li et al., 1971) or to a cofactor (Manavalan et al., 1986) or at a different pH (Labhardt, 1982). The structure of the protein when free in solution may very well differ from that in the DNA complex. Since the preparation of the PLP core involves a number of denaturing steps, there might be concern that the core refolds, before CD studies, into a globular structure different from that existing in the native protein. However, this would be likely only if the tails interact strongly with the core. The fact that the tails seem to be highly extended (see below) argues against this possibility.

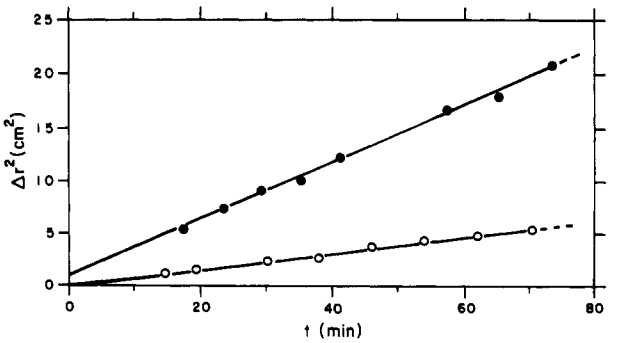

FIGURE 6: Plot of Δr^2 vs. time used to calculate the diffusion coefficient (Chervenka, 1969) for the whole protamine-like component (○) and for its trypsin-resistant core (●). Δr is given in centimeters, and it is a measure of the broadening of the boundary formed by the synthetic boundary cell in the ultracentrifuge, which is measured as a function of time in seconds.

Tertiary Structure. In order to analyze and compare the shape of the trypsin-resistant core with the whole protamine-like protein, we have measured the diffusion coefficient of both in the analytical ultracentrifuge. Figure 6 shows the kind of graph employed to evaluate these constants (Chervenka, 1969); the slopes of the lines shown in that figure are proportional to the diffusion coefficients. By using this methodology, we have estimated $D_{20,w} = 14.4 \times 10^{-7}$ cm^2·s^{-1} for the core and $D_{20,w} = 4.0 \times 10^{-7}$ cm^2·s^{-1} for the whole protein. Note that these values are corrected to standard conditions (20 °C and water); they are not, however, extrapolated to zero concentration of the macromolecular solute. The concentration was, however, the same in both cases (≃6.5–7.5 interference fringes, i.e., ≃1.5–1.8 mg/mL). The value for $D_{20,w}$ found for the whole molecule is consistent with the value of $D^0_{20,w} = 6.0 \times 10^{-7}$ cm^2·s^{-1} found for histone H5 with a molecular weight of 20 900 (Champagne et al., 1970) when taking into account the molecular weight difference. These diffusion coefficients have allowed us to estimate some structural parameters.

The diffusion coefficient can be used to calculate an observed frictional ratio, $(f/f_0)_{obsd}$:

$$(f/f_0)_{obsd} = \frac{RT}{6\pi\eta_{20,w}\mathcal{N}D_{20,w}(3\bar{M}v/4\pi\mathcal{N})^{1/3}} \quad (2)$$

where $\mathcal{N}$ = Avogadro's number, $\eta_{20,w}$ = the viscosity of water at 20 °C, M = the molecular mass, and $\bar{v}$ = the partial specific volume. Using, for the core peptide, $M = 7615$ g mol^{-1} and $\bar{\delta} = 0.74$ cm^3 g^{-1} (estimated from the amino acid composition), we find $(f/f_0)_{obsd} = 1.13$. This in itself indicates that the core is a compact molecule, but even this estimate of f/f_0 is probably high because of hydration. If a molecule is hydrated to the extent of δ g of H_2O/g of protein, the shape-dependent part of f/f_0 is given by (Tanford, 1961)

$$f/f_0 = (f/f_0)_{obsd}(1 + \delta/\bar{v}\rho)^{1/3} \quad (3)$$

Using a value of $\delta = 0.37$ g/g, estimated by the method of Kunz and Kauzman (1971), we find $f/f_0 = 1.00$. Clearly, the core has a compact structure. If the core is in fact a sphere, its radius of gyration is then calculated to be 13.1 Å. This

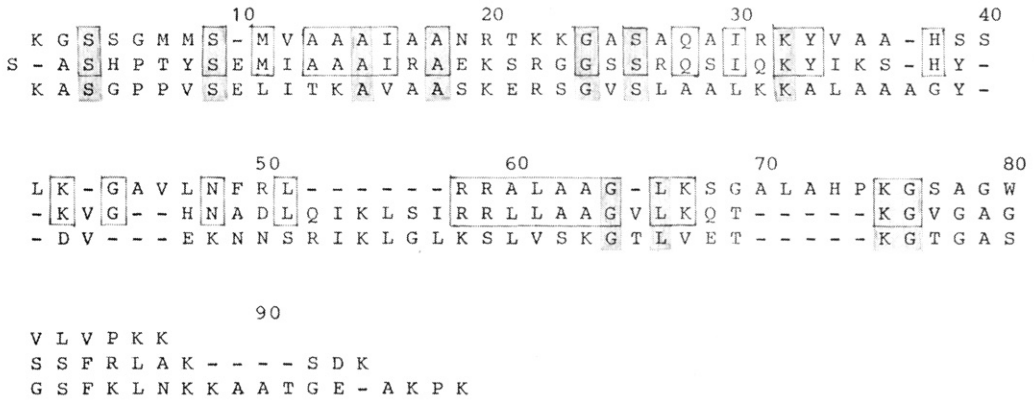

FIGURE 7: Sequences of the trypsin-resistant core peptides from (top) protamine-like protein from *Spisula*, (middle) histone H5 from chicken erythrocyte (Briand et al., 1980), and (bottom) histone H1 from calf thymus (Liao & Cole, 1981). The sequences have been aligned for maximum coincidence. The boxes represent homologous sequences between *Spisula* protein and histone H5, and the shaded regions correspond to sequences which are common to the three peptides.

is consistent with the value of 14.5 Å reported for the compact core of H5 (79 amino acid residues) from neutron scattering (Allan et al., 1980).

These results contrast markedly with those calculated for the whole protein. Using the value of $D_{20,w}$ reported above, the molecular mass of 33 500 g mol^{-1} (Ausio & Subirana, 1982a), and estimated values of $\bar{v} = 0.709$ cm^3 g^{-1} and $\delta = 0.41$ g/g, we find $f/f_0 = 1.65$. From this, we conclude that the whole PLP molecule is very asymmetric.

Indeed, on the basis of a prolate ellipsoid model for the shape of the molecule, one may calculate the axial ratio a/b from

$$f/f_0 = \frac{(1 - b^2/a^2)^{1/2}}{(b/a)^{1/2} \ln \{[1 + (1 - b^2/a^2)^{2/3}]/(b/a)\}} \quad (4)$$

The value thus obtained, $a/b = 11.7$, is very similar to the ratio $a/b = 9$ found for the whole histone H5 (Aviles et al., 1978).

Comparison of the above results strongly suggests that the N- and C-terminal tails of the molecule are highly extended in solution. It seems unlikely, from the CD data, that these tails are entirely "random coil", since identifiable secondary structure remains. The existence of secondary structure in the tail region is unusual in the H1 protein class, and to our knowledge, a similar situation has only been reported so far for the sperm-specific histone H1 of the sea urchin *Sphaerechinus granularis* (Giancotti et al., 1981).

Discussion

In the present paper, we have analyzed the different levels of structure of the trypsin-resistant core from the protein which accounts for almost 75% of the basic proteins associated with the DNA in the sperm of the surf clam *Spisula solidissima* (Ausio & van Holde, 1986).

From this stoichiometry and from "in vivo" studies about the association of this protein with DNA (Ausio & van Holde, 1987), it is clear that the protein behaves like a protamine in spite of its amino acid composition and molecular weight (Ausio & Subirana, 1982a) which are clearly different from those found in the "typical" fish protamines (Ausio et al., 1973). On the other hand, the presence of a trypsin-resistant core and the amino acid composition of the whole protein (Ausio & Subirana, 1982a) might suggest some relationship to the histone H1 family. We are therefore clearly confronting a situation where it is really difficult to discern whether we are dealing just with protein structural analogy or with a real homology. The following detailed discussion at each of the different structural levels should help in clarifying such ambiguity.

Figure 7 shows the sequence (primary structure) of the core when compared to the known sequences for histone H1 from calf thymus (Liao & Cole, 1981) and from histone H5 from chicken erythrocyte (Briand et al., 1980). We have chosen the first one because of its general representation of the whole H1 protein family and the H5 histone because it appears in a biological situation which somewhat resembles that found in the sperm (e.g., a terminally differentiated system). The boxes in Figure 7 indicate the identities existing in the sequence of this region, which has been termed the most well conserved within the histone H1 family. The quantification of such boxes indicates a "*Spisula* core"–"H5 core" sequence coincidence of 39%, "H1 core"–"H5 core" 37%, and "*Spisula* core"–"H1 core" 21%. The whole sequence coincidence for the three proteins when taken together is only 12.5%. It is thus clear that at this structural level the core from the protamine-like protein from *S. solidissima* resembles more that of H5 than that of histone H1. At the next structural level, the CD data (see Table I) for the protamine reveal the existence of 15% antiparallel β-sheet and very low α-helix content. These values are almost multiplied by a factor of 2 in the core region, except for the β-turn which drastically decreases in this latter case. The existence of $\simeq$35% of antiparallel β-sheet in the core is especially important in a context of the comparison between the whole PLP and the proteins conventionally considered to be of the histone H1 family. Neither the canonical H1 histone nor the more specialized histone H5 shows any indication for the existence of any β-structure in their trypsin-resistant cores (Bradbury et al., 1975; Crane-Robinson et al., 1976). Instead, the α-helix content for the core regions of such proteins is $\simeq$30% (Giancotti et al., 1981; Aviles et al., 1978) compared to the 10% observed for the core of the protamine-like protein from *S. solidissima*.

Since we know the sequence for the core region, we have attempted but failed in performing a structural assignment to different parts in that region based on predictive methods. Table II summarizes such an attempt. As a matter of fact, none of these methods fits the structure experimentally obtained by circular dichroism. As has been pointed out before, the lack of coincidence between the predictive methods and the experimentally observed values may arise for several reasons, although most likely this is simply due to the very peculiar nature of the protein analyzed here. From all the methods used (see Table III), only that of Burgess et al. (1974) comes at all close to the experimental results. Nevertheless, we have used the graphic method of Chou and Fasman (1978) (Fasman et al., 1977) (see Figure 8) to compare graphically the structural predictions for the cores of calf thymus H1

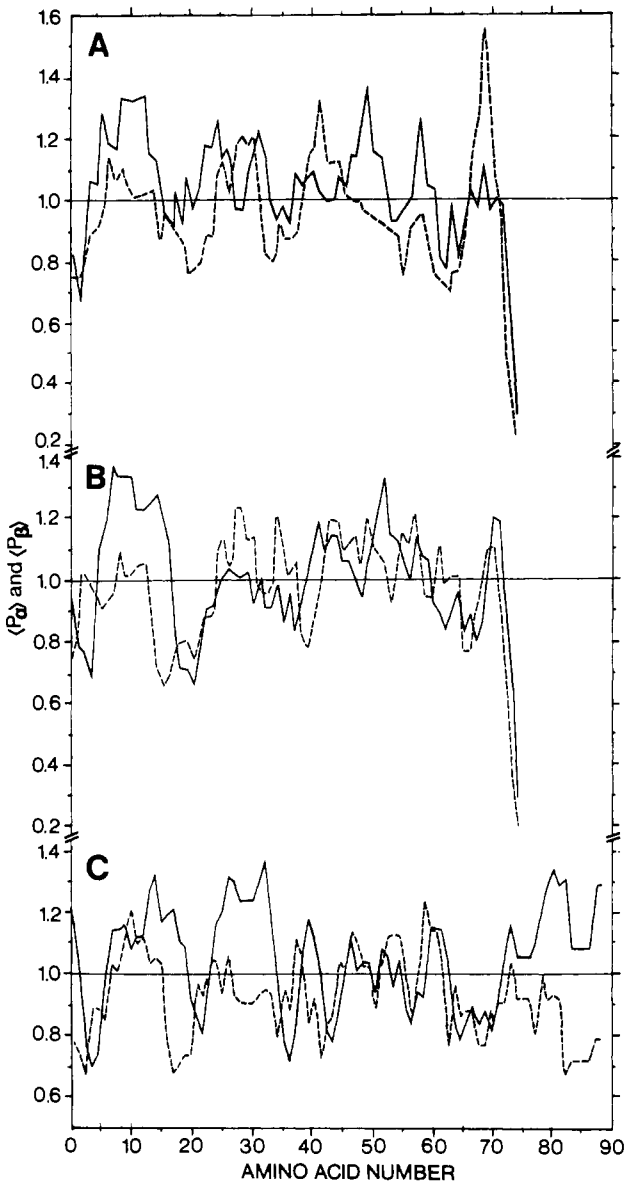

FIGURE 8: Conformational profiles of (A) the protamine-like core, (B) the histone H5 core (residues 22–100), and (C) the histone core of H1 (residues 36–121) from rabbit thymus as redrawn from Fasman et al. (1977). $\langle P_\alpha \rangle$ and $\langle P_\beta \rangle$ stand for the helical and β-sheet potential as defined and calculated according to Chou and Fasman (1978).

histone, chicken erythrocyte H5, and *S. solidissima* PLP. Since, as has just been pointed out, this method does not agree well with the experimental data, the aim of this figure is only for the purpose of further comparison of all these proteins. As can be seen in Figure 8, it is very difficult to find any structural correlation aside from the strong α-helical segment located at the N-terminal end of all three peptides. In fact, Chapman et al. (1978) have argued on this basis that it is possible that the globular structure of histone H5 has evolved separately from that of histone H1.

Aside from the moderate amino acid sequence similarities, the only feature that is common to all these proteins lies in the compact tertiary structure of their trypsin-resistant cores which in turn are flanked by two less structured "tails" at the N- and C-terminal regions of these proteins. The close resemblance of certain short peptide regions in the *Spisula* PLP and H5 (Figure 7) could be taken to suggest an evolutionary relatedness between these proteins. However, there is a paradox involved in such an interpretation. The avian species containing H5 and the molluscan species containing these protamine-like proteins lie on entirely different branches (deuterostome and protostome) of the animal evolutionary tree. Any common ancestor would lie far back at the branching point of the evolutionary paths on which these two proteins are suddenly expressed. Put in another way, there is no direct line of descent from molluscs to chickens.

From the results discussed so far, it seems clear that it is, in principle, not possible to assign the protamine-like component of *Spisula* to the histone H1 family based only on its chemical composition (high contents in lysine and alanine) an on the common existence of a globular trypsin-resistant core. The first of these characteristics (the amino acid composition) may be intrinsic to the specific function assigned to this kind of structural spermatic nuclear protein, as it is reflected by their conserved values among different bivalve species (Ausio, 1986), even in proteins such as the protein ϕ1 from *Mytilus* (mussel) for which there is no evidence at all either for secondary or for tertiary structure (Puigdomenech et al., 1975). On the other hand, the existence of an internal resistant trypsin core may be a more universal protein condition for structural proteins above a certain critical molecular mass. In fact, not only histone H1 but also all the other histones exhibit some globular folded region. Alternatively, the existence of such an internal core may favor the deposition of the protein on the DNA in a certain directionally established arrangement similar to that found for histone H5 in the nucleosomal filament (Lennard & Thomas, 1985). Further comparison of the sequences for the N- and C-terminal regions of the protamine-like protein (Ausio et al., unpublished results) to the corresponding regions both in H1 and in H5 histones may help to finally establish whether there is only an analogy or a real homology between this protein and the proteins from the histone H1 family.

REFERENCES

Allan, J., Hartman, P. G., Crane-Robinson, C., & Aviles, F. X. (1980) *Nature (London) 288*, 675–679.

Ausio, J. (1986) *Comp. Biochem. Physiol., B: Comp. Biochem. 85B*, 439–449.

Ausio, J., & Subirana, J. A. (1982a) *J. Biol. Chem. 257*, 2802–2805.

Ausio, J., & Subirana, J. A. (1982b) *Biochemistry 21*, 5910–5918.

Ausio, J., & van Holde, K. E. (1987) *Eur. J. Biochem.* (submitted for publication).

Ausio, J., Sasi, R., & Fasman, G. D. (1986) *Biochemistry 25*, 1981–1988.

Aviles, F. X., Chapman, G. E., Kneale, G. G., Crane-Robinson, C., & Bradbury, E. M. (1978) *Eur. J. Biochem. 88*, 363–371.

Ayres, W. A., Small, E. W., & Isenberg, I. (1974) *Anal. Biochem. 58*, 361–367.

Bohm, E. L., Strickland, W. N., Strickland, M., Thwaits, B. H., van der Westhuizen, D. R., & von Holt, C. (1973) *FEBS Lett. 41*, 219–222.

Bradbury, E. M., Cary, P. D., Chapman, E. G., Crane-Robinson, C., Danby, S. E., Rattle, H. W. E., Boublik, M., Palau, J. & Aviles, F. J. (1975) *Eur. J. Biochem. 52*, 605–613.

Briand, G., Kmiecik, D., Sautiere, D., Wouters, D., Borie-Loy, O., Biserte, G., Mazen, A., & Champagne, M. (1980) *FEBS Lett. 112*, 147.

Burgess, A. W., Ponnuswamy, P. K., & Scheraga, H. A. (1974) *Isr. J. Chem. 12*, 239–286.

Champagne, M., Pouyet, J., Ouellet, L., & Garel, A. (1970) *Bull. Soc. Chim. Biol. 52*, 377–390.

Chapman, G. E., Aviles, F. J., Crane-Robinson, C., & Bradbury, E. M. (1978) *Eur. J. Biochem. 90*, 287–296.
Chervenka, C. H. (1969) *A Manual of Methods for the Analytical Ultracentrifuge*, Spinco Division, Beckman, Palo Alto, CA.
Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.
Crane-Robinson, C., Danby, S. E., Bradbury, E. M., Garel, A., Kovacs, A.-M., Champagne, M., & Daune, M. (1976) *Eur. J. Biochem. 67*, 379–388.
Elwell, M. L., & Schellman, J. A. (1977) *Biochim. Biophys. Acta 494*, 367–383.
Fasman, G. D., Chou, P. Y., & Adler, A. J. (1977) *Mol. Biol. Mamm. Genet. Appar. Proc. Int. Symp. 1*, 1–52.
Garnier, J., Osquthorpe, D. J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.
Giancotti, V., Russo, E., Cosini, S., Cary, P. D., & Crane-Robinson, C. (1981) *Biochem. J. 197*, 655–660.
Giancotti, V., Russo, E., Gasparini, M., Serrano, D., Del Piero, D., Thorne, W. A., Cary, P. D., & Crane-Robinson, C. (1983) *Eur. J. Biochem. 136*, 509–516.
Hartman, P. G., Chapman, G. E., Moss, T., & Bradbury, E. M. (1977) *Eur. J. Biochem. 77*, 45–51.
Hennessey, J. P., Jr., & Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085–1094.
Johnson, W. C., Jr. (1971) *Rev. Sci. Instrum. 42*, 1283–1286.
Kuntz, I. D., Jr., & Kauzmann, W. (1974) *Adv. Protein Chem. 28*, 239–345.
Labhardt, A. N. (1982) *J. Mol. Biol. 151*, 331–355.
Laemmli, U. K. (1970) *Nature (London) 1227*, 680–685.
Lennard, A. C., & Thomas, J. O. (1985) *EMBO J. 4*, 3455–3462.
Li, H. J., Isenberg, I., & Johnson, W. C., Jr. (1971) *Biochemistry 10*, 2587–2593.
Liao, L. W., & Cole, R. D. (1981) *J. Biol. Chem. 256*, 3024–3029.
Libertini, L. J., & Small, E. W. (1985) *Biophys. J. 47*, 765–772.
Lim, V. I. (1974) *J. Mol. Biol. 88*, 873–894.
Manavalan, P., Mittelstaedt, D. M., Schimerlik, M. I., & Johnson, W. C., Jr. (1986) *Biochemistry 25*, 6650–6655.
Mosteller, F., & Tukey, J. W. (1977) in *Data Analysis and Regression*, Addison-Wiley, Menlo Park, CA.
Mulvery, R. S., Gualtieri, R. J., & Beychok, S. (1974) *Biochemistry 13*, 782–787.
Panyim, S., & Chalkley, R. (1969) *Arch. Biochem. Biophys. 130*, 337–346.
Puigdomenech, P., Cabre, O., Palan, J., Bradbury, E. M., & Crane-Robinson, C. (1975) *Eur. J. Biochem. 59*, 237–243.
Ramachandran, L. K., & Witkop, B. (1967) *Methods Enzymol. 11*, 283–299.
Russo, E., Giancotti, V., Crane-Robinson, C., & Geraci, G. (1983) *Int. J. Biochem. 15*, 487–493.
Subirana, J. A., Cozcolluela, C., & Palau, J. (1973) *Biochim. Biophys. Acta 317*, 364–379.
Tanford, C. (1961) *Physical Chemistry of Macromolecules*, Wiley, New York.
van Holde, K. E. (1971) *Physical Biochemistry*, Prentice-Hall, Englewood Cliffs, NJ.

Equilibrium Binding of Myristoyllysophosphatidylcholine to Bovine Myelin Basic Protein: An Example of Ligand-Mediated Acceptor Association†

Alexander Gow, Donald J. Winzor, and Ross Smith*
Department of Biochemistry, University of Queensland, St. Lucia, Queensland 4067, Australia
Received July 15, 1986; Revised Manuscript Received October 29, 1986

ABSTRACT: The interaction of myristoyllysophosphatidylcholine with bovine myelin basic protein at pH 7.4 and 4.5, I = 0.48, has been investigated by a recycling partition equilibrium technique with Bio-Gel P-2 as the gel phase. Important points to emerge from this direct binding study are (i) that it is a monomeric (not micellar) amphiphile that binds to myelin basic protein, (ii) that the amphiphile binds preferentially to the monomeric form of myelin basic protein, (iii) that this binding to monomer is highly cooperative, (iv) that the similarity of binding behavior in the two environments tested is consistent with the dominance of a hydrophobic contribution to the protein–amphiphile interaction, and (v) that the self-association of myelin basic protein in the presence of phospholipid [Smith, R. (1982) *Biochemistry 21*, 2697–2701] must reflect the aggregation of a protein–amphiphile complex(es) coupled with concomitant release of some lipid. These findings are then related to earlier nuclear magnetic resonance and circular dichroism studies in which the results were interpreted on the basis that myelin basic protein bound preferentially to micellar phospholipid.

Central nervous system myelin contains two predominant proteins in association with an unusually high lipid content. One of these proteins, the proteolipid protein, displays many of the properties considered characteristic of an intrinsic membrane protein that is retained within the lipid bilayer by hydrophobic interactions. The other major protein, myelin basic protein (MBP),[1] is less easily classified. Evidence for its classification as an extrinsic membrane protein associated

† This work was supported by National Health and Medical Research Council and University of Queensland Special Project Grants (to R.S.) and also in part by funds from the Australian Research Grants Scheme (to D.J.W.).

[1] Abbreviations: cmc, critical micelle concentration; DMPC, dimyristoylphosphatidylcholine; lysoMPC, myristoyllysophosphatidylcholine; MBP, myelin basic protein; NMR, nuclear magnetic resonance.

0006-2960/87/0426-0982$01.50/0