NMR and Computational Studies on the Conformational Folding of the Biomineralization Template, Phosphophoryn

Thesis by John Spencer Evans

In Partial Fullfillment of the Requirements for the Degree of Doctor of Philosophy

California Institute of Technology, Pasadena, California

1993 A.D.

(Defended October 12, 1992)

"The horror, the horror!....exterminate all the brutes!"

--Joseph Conrad, "Heart of Darkness"

ABSTRACT: Bovine Dentine Phosphophoryn (BDPP) is a member of the "Asp-rich" superfamily of template macromolecules known as polyelectrolyte mineral matrix proteins (PMMPs). Using 1-D and 2-D NMR multinuclear spectroscopy, protein sequencing, solid-phase peptide synthesis, and molecular modeling, we investigated the capacity of BDPP to fold in solution under certain conditions. It is believed that the folding properties of a template macromolcule are important in the formation of an inorganic mineral phase. In this report, we have established the following regarding BDPP structure:

I. BDPP sequence organization can be conceptualized as three types of domains: *polyelectrolyte calcium binding domains* (PCBD), *hinge domains* (HD), and the *hydrophobic domains* (HC). The PCBD regions can be further subdivided into $(Asp)_n$, $(PSer)_n$, and $(PSerAsp)_n$-containing homopolymer and heteropolymer sequence stretches. The HD contain predominantly neutral or uncharged amino acids such as Ser, Gly, and Pro. These HD domain sequences are flanked on either side by PCBD regions.

II. BDPP contains a number of Lys sidechains (44/1000 residues) which are believed to form ion pairs with either Asp, Glu, or PSer residues in the protein. The function of this salt-bridging is not understood at this time, but it may be responsible for maintaining the protein template molecule in a conformation that retains a high negative charge density.

III. At low pH, or in the presence of divalent cations, BDPP assumes a global conformation that is condensed in particle size. At neutral pH under conditions of low ionic strength and in the absence of divalent cations, this global conformation converts to an extended form. The folding transition between these two conformers is mediated by conformational change in the HD regions which are flanked by specific PCBD sequence regions.

IV. The PCBD regions possess some degree of tertiary and secondary structural organization in the absence of divalent cations at low ionic strength. This folding permits the surface charge density of these regions to remain high, relative to that of a random coil conformer.

V. BDPP exhibits a selectivity in terms of divalent cation binding sites. Under conditions of low ionic strength and divalent cation depletion, the addition of Cd (II) to BDPP leads to binding at various PCBD sequence stretches, according to the following order:

$$(PSer)_n > (PserAsp)_n > (Asp)_n$$

VI. Modeling studies conducted on PCBD sequence peptides [$(Asp)_{20}$, $(PSer)_{20}$, and $(PSerAsp)_{10}$] indicate that there is a sequence preference for certain conformers in the presence of Na^+, i.e., $(Asp)_{20}$ forms "supercoils", $(PSer)_{20}$ forms "hairpins", and $(PSerAsp)_{10}$ forms "spirals" or "distorted" hairpins. Each of these conformers features some degree of sidechain

folding and/or peptide backbone secondary structure, in support of Mann's hypothesis, as well as the experimental data obtained for BDPP.

VII. A peptide mimetic which represents a PCBD-Hinge-PCBD motif of rat α-phosphophoryn, was constructed using FMOC solid-phase peptide synthesis. Using 2-D NMR spectroscopy and monitoring α-CH and sidechain β-CH_2 proton chemical shifts, we can demonstrate that this peptide mimetic folds or collapses under conditions of low pH and divalent cation addition, in a manner similar to that observed for intact BDPP.

Ac*know'ledg*ment (ak-nol'ej-ment),n. 1. the act of admitting; confession

I found this definition somewhere between **aclastic** and **acknowledger** in *Webster's New Universal Unabridged Dictionary.* I 've never come across the word aclastic before, and I haven't the faintest idea what it means; but I'll lay odds that acknowledger pretty much describes me at this moment. I'm in the process of finishing the assembly of this thesis, and for some reason I put off writing this section until the last moment. Call it sloth, I guess.

At any rate, this year and this work have been very good to me in a number of ways. Things are falling into place. The Pittsburgh Pirates have just won their third straight NL East Division title. My long term job prospects are brightening. And, the field of Biomineralization is starting to take off and become accepted in the Scientific Community (did that need to be capitalized?). So it is appropriate at this time to *admit* -- no, dammit, *testify* is more the word--that a lot of good people have had a lot to do with all of this success. I just happened to be there.

First, a word of thanks to my advisor, Sunney I. Chan. I came into Sunney's group back in the Spring of '88, with a D.D.S. degree, no Ph.D, a background in Cell Biology, and not much of a chemistry background. Sunney basically took a risk on me -- he gave me the support, the lab space, and left me alone to do whatever it was I wanted. Either I was going to succeed somehow, or fail miserably. I didn't fail, and it is to his credit that I succeeded as well as I have. When it came time to become a Ph.D candidate, Sunney backed my choice and was very supportive of my efforts here. His suggestions and comments helped focus my efforts, and gave me a different perspective from my own--a valuable thing, believe me. To him I owe a lot of thanks, and a multi-course Chinese dinner somewhere down the line. Continued success to you, Sunney!

Who's next? Well, I collaborated with a number of people in the 4.5 years here at Caltech. Professor Bill Goddard, with whom I spent a lot of time learning and testing computational methods, was like a second advisor to me. He was patient many times as I struggled to learn theoretical chemistry (and did!). He is a very creative scientist with unlimited energy and enthusiasm, a good role model for anyone who aspires to a career as a Professor. Bill's graduate student, Alan

Mathiowetz, who is also graduating soon, developed the Monte-Carlo program which I applied to the polyelectrolyte problem. Siddharth Das Gupta and Jim Gerdy from Bill's group were also helpful with ideas, questions, etc. Ken Graham, Peter Dervan's student (who's also graduating this month), ushered me into the synthetic peptide world -- we fought long and hard to make a simple 10-mer poly-L-(Asp) peptide using t-Boc chemistry, and failed miserably at it. But it steered me away from a method that wouldn't be of any use in this field. However, Tom Chiu, an undergraduate from Yale who spent a summer with us, made INTERV1 a success using FMOC chemistry. Win a few, lose a few, eh? And a BIG thanks to Tammy Bauer, Jane Sanders, and Jon Racs of the Hood Group, who sequenced, analyzed, and sprayed everything I gave them.

A special thanks has to go to the Chan Group. No, please -- don't try to stop me. Laboratories are notorious "revolving doors": people come and go so fast you sometimes don't get to know them. But I did get to know a few here, and their friendship made it a pleasure to come to the lab each day and get down to business. Special thanks go to Randy Larsen, Elise Gabriel, Walther Ellis, Joel (Jor -El) Morgan, Sig, Ted, Dorothy (from the "Land of DOZ"), Jack, Qizhi, Mr. Pan, Scott Ross, Zhuyin (and "Baby"), Mike Stowell, Gary Mines, Bog Kuszta, Zehua, Paul Smith, Wei Li, Drew Schiemke, Wa-To, Nancy Vogelaars....we'll be here all night if I keep this up. A special thanks and good wishes to the "New Wave" of NMR graduate students now in place: Silvia, Ruth Ann Chan and her Soda Pop Shop...and of course Dennis Anderson, who, to quote the Legend, Dr. Robert J. Driscoll, "was a constant presence in the Lab during my last months here in the group..."

Finally, I am indebted to the following people. First, to Heinz Lowenstam, Professor emertus of the Geology Division at Caltech, and the "Godfather" of Biomineralization. His life is filled with knowledge and a true pioneering spirit, one which I try to emulate. He is one of Caltech's true "living legends". Go see him sometime and say "Hi" for me. Next, to Dr. Thomas J. Valega, Director of the National Institute of Dental Research, who gave me the NRSA fellowship, and in essence allowed me to do whatever I wanted with it, including extending it for 4.5 years to get a Ph.D. I truly owe him a dinner! Special thanks go to my first research advisor, Dr. Arthur Veis, of Northwestern, whose support and kind words

have helped me in my career more than anything else. Lastly, I want to express my love and thanks to two women who have been my strength and source of love over the years. To my mother, who as a single parent under trying conditions raised my sister and I to be as creative as we both are. I love you, Mom (and you, too, Hollis). And to my wife, Annette, who supported my decision to get this degree, lived through all the long hours, weekends...well, you know what a graduate student has to do. She most of all deserves my thanks and reciprocal support. I love you, Babe!

--Pasadena, California
September 30, 1992

Table of Contents

List of Tables and Figures

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Chapter Six

Chapter Seven

Chapter Eight

Chapter Nine

One

INTRODUCTION

BACKGROUND AND SIGNIFICANCE

The Occurrence of Biominerals in Nature

Biomineralization, as defined by Lowenstam [1981], is a cell-mediated process in which an organic polymer matrix, comprised of several classes of macromolecules, acts as a scaffolding or template for the synthesis of an inorganic solid phase. The resulting biocomposite material has been hypothesized to perform at least three necessary functions for life: (1) *biomechanical support*, (2) *ion homeostasis*, (3) *gravity sensor*, and (4) in the instance of biomagnetic minerals, *organism orientation in the Earth's magnetic field* [Williams, 1984, 1988; Skiles, 1985; Yorke, 1985; Mann *et al*, 1991; Blakemore,1975; Blakemore *et al.*, 1981; Blakemore *et al.*, 1985]. The occurrence of biominerals in Nature is quite diversified; nearly every kingdom possesses representative organisms which produce an inorganic matrix [Table 1; see Lowenstam and Weiner [1989] for a complete review]. To date, the biominerals discovered thus far encompass 60 different inorganic mineral phases of varying composition. These include silicates, $[Si(OSi)_n(OH)_{4-n}]_m$, and the phosphates, sulfates, halides, oxalates, and oxides of Groups Ia, IIa, and several *d*-block metals [for a review, see Lowenstam and Weiner, 1989, Wilbur and Simkiss, 1989; Mann, 1988b]. Of these, the silicates and calcium-based minerals represent the most widely occurring group of biominerals in the ecosphere [Lowenstam and Weiner, 1989; Williams, 1984]. In addition to the inorganic minerals, several biogenic organic crystals have been identified (Table I); these organic crystals apparently fulfill functions analogous to their inorganic counterparts in Nature.

Table I: Overview of Biogenic Inorganic Minerals.

Adapted from Lowenstam and Weiner (1989), with amendments from recent data obtained from Zierenberg and Schiffman (1990) and Macaskie *et al.*, (1992).

Table I: Distribution of Biogenic Minerals[0]

Mineral Name[1]	Formula	Occurrence (Kingdom)[2]
Ca		
Calcite	$CaCO_3$	M, Pr, F, Pl, A
Aragonite	$CaCO_3$	M, Pr, Pl, A
Vaterite	$CaCO_3$	Pl, A
Monohydrocalcite	$CaCO_3*H_2O$	M, A
AHC[3]	$CaCO_3*nH_2O$	M, F
Fluorite	CaF_2	A
Amorphous fluorite	CaF_2	A
ACP[4]	$Ca_x(PO4)_y$	Pr, A
DCPD[5]	$CaHPO_4*2H_2O$	A
beta-TCP[6]	$Ca_3(PO_4)_2$	A
DCPA[7]	$CaHPO4$	A
OCP[8]	$Ca_8(HPO_4)_2(PO_4)_4*5H_2O$	A
HAP[9]	$Ca_{10}(PO_4)_6(OH)_2$	M, Pr, F, A
Amorphous pyrophosphate	$Ca_2(P_2O_7)*nH_2O$	A
Earlandite	$Ca_2(C_6H_5O_7)$	Pl, A
Calcium tartrate	$Ca(C_4H_4O_6)$	Pl
Calcium malate	$Ca(C_4H_4O_5)$	Pl
Caclium oxalate	$Ca(C_2O_4)$	Pr, Pl, A
Weddelite	$Ca(C_2O_4)*2H_2O$	F, Pl, A
Whewellite	$Ca(C_2O_4)*H_2O$	Pr, F, Pl, A
Gypsum	$CaSO_4$	Pr, A

Table I: Distribution of Biogenic Minerals

Mineral Name	Formula	Occurrence (Kingdom)
Mg		
Glushinskite	$Mg(C_2O_4)*2H_2O$	F
AMCP[10]		A
Whitlockite		M
Struvite	$[(NH_4)MgPO_4]*6H_2O$	M, A
Protodolomite	$MgCa(CO_3)_2$	A
$Ca_3Mg_3(PO_4)_4$	$Ca_3Mg_3(PO_4)_4$	A
Ba		
Barite	$BaSO_4$	Pr, A
Sr		
Celesite	$SrSO_4$	Pr
K		
Jarosite	$KFe_3(SO_4)_2(OH)_6$	M
Hieralite	K_2SiF_6	A
Na		
Sodium urate	Na	A

Table I: Distribution of Biogenic Minerals

Mineral Name	Formula	Occurrence (Kingdom)
Zn		
Sphalerite	ZnS	M
Wurtzite	ZnS	M
Si		
Opal (silicate)	$[Si(OSi)_n(OH)_{4-n}]_m$	Pr, Pl, A
Mn		
Todorokite	$[(Mn,Ca,Mg)Mn_3]^{4+}O_7 * H_2O$	M
Birnessite	$Na_4Mn_{14}O_{27} * 9H_2O$	M
Maganese Oxalate	$Mn_2(C_2O_4) * 2H_2O$	F
Pb		
Galena	PbS	M
Hydrocerussite	$PbCO_3$	F
Cu		
Chalcopyrite	$CuFeS_2$	M
Copper (II) Oxalate	$Cu(C_2O_4) * nH_2O$	Pr, Pl, A

Table I: Distribution of Biogenic Minerals

Mineral Name	Formula	Occurrence (Kingdom)
Ag		
Prousite	Ag_3AsS_3	M
Peareite	$Ag_{14.7 - x}Cu_{1.3 + x}As_2S_{11}$	M
U		
HUP	HUO_2PO_4	M
Fe		
Magnetite	$[Fe(II)][Fe(III)]_2O_4$	M, Pr, A
Goethite	α-Fe (III)O[OH]	F, A
Lepidocrocite	γ-Fe (III)O[OH]	A
Ferrihydrite	$9\ Fe_2O_3{*}9H_2O$	M, F, Pl, A
Amorphous ilmenite	$FeTiO_3$	M
AlO[11]	Fe_2O_3	A
Vivianite	$[(Fe(III))_3(PO_4)_2{*}8H_2O]$	M, A
AHFe (III) P[12]	Fe (III) $(PO_4){*}nH_2O$	M, A
Greigite	$Fe(II)[Fe(III)]_2S_4$	M
Mackinawite	$[Fe,Ni]_9S_8$	M
Pyrite	FeS	M
Hydrotroilite[13]	$Fe(II)_nS_n{*}nH_2O$	M
$KNa_3[Fe_{15}Mg_2]^{5+}[PO_4]_3[OH]_3$	$KNa_3[Fe_{15}Mg_2]^{5+}[PO_4]_3[OH]_3$	A

---

0 Adapted from Lowenstam and Weiner, 1989.

[1]Geochemical/chemical classification and formula, with the exception of the calcium phosphates (see footnotes 4-9), taken from *Encyclopedia of Minerals, 2nd Edition* (Roberts, M.L., Campbell, T.J., and Rapp, G.R. Jr., Eds.) Van Nostrand Reinhold Company, New York, New York, 1985.

[2]Abbreviations: M = monera; Pr = prototcista; F = fungi; Pl = plantae; A = Animalia

[3]Amorphous hydrous calcite

[4]Amorphous calcium phosphate; structure variable. After Nancollas (1988).

[5]Dicalcium phosphate dihydrate; also known as *brushite*, ibid.

[6]beta-tricalicum phosphate; also known as *whitlockite*, ibid.

[7]Dicalcium phosphate; also called *monetite*, ibid.

[8]Octacalcium phosphate, ibid

[9]Hydroxyapatite. The apatites comprise several classification, depending on their substitution. *Dahllite* refers to carbonate substituted apatite, with the CO_3^{2-} group occupying lattice sites of either the hydroxyl (type ?) or the phosphate (type ?). *Francolite* refers to fluoroapatite [$Ca_{10}(PO_4)_6(F)_2$], with fluorine substitution at the OH sites. The general chemical formula for biogenic apatites can be written as (Nancollas, 1988):

$$(Ca, Sr, Mg, Na, H_2O, \{\})_{10}(PO_4, HPO_4, CO_3, P_2O_7)_6(OH, F, CO_3, Cl, H_2O, O, \{\})_2$$

where "{}" refers to lattice vacancies.

[10]Amorphous magnesium calcium phosphate.

[11]Amorphous iron oxide

[12]Amorphous hydrous iron phosphate

[13]Colloidal hydrous ferrous sulfide; species status uncertain.

All biogenic minerals exist as hydrates, in which the water molecules exist as surface-absorbed species [Young, 1975; LeGeros *et al.*, 1989; Lerner *et al.*, 1989, Mann, 1988b; Tomazic *et al.*, 1989; Menanteau *et al.*, 1984; Misra, 1986; Wilbur and Simkiss, 1989], or, as integral constituents of the lattice [Brown, 1966; Nelson and McLean, 1984; Meyer and Eanes, 1978; Young and Brown, 1982; Mooney and Aia, 1961]. Of the known biogenic minerals, approximately 80% exist in the crystalline state, the remaining 20%, in the paracrystalline or amorphous state [Lowenstam and Weiner, 1989; Wilbur and Simkiss, 1989]. The distinctions between amorphous and crystalline phases lie primarily in the presence or absence of lattice ordering in the mineral phase, where amorphous materials have been designated as those which possess no long-range translational and orientational order ($\geq$ 20 Å) , but exhibit short-range (2 - 5 Å) or medium-range (5 - 20 Å) ordering [Elliot, 1991]. The widespread presence of amorphous minerals raises interesting questions regarding their existence. In several systems, such as the calcium phosphates, amorphous minerals such as amorphous calcium phosphate [molecular formula $Ca_3(PO_4)_2$ or $Ca_3(PO_4)_{1.87}(HPO_4)_{0.2}$, abbreviated as ACP] act as precursor phases in the formation of a crystalline mineral, hydroxyapatite [molecular formula: $Ca_{10}(PO_4)_6(OH)_2$, abbreviated as HAP][Heughebaert *et al.*, 1990a,b; Iijima *et al.*, 1991]; in other examples, amorphous minerals are the primary inorganic solid phase that forms [e.g., amorphous ferric phosphate hydrate, $Ca_{0.23}Mg_{0.20}Fe_{1.25}(PO_4)(OH)_{1.58}{}^*(H_2O)_{1.88}$] [Lowenstam and Rossman, 1975]. The role and evolutionary significance of amorphous minerals in biomineralization has yet to be fully elucidated, but it is believed that the structural and fracture properties that amorphous materials afford over crystalline materials is an important one [Berman *et al.*, 1990].

Control and Regulation of the Biomineralization Process.

In general, the process of biomineralization has evolved into several key events that allow for the control of inorganic mineral phase propagation (Lowenstam and Weiner, 1989).[1] These key steps are:

(1) *Compartmentalization.* In most of the biogenic mineral systems studied thus far, cells delineate a space whose boundaries are selectively permeable. This space is either lined with cells, or, is a non-cellular vesiculated or membrane-bound compartment. This compartmentalization allows selective concentration of specific cations and counteranion pairs which will constitute the inorganic mineral phase.

(2) *Supersaturation.* The net concentrations of the cation and counteranion are raised to levels which exceed the K_{sp} of the inorganic ion pair. This condition may result in precipitation of the ion pair from solution; or, it may lead to the formation of a metastable solution from which precipitation will occur.

(3) *Initiation/Nucleation.* The inorganic ion pair precipitate is formed. Upon this nucleation foci, cation and counteranion apposition occurs. The initiation of foci formation and apposition

---

[1] Two types of processes that nucleate and control crystal growth in Nature have been defined (Mann, 1988b). The first, called *Chemically-controlled*, involves regulation of mineralization by such factors as ion transport rates, levels of ion saturation, pH, diffusion, etc. The second, *Matrix-controlled*, involves template interactions. These can further be subdivided into *matrix-induced* (i.e., specific activation of nucleation at the matrix surface, but structure and orientation are determined by chemical factors) and *matrix-mediated* (i.e., characterized by a high level of crystallochemical specificity in the structure and orientation of the incipient nuclei). In this report, we will be discussing matrix-controlled biomineralization, with no distinction being made between matrix-mediated or matrix-induced.

may be regulated by initiator-regulator template macromolecules, or smaller organic molecules that can act as inhibitors or accelerators of apposition. This process may also be chemically-controlled by inducing changes in local ion concentrations and pH [Mann, 1988b].

(4) *Directional Growth*. The appositional growth of a crystal or amorphous phase may be epitaxial in nature, where selective elongation along one or more of the crystal axes takes place. This directionality has been shown to be regulated by the selective binding of initiator-regulator template macromolecules to a crystal or mineral face [Addadi and Weiner, 1987; Addadi *et al.*, 1987; Weissbuch *et al.*, 1991].

The Role of Organic Macromolecules in the Regulation of Biomineralization.

Clearly, Nature has evolved a means of controlling the formation and composition of inorganic compounds. As discussed above, one way of regulating the processes of nucleation and directional growth involves the participation of organic macromolecules in the matrix. These organic macromolecules can be categorized into two fundamental groups: (1) *scaffolding* and (2) *templates*.

As the name implies, *scaffolding* macromolecules are the framework in and around which the inorganic mineral phase forms. The combination of an inorganic mineral with a organic polymer matrix results in a composite material whose compressive, tensile, shear, and elastic properties are strikingly different from those properties of the original inorganic mineral phase. In some instances, such as the biogenic calcium carbonates (aragonite and calcite, crystalline phase), these modified properties approach that of glasses [Moradian-Oldak *et al.*, 1991; Berman *et al.*, 1988, 1990]. Some examples of scaffolding macromolecules include Type I

collagen, β-silk fibroin, and β-chitin (Table II). In general, the scaffolding macromolecules comprise a significant percentage (70% or greater) of the total macromolecule composition in the extracellular matrix [Veis, 1982].

The second class of organic molecules, *templates,* are a diverse group which are found in low abundance in the extracellular matrix (<10%)[Veis, 1982; Mann, 1988b]. Their role is paradoxical, in that they are responsible for the initiation of the inorganic mineral phase, and, they also regulate or limit the propagation of the inorganic phase, such that there is epitaxial or directional crystal growth. These molecules have been demonstrated to bind specific cations with high affinity and selectivity, and, also bind to the surface(s) of a forming mineral phase [Williams, 1984; Addadi and Weiner, 1985]. That they are represented in small amounts in the extracellular matrix indicates that the role of these macromolecules is a catalytic one. This has been also confirmed by in vitro experimentation, where nanomolar levels of these macromolecules are required to initiate crystal growth [Addadi *et al.*, 1987, 1991; Addadi and Weiner, 1985; Wheeler *et al.*, 1981; Berman *et al.*, 1990; Berman *et al.*, 1988].

Template Macromolecules

Conceptually, a template can be thought of as a molecular surface that serves as an interface for molecular recognition, i.e., *it possesses active sites that interact stereospecifically with other atoms or molecules* [Weissbuch *et al.*, 1991][Figure 1,2]. This can serve two purposes. For the nucleation event in biomineralization, these active sites on the template will serve as binding sites for metal ions, and an inorganic lattice begins to propagate in an epitaxial direction from the initial forming face of the metal atom array. For inhibition, the template binding to an array of exposed

Table II: List of Scaffolding Macromolecules.

Adapted from Lowenstam and Weiner (1989), Veis (1988), Mann (1988b).

Table II

Scaffolding Macromolecules

Macromolcule	Tissue Source	Biomineral[1]
Type I Collagen	Bone, Dentine, Tendon[2]	ACP, DCPD, β-TCP, OCP, HAP
Type II Collagen	Cartilage	ACP, DCPD, β–TCP, OCP, HAP
α-Chitin[2]	Cuticle (*Arthropoda*)	Calcite ($CaCO_3$)
	Tubes (*Annelida*)	Calcite/aragonite ($CaCO_3$)
β-Chitin[3]	Shell (*Mollusca*)	Calcite/Aragonite ($CaCO_3$)
"Silk-Fibroin" Protein[4]	Shell (*Mollusca*)	Calcite/Aragonite ($CaCO_3$)
"Glycine-Rich" Protein[5]	Shell (*Articulata*)	Calcite ($CaCO_3$)
"Filamentous" Protein[6]	Spicule (*Porifera*)	$[Si(OSi)_n(OH)_{4-n}]_m$
"Chitin-Protein[7]	Cuticle (*Ectoprocta*)	Calcite ($CaCO_3$)
Complex"	Shell (*Brachiopoda*)	Francolite (HAP)

---

[1] Predominant mineral phase(s) affiliated with the denoted macromolecule.

[2] May be found in complex with protein(s).

[3] May be found in complex with protein(s).

[4] Silk-fibroin-like protein has been identified to have repeating β-sheet structure arising from Gly-X-Y repeat sequence, similar to fibroin. This protein is found complexed with β-chitin in Mollusca shell matrix.

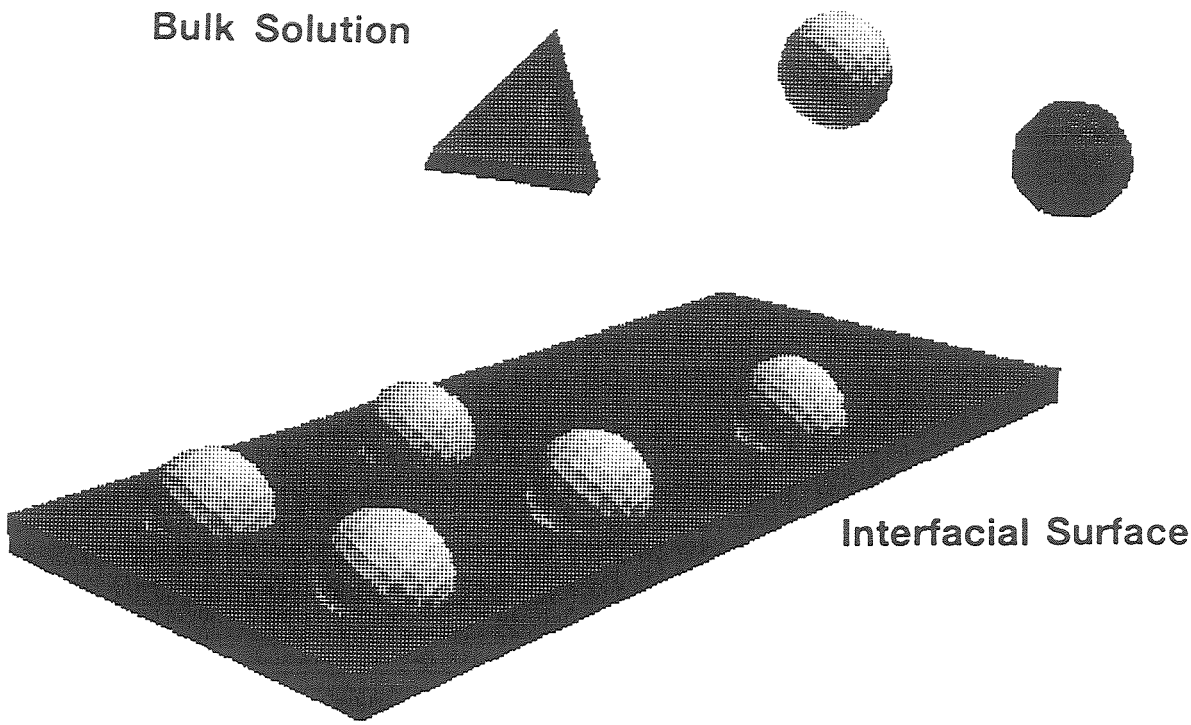
[5] Insufficient data available.

[6] Insufficient data available.

[7] Insufficient data available.

Figure 1: Conceptualization of a Template.

A template (represented by a planar surface) can be thought of as a molecular surface that serves as an interface for molecular recognition, i.e., *it possesses active sites that interact stereospecifically with other atoms or molecules* . Here, only atoms that possess specific properties (e.g., valence state, charge/mass ratio) will occupy sites on the template.

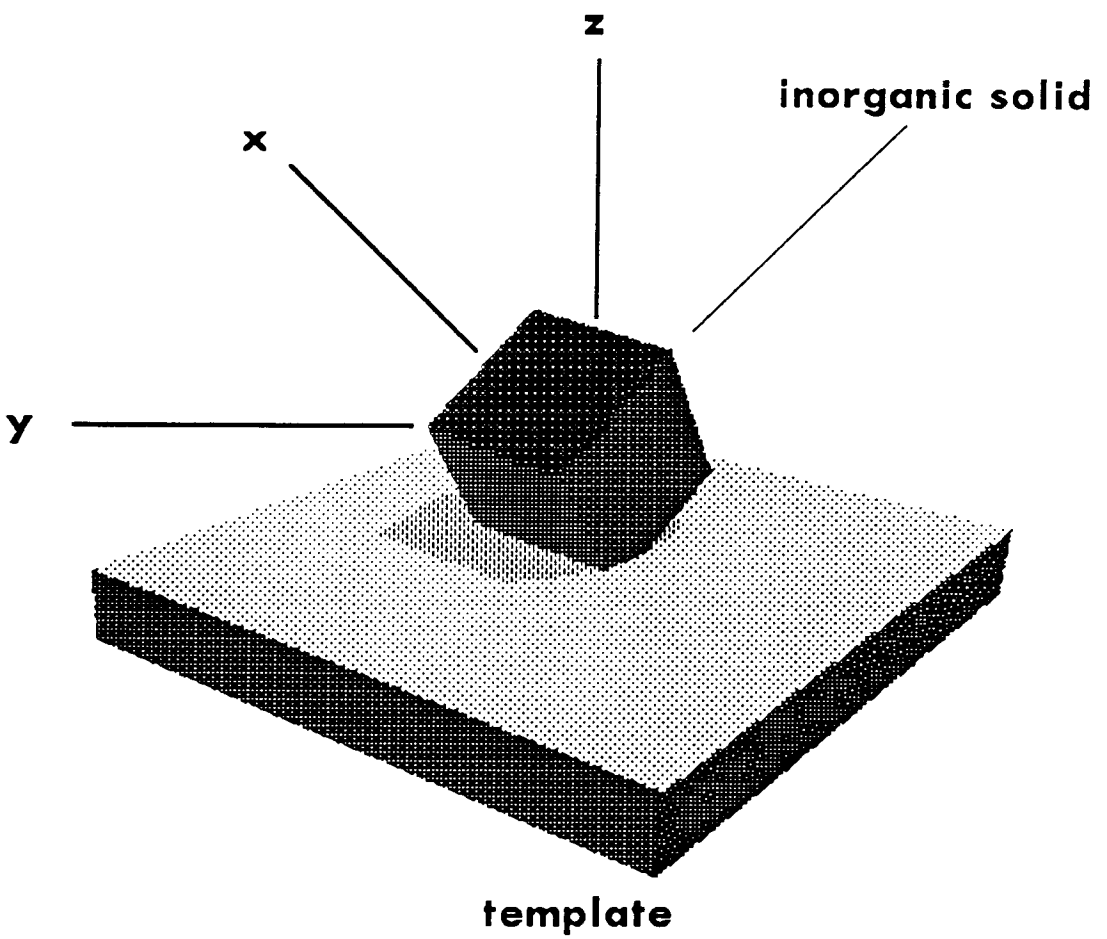


Template: molecular surface that serves as an interface for molecular recognition, i.e.,

it possesses <u>active sites</u> that interact stereospecifically with other atoms or molecules

Figure 2: Conceptualization of Template Function in Biomineralization.

Pictoral representation of a inorganic solid nucleating off of a complementary face of a molecular template surface. Normally, metal atoms deposit onto lattice sites on the template; these sites are usually anionic atoms that mimic a corresponding array of electronegative atoms in a crystal or amorphous solid lattice. Nucleation then proceeds epitaxially along a defined axis (in this case, the *z*-axis; in crystallographic terms, it would be the *c*-axis).


z
inorganic solid
x
y
template

metal atoms on a crystal surface can act to prevent further counteranion-cation apposition at this blocked face. In biomineralization, organic template macromolecules[2] fall into several distinct classes:

(1) *Anionic polysaccharides.*

(2) *Lipids and Fatty Acids.*

(3) *Hydrophobic proteins.*

(4) *Polyelectrolyte Mineral Matrix Proteins (PMMP).*

For the purposes of subsequent discussion, a brief overview of classes (1) through (3) will given below, with a more detailed examination of acidic or anionic proteins to follow. The reason for including this discussion of the template macromolecule classes (1-3) first is twofold: (1) the ion and mineral binding chemistry of anionic polysaccharides and fatty acids/lipids may be similar to those exhibited by the polyelectrolyte mineral matrix proteins, and, (2) the PMMPs possess hydrophobic domains that may be involved in biomineral regulation.

(1) Anionic Polysaccharides. Complex polysaccharides containing anionic chemical groups have been identified in a number of mineral matrices (Lowenstam and Weiner, 1989). These anionic polysaccharides may have a dual role in biomineralization: to increase the net concentration of cations, and restrict the availability of cations for mineral deposition. The best studied of these systems is the cartilage matrix of chordata, where anionic

---

[2]By the term *macromolecule* we are limiting our discussion to templates whose molecular weights are in excess of 0.5 kD. It should be noted that smaller molecules also perform the function of templates, e.g., dicarboxylic acids (Mann, 1988b).

polysaccharides are found to exist with proteins in a multimeric complex called *proteoglycan* (PG)(Schwartz *et al.*, 1985; Veis, 1988). A Ca^{2+} binding anionic polysaccharide has also been identified in the coccoliths of *Emiliania huxleyi* (de Jong *et al.*, 1976)

The PG complex, possessing an average molecular weight of 2000 to 3000 kD, is comprised of several macromolecules: (a) Hyaluronic acid. (b) Core Protein. (c) Glycosaminoglycans. (d) Link Protein (Hardingham and Muir, 1973; Hascall and Heinegard, 1974; Schwartz *et al.*, 1985; Hassel *et al.*, 1986). Of particular interest are the glycosaminoglycan, or GAG, polymer chains. These are covalently attached to the central core protein, which in turn attaches non-covalently to the hyaluronic acid polymer chain. Several types of anionic polysaccharides comprise the GAG complex: (1) chondroitin-4- and chondroitin-6-sulfates. Chondroitin sulfates are polyanionic polymers of a repeating disaccharide comprised of alternating glucuronic acid and sulfated N-acetylgalactosamine residues. In chondroitin sulfate proteoglycan (CSPG), the average chondroitin chain has a molecular weight of 20 kD, and up to 80-100 chains are found to be covalently attached to the core protein (Krueger *et al.*, 1990; Lohmander *et al.*, 1980; Nilsson *et al.*, 1982). (2) keratan sulfate. Keratan sulfate is a repeating heteropolymer of galactose and sulfated N-acetylglucosamine, of an average molecular weight of 4-8 kD, with approximately 15-20 keratan sulfate chains linked to a single core protein (Schwartz *et al.*, 1985; Krueger *et al.*, 1990; Lohmander *et al.*, 1980; Nilsson *et al.*, 1982). Both GAGs are linked to the core protein via seryl or threonyl-*O* linkages. Thus, carboxylic acid and sulfate monoesters are the two anionic groups which comprise the GAG molecule.

The functions of the proteoglycans are inexplicably tied to the anionic carbohydrate moieties and their polyelectrolytic properties. The large hydrodynamic size of PG is created by water molecule aggregation around the polyanionic regions of the PG molecules

[Muir, 1980]. The charge repulsion of the PG complex results in the orientation of GAG sidechains in a spatially extended conformation. These two properties result in modifying the cartilage matrix to a compressible, load-bearing material [Poole *et al.*, 1989]. In a quadrupolar multinuclear NMR ion study, Lerner and Torchia (1986) investigated the ion binding properties of cartilage PG and chondroitin sulfates. Using ^{23}Na, ^{39}K, ^{25}Mg, and ^{43}Ca ions and measuring quadrupolar relaxation rates, it was found that ^{39}K and ^{23}Na did not interact strongly with either PG aggregate or chondroitin sulfate in solution, but that the PG aggregate bound both ^{25}Mg and ^{43}Ca ions at reduced affinities [Lerner and Torchia, 1986]. This would indicate that the ionic dependence of the mechanical properties of cartilage is due primarily to outer sphere electrostatic shielding and not due to specific ion affinities. Interestingly, phosphate anion can readily displace calcium ions from PG molecules, and this "ion-exchange" effect of PG molecules has been purported to mediate the release of Ca^{2+} ion for calcium phosphate precipitation [Hunter, 1987]. The observations that PG aggregate binds to hydroxyapaptite [Hunter *et al.*, 1987], and inhibits the growth of HAP seed crystals <u>in vitro</u> [Chen and Boskey, 1985], suggests that the PG macromolecule may regulate the growth of HAP in cartilage matrix. When proteoglycans interact with hydroxyapatite in vitro, the S-O modes of the sulfate groups in the IR spectrum disappear, suggesting that the interaction between PG and hydroxyapatite is via the sulfate groups [Embery and Rolla, 1980]. In an earlier work, Chen *et al* (1984) demonstrated that PG monomer A1D1D1 and chondroitin-4-sulfate were less effective in inhibiting HAP growth in solution, but that the PG aggregate A1 fraction, containing multimers of A1D1D1, exhibited greater inhibition of HAP growth. Using chondroitinase AC digestion, chondroitin-4-sulfate and chondroitin-6-sulfate chains were cleaved from PG aggregates; these "clipped" PG preparations were shown to inhibit β-tricalcium phosphate formation in solution to a greater extent than isolated chondroitin sulfate chains and cathepsin D

protease-digested proteoglycans [Sajdera and Hascall, 1969; Dziewiatkowski and Majznerski, 1985].

These results, together with the results obtained from other in vitro experiments on cartilage PG monomer and aggregate fractions, provide evidence in support of a hypothesis that PG structural modification precedes the onset of cartilage mineralization [Chen *et al.*, 1984]. This modification may involve extracellular matrix proteases or glycosylases, which "clip" away the PG chains as the cartilagenous matrix undergoes maturation. It is also possible that other chemical modifications occur to PG molecules in the matrix. With a decrease in PG aggregate concentration, HAP synthesis from supersaturated solution begins to increase in the absence of PG inhibitor. Decreases in PG concentration have been observed in vivo in cartilage matrix as it undergoes maturation [Blumenthal *et al.*, 1979; Howell *et al.*, 1969; Cuervo *et al.*, 1969; Howell and Pita, 1976].

(2) Lipids and Fatty Acids. The basic concept of fatty acid/lipid template macromolecule-controlled crystal growth is that the long-chain hydrophobic moieties stabilize the organic phase into one or more packing orientations, and that the polar headgroups act to bind, nucleate and structurally organize the mineral phase [Mann, 1988]. Since a considerable amount of research has been done in this particular area of template-controlled crystal growth, it is appropriate at this point to examine this particular phenomenon in more detail.

In vivo evidence for lipid/fatty acid-induced mineralization has been observed in a number of calcium phosphate-based biogenic minerals. Ultrastructural studies of the calcium phosphate (HAP) mineral matrix (bone, cartilage) have established the presence of lipid vesicles, termed *matrix vesicles*, which may be involved in mineral deposition [Boyan *et al.*, 1989; Bonucci, 1970; Wu *et al.*, 1991a,b; Shapiro *et al.*, 1966; McLean *et al.*, 1987; Genge *et al.*,

1988; Genge *et al.*, 1990]. It is not clear what the exact role, if any, these vesicles perform in extracellular biomineralization. Transmission electron micrographs of these matrix vesicles have shown the presence of calcium phosphate mineral deposits in the interior of the vesicles [Eanes *et al.*, 1984; McLean *et al.*, 1987; Arsenault, 1985; Bonucci, 1970], and this has been taken to mean that the matrix vesicles themselves may be responsible as nucleators of the initial phase(s) of calcium phosphate minerals.

Borrowing from the in vivo observations of vesicle-mediated calcium phosphate precipitation, several in vitro investigations utilizing phosphatidylserine, phosphatidylcholine, and dicetyl phosphate-containing synthetic liposomes have shown that when the vesicles are first encapsulated with inorganic phosphate and then inoculated with Ca^{2+} ionophores (e.g., X-537A) in the presence of Ca^{2+}, a calcium phosphate precipitate forms in the interior and the exterior surface of the vesicles [Eanes and Hailer, 1987; Eanes and Hailer, 1985; Eanes *et al.*, 1984). This initial calcium phosphate mineral was observed to be an unstable amorphous phase, which converted to HAP over time [Eanes *et al.*, 1984]. This pattern of mineral deposition is analogous to what has been observed in vivo. In a different approach, Mann *et al.* (1986) utilized phosphatidylcholine unilamellar vesicles (300 Å diameter) to study membrane-mediated iron oxide mineral formation. Metal ions (Fe^{3+} or Fe^{2+} chloride salts) were encapsulated in liposomes via sonication at low pH, and the vesicles were then transferred to alkaline media (pH 12). Electron micrographs, absorbance spectroscopy at 420 nm, and X-ray microprobe analysis revealed that intravesiclar Fe (III) solutions crystallized as poorly ordered sphericular geothite (α-FeOOH), whereas Fe (II) solutions produced intravesicular mixed valence magnetite crystallites [$Fe(II)Fe(III)_2O_4$]. One explanation for this differential precipitation may lie in a diffusion-controlled Fe oxidation rate, i.e., the control of OH^- diffusion across the membrane bilayer and hence the rate of Fe oxidation is regulated by the efflux rate of Cl^- ion diffusion [Mann *et al.*, 1983;

Mann and Williams, 1983). This behavior is supported by the observation that the Fe oxidation rate is first-order (Mann *et al.*, 1986). In subsequent work, Mann and co-workers explored the chemistry of lipid headgroup-controlled $CaCO_3$ crystal growth by utilizing Langmuir monolayers of stearic acid (Mann *et al.*, 1988; Mann *et al.*, 1991), octadecylamine (Heywood *et al.*, 1991), cholesterol (Mann *et al.*, 1991; Rajam *et al.*, 1991) and octadecanol (Rajam *et al.*, 1991). In the presence of supersaturated calcium bicarbonate subphases in which the $[Ca^{2+}]$ was varied, monolayers with carboxylate headgroups induce either the oriented nucleation of calcite (at $[Ca^{2+}]$ = 9 mM) with the {110} axis perpendicular to the organic surface, or the formation of oriented vaterite (at $[Ca^{2+}]$ = 4.5 mM) with the {001} axis normal to the monolayer. Surprisingly, the octadecylamine monolayer surface gave rise to vaterite in two morphologies (Type I and Type II) which were independent of the $[Ca^{2+}]$, and apparently represent two specific $CaCO_3$ nucleation orientations on the organic surface. Octadecanol films were found to inhibit crystallization, and cholesterol monolayers gave rise to random non-oriented calcite deposition analogous to control experiments. These experiments demonstrated: (1) inorganic solid phase formation is cation concentration dependent, (2) a high charge density at the monolayer surface could kinetically favor one form of mineral over another through stabilization of disordered clusters of excess charge (i.e., more CO_3^{2-} ion on the octadecylamine surface, thus preferentially stabilizing the more open structure that is characteristic of vaterite, and (3) the electrostatic interactions between crystal nuclei and the organic surface are influenced by structural relationships at the interface.

The in vitro work of Eanes and Mann have elucidated several exciting concepts regarding template macromolecules: namely, that simple amphipathic organic monolayer films can nucleate and orient inorganic solids. However, as of yet there are no two- or three-dimensional structural details regarding these monolayer surfaces. Further investigations will undoubtedly establish the

headgroup orientation and packing of these monolayers and of lipid bilayers which nucleate mineral phases.

(3) Hydrophobic Proteins. It has been noted that a certain hydrophobic mineral matrix protein, *amelogenin*, is found in association with the inorganic calcium phosphate mineral phase hydroxyapatite (HAP) of tooth enamel. Protein sequencing studies have established that amelogenin possesses conserved amino acid sequence domains which are predominantly non-polar (Gibson *et al.*, 1991a,b; Snead *et al.*, 1985; Fincham *et al.*, 1989). Several isoforms of amelogenin exist as a result of either genetic polymorphism or transcriptional alternative splicing mechanisms (Shimokawa *et al.*, 1987; Breitbart *et al.*, 1987; Young *et al.*, 1987). Amelogenin displays a strong binding affinity for HAP (Aoba *et al.*, 1989; Aoba and Moreno, 1991), and this binding affinity has been localized to the highly conserved 5 kD N-terminal domain of the protein (named the **T**yrosine **R**ich **A**melogenin **P**olypeptide, or **TRAP** fragment)(Fincham *et al.*, 1981; Fincham *et al.*, 1983; Fincham *et al.*, 1989). The TRAP fragment, a 45 residue amino acid sequence, contains 6 Tyr residues at sequence positions 12, 17, 26, 34, 37,and 39, 1 Thr residue at position 21, 3 Ser residues at positions 16, 28, and 33, and 3 His residues at positions 6, 9 and 32. It is believed that these sidechains may mediate the binding of TRAP to the HAP surface, perhaps by a combination of hydrogen bonding and ligand coordination interactions. The 10 residue hydrophilic C-terminal domain of the amelogenin protein, TDKTKREEVD, also affects the ability of the protein to bind to HAP (Aoba and Moreno, 1991). No experimental evidence is available on how these hydrophobic proteins exert control over crystal growth, although enamel fluid studies have demonstrated that enzymatic degradation of the 25 kD amelogenin to the 20 kD and 13 kD forms changes the affinity of the protein for HAP, and, its ability to regulate crystal growth (Aoba and Moreno, 1991). Various attempts at defining the secondary and tertiary structure have met with limited success or are inconclusive (Bonar *et al.*, 1965,1984;

Renugopalakrishnan *et al.*, 1986, 1988, 1989; Zheng *et al.*, 1987; Jodaikin *et al.*, 1986). However, various theories put forth by Sikes and Wheeler (1988a,b; 1989) and others (Schlesinger and Hay, 1979; Hay *et al.*, 1979) suggest that a hydrophobic matrix protein could "sit" on the crystal surface and block the diffusion of cations and anions to the forming face of the crystal, thus limiting apposition at that face and allowing propagation along the crystallographic axis representing the non-blocked face(s) of the crystal. It has also been shown that hydrophobic α-amino acids such as leucine, valine, and phenylalanine form ordered clusters on the water surface that act as epitaxial matrices for oriented nucleation of glycine (Weissbuch *et al.*, 1990). Perhaps, in an analogous fashion, the hydrophobic amino acids of TRAP form epitaxial matrices for calcium phosphate crystal growth.

(4) *Acidic Macromolecules or Polyelectrolyte Mineral Matrix Proteins (PMMP)* Perhaps the best-studied and the most complex group of template macromolecules are the so-called *acidic proteins*, hereafter referred to as the *Polyelectrolyte Mineral Matrix Proteins* (abbreviated as PMMP). As a class, these proteins share several unique characteristics:

(1) Significant proportion (≥50%) of amino acid sidechain R groups that possess carboxylate, dicarboxylate, and hydroxyl chemical groups, or, are post-translationally modified with anionic moieties (e.g., PO_4^{-3}, SO_4^{-3}-modified monosaccharides, PO_4^{-3} modified monosaccharides)
(2) Presence of hydrophobic domain(s)
(3) Glycosylation at Ser or Thr sidechains.
(4) Selective binding of certain divalent cations with high affinity.
(5) Selective binding to inorganic mineral phases.
(6) Present in low abundance in mineralized tissues.

(7) Found in association with scaffolding macromolecules.

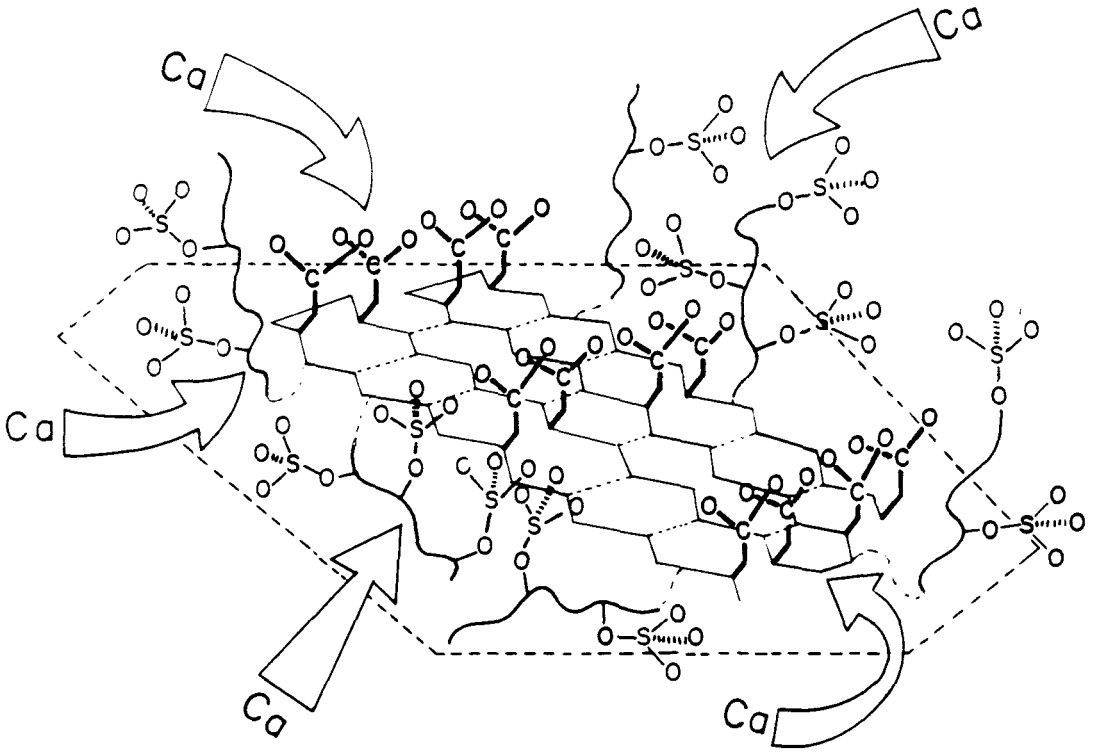
The chemical mechanism by which the PMMPs exert control over biomineralization has not been fully elucidated. From in vitro Ca^{2+} ion binding studies (Crenshaw, 1972; Krampitz *et al.*, 1976; Wheeler and Sikes, 1984) and crystal binding experiments conducted on PMMPs isolated from calcium carbonate systems (Addadi and Weiner, 1985; Addadi *et al.*, 1987), there is clear evidence for PMMP localization at specific crystallographic faces of an inorganic solid, as well as assuming a specific secondary structure (β-sheet) upon binding Ca^{2+}. In their seminal work, Addadi and Weiner (1985) demonstrated that the Asp-rich protein, p12, extracted from the calcite or prismatic layer of the bivalve, *Mytilus californianus*, exhibited stereoselective adsorption or binding onto a specific crystal face of organic and inorganic Ca^{2+} crystals: calcium fumarate trihydrate, calcium malonate dihydrate, calcium tartate tetrahydrate, calcium maleate monohydrate, and calcite. This selective mode of protein binding resulted in the apposition of crystal growth along a particular axis.In calcite, this apposition occurred on the {001} plane along the *c*-axis of the crystal. This face contains the planar carbonate molecules, and it is hypothesized that the Asp and Glu carboxylate sidechains of the p12 protein approach the crystal face in a perpendicular orientation to the calcium plane. In the field of Fe biomineralization, the role of carboxylate amino acid sidechain groups in mineral formation was elegantly demonstrated in the site-directed modification study of Wade *et al.* (1991), in which recombinant human ferritin L and H chains possessed modifications at the ferroxidase site, and the putative nucleation site on the inner surface (Glu61, Glu64, Glu67 → Ala). This resulted in the generation of a ferritin molecule variant which exhibited complete inhibition of Fe uptake and deposition.

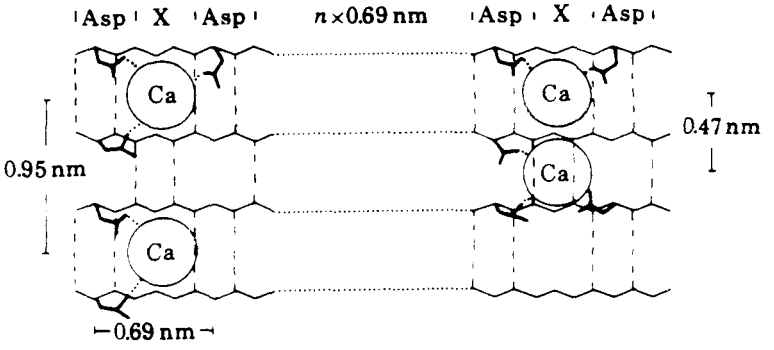
A Model for PMMP Regulation of Biomineralization.

A model of PMMP binding and ordering of cations and crystal surfaces has been invoked by Weiner [1984], Weiner and Traub [1984], and refined by Addadi *et al.*, [1987; see also Lowenstam and Weiner, 1989](Figure 3). The PMMP is viewed as two separate surfaces: sulfonated carbohydrate moieties and polycarboxylate arrays. The sulfonated carbohydrate residues attract divalent cations via electrostatic interactions, thus elevating the local divalent cation concentration in the vicinity of the PMMP protein surface. The carboxylate groups of the protein then bind the metal ions (in this case, Ca^{2+}) in a three-dimensional stereospecific **array**; this array is achieved by arranging the peptide conformation into a β-sheet configuration. The carboxylate groups are arranged in a row perpendicular to the Ca^{2+} atoms, such that the distance between carboxylate groups is 6.6 to 6.9 Å. The β-sheet secondary structure permits the staggered arrangement of Asp and Glu COOH groups, such that Ca^{2+} atom coordination involves intraprotein axial ligand oxygen donors. This intrapeptide coordination stabilizes the Ca^{2+}-protein complex, permitting the formation of a well-ordered three-dimensional COO array on the protein surface, which then mimics the carbonate {001} face of the calcite crystal [known as the *stereochemical effect*]. Further Ca^{2+} ion apposition occurs, and the crystal phase undergoes directional apposition. In a similar manner, the PMMP binds to Ca^{2+} ions localized on the face of a forming crystal. This binding precludes further solution Ca^{2+} apposition from solution, and crystal growth ceases at that face. Since the molecular weights of the known calcite and aragonite PMMPs range between 5 and 20 kD [Weiner, 1983], the adsorbed proteins most certainly span a large number of unit cells when it adsorbs onto a crystal surface. As crystal growth occurs, these adsorbed proteins may become entrapped or occluded by overgrowth of the crystalline phase; this entrapment occurs at boundaries of ordered blocks within the crystals (known as mosaic blocks)[Addadi and Weiner, 1988]. As a result of selectively adsorbing onto specific

Figure 3: **The Proposed Addadi-Weiner-Traub Model of PMMP-Induced $CaCO_3$ Nucleation.**

The **top figure** represents the nucleation site for mollusc shell nucleation of $CaCO_3$. The PMMP protein adopts a β-sheet conformation and adsorbs onto a solid substrate, such as a scaffolding macromolecule (β-silk fibroin, in this case). The covalently bound sulfated polysaccharides attract counterions, which are then coordinated by carboxylate groups of Asp and Glu in the protein. Arrows show the influx of Ca ions. Reproduced from Addadi and Weiner, 1987. The **bottom figure** is a schematic illustration of the possible modes of calcium binding to the Asp rich b-sheet domains in the calcitic PMMP protein. Ligand coordination is hypothesized to be either bidendate or tridendate with respect to carboxylate interactions (i.e., 2 - 3 Asp residues/Ca atom) Reproduced from Weiner and Traub, 1984.


Ca
Ca
Ca
Ca
Ca


Asp
X
Asp
n×0.69 nm
Asp
X
Asp
Ca
Ca
Ca
Ca
0.95 nm
0.47 nm
0.69 nm

calcite crystal planes, the intracrystalline proteins have been shown to either create a more brittle crystalline phase, as in mollusc shell calcite [Addadi and Weiner, 1988], or strengthened, as in sea urchin tests [Addadi and Weiner, 1988; Berman *et al.*, 1988].

Support for the PMMP β-sheet array model has come primarily from FT-IR spectroscopy, CD, and X-ray diffraction. Using HPLC-purified protein fractions isolated from a variety of mollusc shell matrices (*gastropoda*, *bivalva*, and *cephalodoa*), it has been shown that the Asp-rich PMMP proteins undergoes conformational change upon addition of Ca^{2+}, adopting a β-sheet conformation, as evidenced by the frequency of the amide I absorption band (1623-1629 cm^{-1})[Worms and Weiner, 1986]. The poly-L-$(Asp)_n$ sequence can act either as a nucleator, when immobilized to a solid surface, or, as an inhibitor of nucleation when it is free in solution. [Campbell *et al.*, 1989]. These investigators noted that the sodium salt forms of these proteins did not assume a β-sheet conformation, but perhaps an α-helix or random coil conformation. Addadi *et al.*, (1987) employed sulfonated polystyrene films as a surface for nucleation, thus mimicking the organic template of sulfonated polysaccharides that p12 and other prismatic/calcite Asp-rich PMMPs utilize in nucleation chemistry in the bivalve shell. The resulting calcite crystals which formed on these films were found to nucleate off of the {001} face. Under certain conditions of polystyrene sulfonation (8 hr, 24 hr exposure to 98% H_2SO_4), the addition of poly-L-$(Asp)_n$ (average MW: 6 kD), poly-acrylic acid (average MW: 5 kD), and PMMP (p21 protein) resulted in increased formation of {001}-oriented calcite crystals on the surface of the polystyrene film. This result implied that an optimal proportion of sulfonates and carboxylates are necessary for {001} oriented calcite nucleation. Interestingly, FT-IR analysis confirmed that the poly-L-$(Asp)_n$ peptides assumed a β-sheet conformation when bound to the calcite crystals on the sulfonated films. Poly-L-$(Glu)_n$ and Poly-acrylic acid were found not to assume an ordered β-sheet conformation, and concommittantly, did not orient calcite crystals as well as poly-L-$(Asp)_n$ did. In a later

study, Addadi *et al.* (1991) examined the effect of secondary structure on calcium fumarate crystal-polypeptide interactions. In this study, Poly-L-Asp (MW 20 kD) predominantly assumed the β-sheet conformation at pH 7 to 9 in the presence of Ca^{2+}, and, limited the growth of calcium fumarate crystals at 18° and 35° C by binding to the {010} face. Conversely, poly-L-Glu (MW 20 kD) assumes the random-coil conformation in the presence of Ca^{2+}, and delays but does not regulate the growth of calcium fumarate crystals at 18° C, but does so at 35° C. Polymer size influenced the conformation of the peptide: Poly-Asp, MW 6 kD, was discovered to be comprised of 40%:60% mixture of β-sheet:random coil in the presence of Ca^{2+}, whereas the 20 kD polymer is predominantly β-sheet [Addadi *et al.*, 1991]. Circular dichroism studies conducted on $(Glu\text{-}PSer)_{15}$ in the presence and absence of calcium, revealed that a random-coil to β-sheet conformational transition occurred upon the addition of Ca^{2+} [Lee and Veis, 1980]. A preliminary study utilizing atomic force microscopy (AFM) to image poly-L-$(Asp)_n$ peptides bound to $CaCO_3$ mineral surfaces also appears to confirm the specificity of peptide binding on the mineral surface: namely, the peptides bind in parallel linear arrays along the *c*-axes of the flat regions of basal planes [Sikes, 1991].

The concept that anionic proteins can act as templates for the formation of inorganic solids is a fascinating problem that exists at the interface between biology and chemistry, as well as in medical science. Recently, a PMMP protein, uropontin, was identified in human urine samples as an Asp-rich 72 - 50 kD protein that inhibited the growth of calcium oxalate crystals that lead to urinary tract stone disease [Shiraga *et al.*, 1992; Hoyer and Daikhin, 1992]. The application of protein templates as catalytic sites for the formation of novel inorganic materials is already on the horizon. As demonstrated by Mann and Meldrum [1991], and Meldrum *et al.*, [1991], the iron-storage protein ferritin can be adapted to generate nanometer-sized FeS particles by in situ reaction of the iron oxide core of the native ferritin with H_2S/Na_2S. In the same study, other

experiments with redox-active metal ion species such as Ti (III), Co (II), Cr (II), UO_2^{2+}, and Mn (II) demonstrated that the apoferritin protein will generate oxides for Me (II) and the UO_2^{2+} ion. Thus, supramolecular protein "cages" have the potential to act as constrained reaction environments in the synthesis of inorganic materials of nanometer dimension [Meldrum *et al.*, 1991]. Therefore, one of the important goals in the study of template macromolecules is understanding the chemistry of protein-mediated nucleation.

From the foregoing, the PMMP proteins are clearly an intriguing class of proteins which possess numerous "active" sites for molecular recognition of divalent cations or crystal surfaces. Combining many of the chemical properties of the other templates macromolecules, the PMMPs offer an interesting and challenging molecular system for studying the chemical mechanism of nucleation and inhibition.

Bovine Dentine Phosphophoryn (BDPP): A PMMP Protein Involved in the Formation of Biogenic Calcium Phosphates.

Despite the differences in the chemical composition of different biomineral phases, the preponderance of PMMPs in Nature suggests that there may be *a common chemical mechanism(s) for protein template-mediated inorganic solid phase nucleation and controlled epitaxial growth*. The ultimate test of this hypothesis will require focusing experimental efforts on PMMP proteins that are representative of different biogenic mineral systems. Our efforts have been directed at a major PMMP macromolecule, **phosphophoryn,** that shares similar characteristics with other Ca^{2+}-associated PMMP proteins. Phosphophoryn and its role in the biomineralization process has been characterized to a greater extent than any of the other known PMMPs.[3] Historically, this has

[3]A similar argument could be made in support of **osteocalcin**, or vitamin-K-dependent bone GLA protein, as the "best-characterized"

occurred for a very important reason: this template macromolecule is involved in the formation of inorganic calcium phosphate mineral phases in vertebrate tissue (i.e., tooth dentine) and hence possesses biomedical significance. First reported in 1972 as an EDTA-soluble non-collagenous protein isolate of bovine tooth dentin [Veis *et al.*, 1972], this extraordinary calcium binding protein was subsequently purified from other tooth sources, including human, ovine, rat, and sea urchin [Rahima and Veis, 1988; Veis *et al.*, 1986], phosphophoryn is an extraordinary calcium binding protein. As shown in Table III, the amino acid composition is similar to many of the characterized bivalve PMMPs [Weiner and Hood (1975); Weiner (1983); Borbas *et al.* (1991)]; to Ca^{2+} binding salivary proteins such as statherin [Schlesinger *et al.*, 1989] and acidic proline-rich proteins (APRP)[Moreno *et al.*, 1982; Hay *et al.*, 1988]; to the enamel protein, tuftelin [Deutsch *et al.*, 1991]; and to many of the bone phosphoproteins [Lee and Glimcher, 1981; Uchiyama *et al.*, 1986; Gotoh *et al.*, 1990a,b; Mikuni-Takagaki and Glimcher, 1990; Zhang *et al.*, 1990].[4] It has been found that Asp and *o*-phospho-L-serine (PSer) exist in a nearly 1:1 stoichiometry in the amino acid composition. The bovine and murine phosphophoryns exist as a single isoform, whereas in the rat there are at least three isoforms

---

PMMP, since extensive structural and Ca^{2+} binding studies have been performed on this protein [refs]. However, osteocalcin differs significantly from the PMMP criteria utilized in this report. In addition, very few studies have definitively shown what role, if any, osteocalcin plays in the nucleation or regulation of calcium phosphate crystal growth in vivo.

[4]It appears that phosphophoryn also has a distant "cousin" in the American Oyster, *Crassostrea virginica*! Rusenko *et al.*, (1991) have recently purified and characterized a phosphoprotein from the oyster shell matrix. The amino acid composition of this protein possesses 90% Asp, Ser, and Gly in equimolar amounts, with 10% of the remaining amino acid composition being non-polar and hydrophobic, which appear to be localized at the C-terminus of the protein. Mild acid hydrolysis of this phosphoprotein reveal the presence of Poly-L-$(Asp)_n$ and Poly-L-$(Pser)_n$ domains. And, the oyster shell protein was shown to effectively inhibit $CaCO_3$ precipitation in vitro.

that have been characterized thus far [Tsay and Veis, 1985, Rahima and Veis, 1988]. Depending on the tissue and the circumstances of extraction, 0.2% to 15% of the total phosphophoryn protein has been found to associate either covalently [Kuboki *et al.*, 1984] or non-covalently [Maier *et al.*, 1983; Stetler-Stevenson and Veis, 1986] with the scaffolding macromolecule Type I collagen. This protein-protein association may explain the appearance of matching of the periodicity of the initial calcium phosphate crystals with the periodicity of the collagen fibers [Glimcher and Krane, 1968; White *et al.*, 1979]. The phosphophoryn protein itself may contain intramolecular chemical crosslinks such as N^{ε}-(2-amino-2-carboxyethyl)-lysine and N^{τ}-(2-amino-2-carboxyethyl)-histidine [Kuboki *et al.*, 1984]. Phosphophoryn exists as a single polypeptide chain which binds Ca^{2+} at two classes of sites: high and low affinity [Bovine 30 kD protein fragment: $K_{a_{High}} = 3.6 \times 10^{4}$/moles, 2.8×10^{-6} mol/mg Ca^{2+} sites per protein; $K_{a_{Low}} = 5 \times 10^{2}$/moles, (Lee *et al*, 1977); 30 kD Rat protein fragment: $K_d = 9.0 \times 10^{-8}$ M, 127 mole Ca^{2+}/mole protein; $K_d = 1.1 \times 10^{-5}$ M, 176 mole Ca^{2+}/mole protein (Zanetti *et al.*, 1981)]. The 30 kD protein fragment also exhibits varying binding affinities to Group Ia, IIa, *d*-block, and lanthanide series metal ions [Zanetti *et al.*, 1981; Cookson *et al.*, 1980]. The affinity for metal ions goes as: $Ca^{2+} > Mn^{2+} > Mg^{2+} > La^{3+} > K^{+} > Sr^{2+} > Cs^{+} > Cd^{2+} > Na^{+}$ [Zanetti *et al.*, 1981]. The presence of high ionic strength exerts only a slightly diminishing effect on the $K_{a_{Ca}}$ of phosphophoryn, indicating that the protein binds calcium via ligand coordination and not by simple Coulombic attraction [Stetler-Stevenson and Veis, 1987]. Preliminary sequencing studies of the rat and bovine isoforms [Lechner *et al.*, 1981; Veis, 1988; Sabsay *et al.*, 1991] have revealed that Pser and Asp residues exist in sequence blocks [i.e, $(Pser\text{-}Asp)_n$, $(Pser)_n$, $(Asp)_n$, and combinations thereof][Figure 4]. Evans and Chan [1991], in their 23 residue N-terminal sequence of bovine dentine phosphophoryn, observed the presence of Asp residues interdispersed between non-charged amino acids such as Asn, Ser, and Gly (see Chapter Two of this volume). The structural feature of

Table III: Amino Acid Compositions of Some Representative PMMP Proteins.

Adapted from Veis (1988); Weiner (1983); Shiraga *et al.*, 1992; Worms and Weiner (1986); and Stetler-Stevenson and Veis (1983).

Table III

Amino Acid	BDPP[1]	HDPP[2]	Uropontin [3]	CSP II[4]	N12[5]	P12[6]
Lys	44	32	39	3.5	4.6	4.66
His	6	27	36	0.9	0	0
Arg	3	31	17	1.3	3.15	2.89
Asp	401	273	207	31.7	26.71	31.55
Thr	8	36	66	2.9	7.24	8.17
Ser	457	206	162	7.8	7.47	9.32
Glu	8	36	140	17.3	17.7	19.33
Pro	3	52	55	4.4	8.46	4.86
Gly	33	86	69	14.0	8.20	6.68
Ala	9	46	52	5.8	4.29	3.03
1/2 Cys	2	23	0	0	0	0
Val	3	19	46	3.3	4.21	2.57
Met	0	9	7	0.5	0.53	0.16
Ile	2	9	15	1.5	4.17	3.11
Leu	3	34	47	3.2	2.28	1.26
Tyr	2	14	26	0.7	0	0.18
Phe	1	15	16	1.2	1.49	0.19

[1]From Stetler Stevenson and Veis, 1983. Expressed as residues/1000. Ser value represents total o-phosphoserine and serine content. Asp includes Asn content; Glu includes Gln content.

[2]From Veis, 1988. Expressed as residues/1000. Ser value represents total o-phosphoserine and serine content. Asp includes Asn content; Glu includes Gln content.

[3]From Shiraga *et al.*, 1992; human variant. Expressed as residues/1000. Asp includes Asn content; Glu includes Gln content.

[4]From Worms and Weiner, 1986; representing HPLC purified *Acropora humilis* coral skeleton protein (CSP) . Expressed as mole percent. Asp includes Asn content; Glu includes Gln content.

[5]From Weiner, 1983, represents nacreous-aragonite protein. Expressed as mole percent. Asp includes Asn content; Glu includes Gln content. Obtained from the bivalve *mytilus californianus.*

[6] From Weiner, 1983, represents prismatic calcite protein. Expressed as mole percent. Asp includes Asn content; Glu includes Gln content. Obtained from the bivalve *mytilus californianus.*

anionic residue organization into polyelectrolyte clusters has also been observed for other PMMPs (see above). By inference, it would seem likely that the polyelectrolyte cluster regions of Pser, Asp, and Glu in phosphophoryn are potential sites of Ca^{2+} ion binding. This was partially demonstrated by the alkaline phosphatase dephosphorylation of rat and bovine phosphophoryn, which resulted in the abolishment of high affinity Ca^{2+} binding sites (Lee and Veis, 1980; Zanetti *et al.*, 1981). Thus, Pser residues comprise part or all of the high affinity sites, and by analogy, the Asp and Glu residues most likely comprise the low affinity Ca^{2+} binding sites on phosphophoryn.

The Role of Phosphophoryn in Extracellular Biomineralization

There is some ambiguity regarding the <u>in vivo</u> role that the protein plays in biomineralization. Phosphophoryn forms complexes with Ca^{2+} (1.33 ions/Pser residue) and Mg^{2+} (1.07 ions/Pser residue)(Marsh, 1989a), and these cation-protein complexes have been shown to form a ternary complex with the counteranion, PO_4^{-3} (Lee *et al.*, 1983). The stoichiometry of the Ca-protein-PO_4^{-3} ternary complex has been determined to be similar to that of ACP (Marsh, 1989a,b). The Ca^{2+}-protein complex appears amorphous by electron diffraction, and is observed in the EM to be predominantly linear when the calcium binding sites are 65% saturated; above this level, the protein adopts a folded conformation and protein-protein interactions occur (Marsh, 1989a). In contrast, the Mg^{2+}-protein complex remains linear up to 99% Mg^{2+} site saturation (Marsh, 1989a). In the phosphophoryn-Ca^{2+}-PO_4^{-3} complex, approximately 15% of the bound phosphate is buried in protected domains and is stabilized against ion dissociation (Marsh, 1989b). Marsh (1989a,b) has argued that the phosphophoryn protein, by virtue of its ability to sequester or form ACP clusters on the protein surface, exists <u>in vivo</u> in a mineralized aggregated state that may not be able to catalyze the the direct nucleation of a crystalline calcium phosphate phase (HAP). Instead,

Figure 4: Phosphophoryn Sequence Fragments Derived from Peptide Mapping.

Taken from Sabsay *et al.*, 1991, using the terminology for peptide fragments listed therein. X refers to possibility of phosphoserine.

Peptide Fragment	Sequence
I15	...DDDD...
IV31	...DDXDXDDDDDD...
IV32	...DXXXDD...
12/45	...DIVHPSYRSRTLRRDYMLI...
12/45	...DDDDDDYSDSDSSDSDD...
11	...SXXSXXSSS...

the protein acts as a *sequesterase*, in which the protein-inorganic complex provides a localized microenvironment of ion supersaturation, such that degradation of the protein leads to release of the ions and spontaneous inorganic mineral precipitation. In contrast, other investigations have demostrated that phosphophoryn can induce in vitro HAP formation, either as a free protein species in metastable calcium phosphate solutions (Narwot *et al.*, 1976), or, as a species cross-linked to a rigid polymer matrix in the presence of metastable calcium phosphate (Lussi *et al.*, 1988; Linde *et al.*, 1989). At pH 8.0, the mineral synthesized in the presence of 2.4 mg phosphophoryn/100 mL phosphate solution corresponded to 50-70% HAP which exhibited poorly crystalline structure, as evidenced by FT-IR and X-ray diffraction measurements (Narwot *et al.*, 1976). In the absence of phosphophoryn, ACP forms under these conditions at 25° C (Boskey and Posner, 1973). The matrix-bound phosphophoryn, at a concentration of 2 nmoles/mL, produced a more crystalline HAP phase which exhibited plate-like morphology in EM micrograms (Lussi *et al.*, 1988), thus fueling speculation that phosphophoryn induces nucleation only when the protein is bound to a rigid support, such as a scaffolding macromolecule like Type I collagen. Recent studies by Fujisawa and Kuboki (1990), Addadi *et al.* (1992) and Moradian-Oldak *et al.* (1992a,b) reveal that phosphophoryn specifically interacts with the {010} face of OCP and the {100} face of HAP. Separate studies using atomic force microscopy (AFM) reveal that calcite shell phosphoproteins assumed a "globular" conformation when adsorbed to calcite $CaCO_3$ crystals (Donachy *et al.*, 1992).

Conformational Studies of Phosphophoryn

Preliminary CD, FT-IR, ^{31}P and ^{1}H NMR, and viscometry studies have been conducted on phosphophoryn to ascertain the conformation of the protein in the divalent cation depleted state (Cookson *et al.*,1980; Lee *et al.*,1977; Lee *et al.*,1983; Stetler-

Stevenson and Veis, 1987]. ^{1}H NMR (270 MHz) studies revealed that the 30 kD rat incisor phosphophoryn fragment in the absence of divalent cations exhibited narrow linewidths for Pser, Asp, Glu, Ser, and other amino acids. These narrow linewidths were taken as an indication of considerable segmental mobility in the protein [Cookson *et al.*, 1980]. In this same study, local conformational rearrangement of the protein was demonstrated by the disappearance of amide-NH proton resonances in D_2O at pH 6.5 as the pH was varied. The Asp pK_a's, as measured by ^{1}H NMR pH titration, were found to exist in two groups: one at 4.2 ± 0.2 [representing 80% of the total Asp residues in the protein] and the other at 3.2 pH units [Cookson *et al.*, 1980]. Glutamic acid residues in the 30 kD rat phosphophoryn fragment were found to have pK_a's of 4.8 ± 0.2. The apparent pK_a^2's for Pser residues, measured by ^{1}H and ^{31}P NMR, were found to be 6.6 ± 0.2 and 7.2 ± 0.2 pH units, respectively. These Pser pK_a^2's were within experimental error of the values obtained for 94 kD fetal bovine phosphophoryn in 90% H_2O/10% D_2O [Lee *et al.*, 1983] and low molecular weight bovine phosphophoryn fragment [Lee *et al.*, 1977]. The non-Henderson-Hasselbach titration behavior of 80% of the Asp residues, the Pser residues, and the Glu sidechains in phosphophoryn was explained by invoking the existence of a hydrogen-bonding network, comprised of anionic centers in the protein structure which stabilize the binding of H^+ ions [Cookson *et al.*, 1980]. Cookson *et al.* concluded that at neutral pH in the absence of divalent cations, the phosphophoryn protein adopts a fairly extended, flexible solution structure resembling a negatively charged membrane surface. Additional evidence for a random coil conformation was also presented in two other studies, *viz*, (2) the CD studies of Lee *et al.* (1977), where a small phosphophoryn protein fragment exhibited a weak negative dichroic band at 235 nm and a weak positive dichroic band at 218 nm, considered typical of a random coil molecule, and (2) viscosity measurements of bovine dentin phosphophoryn at high ionic strength in the absence of divalent cations, where the protein behaved as an extended random chain molecule [Stetler-

Stevenson and Veis, 1987). In the presence of solvents (EtOH;Lee *et al.*, 1977) and denaturants (guanidine HCl; Stetler-Stevenson and Veis, 1983), phosphophoryn does not exhibit a random coil conformation.

Phosphophoryn undergoes significant conformational change upon the addition of Ca^{2+} and other divalent cations. EPR spectroscopy was applied to examine metal ion binding stoichiometry by measurement of spectral amplitude of the Mn^{2+} signal (Cookson *et al.*, 1980). These studies revealed the existence of two Mn^{2+} binding site classes in the 30 kD rat protein fragment, consistent with the observations of Lee *et al.*, 1977; Jontell *et al.*, 1981; and Stetler-Stevenson and Veis, 1987. ^{1}H NMR difference spectroscopy of 30 kD rat phosphophoryn protein fragment in the presence and absence of paramagnetic Mn^{2+} showed the β-CH_2 resonance of the Pser residues to be most affected at low Mn^{2+}:protein ratios (<1:5), with the β-CH_2 signal of Asp less affected at the same metal:protein ratio (Cookson *et al.*, 1980). The Mn^{2+} was found to be easily displaced by La^{3+}, and Ca^{2+}, and the cation affinities were sensitive to ionic strength and pH. In this same study, Ca^{2+} binding resulted in the inhomogeneous broadening of ^{1}H NMR linewidths, particularly for Pser and Asp sidechains at low metal ion:protein ratios. With subsequent Ca^{2+} addition, Glu, Ala, and Thr residues also exhibited linewidth broadening, with protein precipitation occurring at calcium:protein concentrations exceeding 45:1. Linewidth broadening was utilized as evidence of reduced segmental motion in the protein as a result of Ca^{2+} complexation and conformational folding (Cookson *et al.*, 1980). Inhomogeneous linewidth broadening of the ^{31}P Pser resonances was also observed as a function of Ca^{2+} addition to phosphophoryn (Lee *et al.* (1977; 1983), with the disappearance of the ^{31}P resonance: (1) at a Ca^{2+}:protein ratio of 1:1 in the absence of inorganic phosphate (Lee *et al.*, 1977), and (2) at a Ca^{2+}:P molar ratio of 1:4 in the presence of inorganic phosphate (Lee *et al.*, 1983). Solid-state infrared spectroscopy was employed to show that the carboxylate carbonyl

stretch at 1625 cm^{-1} of divalent-cation depleted phosphophoryn became reduced upon addition of Ca^{2+} (Lee *et al.*, 1977) Simultaneously, the protein amide I band at 1650 cm^{-1} became more resolved in the presence of Ca^{2+}. Combined with the CD observations which demonstrated that Ca^{2+} addition induced a shift of the phosphophoryn negative dichroic band to 227 nm and the disappearance of the positive dichroic band, it was concluded that the phosphophoryn fragment transformed from a random coil to a more ordered conformation in the presence of Ca^{2+}, although it was noted that the ordered conformation may not be exclusively β-sheet (Lee *et al.*, 1977).

Although progress has been achieved in the study of the phosphophoryn protein over the last ten years, there are still a number of problems which exist. In previous investigations of phosphophoryn ion binding and protein conformation, the protein was either not purified to homogeneity (Marsh, 1989a,b), not totally depleted of divalent cations (Marsh, 1989a,b; Lee *et al.*, 1977, 1983), or was in actuality a degradation peptide fragment of the intact isoform molecule (Lee *et al.*, 1977; Cookson *et al.*, 1980). Most if not all of the spectroscopic studies performed on phosphophoryn were preliminary in nature, due to the lack of primary amino acid sequence data and the size of the protein molecule (>30 kD). Thus, a significant proportion of the data that exists on phosphophoryn is incomplete. As of this writing, no complete amino acid sequence is available for any isoform of phosphophoryn. The acid-catalyzed γ-decarboxylation and subsequent β-elimination of the Asp-Asp peptide linkage (Weiner and Hood, 1975; Sabsay *et al.*, 1991) has thwarted the application of either liquid-phase or gas-solid phase acid peptide sequencing methods to determine the primary amino acid sequence of PMMPs via endopeptidase peptide mapping. Coupled with the inability of cDNA sequencing to determine Ser phosphorylation sites, these limitations have imposed severe restrictions on obtaining the complete amino acid sequence for phosphophoryn.

The Concept of Molecular Complementarity in Template-Metal Ion Interactions: Studying the Tertiary Structure of Phosphophoryn

Clearly, there are many details of the phosphophoryn/PMMP story that have yet to be explored. The most interesting and yet the most poorly understood feature is the secondary and tertiary structure of phosphophoryn and the PMMP class of proteins in general, since this obviously defines the template surface and thus the stereochemistry of metal and complementary anion binding sites. The issue of the template surface has been recently raised in a review article on template-controlled crystal growth [Mann, 1988b]. Mann has suggested that an effective biomineralization template most likely possesses molecular complementarity for metal ions in solution and at the growing crystal interface. According to Mann, complementarity for metal ions can be achieved by the template in the following manner:

1) Create a maximal spatial charge distribution on the template surface.

2) Lower the activation energy for ion cluster formation from solution.

3) Generate a specific stereochemical and space symmetry arrangement of ligand atoms.

4) Induce cooperative and synergistic binding.

In the phosphophoryn protein, all three-space and energetic effects mentioned above would be mediated by the protein secondary and tertiary structure, which creates the stereochemical arrays of carboxylate or *o*-monophosphate anionic moieties in solution. If Mann's hypotheses are correct for template macromolecules in

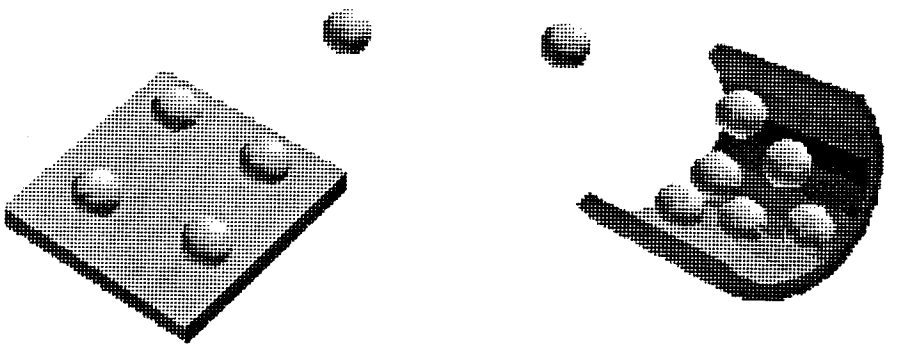
general, then for the protein-mediated nucleation or inhibition processes to be energetically feasible (i.e., have a low $\Delta G_{nucleation}$) and thus be catalytic, we postulate the following (Figure 5): *the template protein (phosphophoryn) must exist in a specific conformational folded state in solution to achieve the highest negative charge density and ligand geometry/symmetry to create active binding sites for the initial step of nucleation or crystal binding*. Furthermore, if the process of ion and crystal binding is cooperative and synergistic in nature, *then once the initial binding step occurs, the phosphophoryn protein must traverse a specific folding pathway to permit the formation of new active sites on the template surface that bind ions with increased affinity* (Figure 5). This process would then continue until the template surface of phosphophoryn reaches the level of saturation binding.

Toward Testing the Concept of Molecular Complementarity in Template-Metal Ion Interactions.

To test this hypothesis, the major goal of this project will be to refine our understanding of how phosphophoryn folds in the presence of Ca^{2+}. To do this, we need to ascertain the initial conformational state(s) that the protein adopts in the absence of Ca^{2+} (denoted henceforth as the *divalent-cation depleted state*), and, to be able to monitor protein conformational change as a function of divalent cation concentration up to the limits of ion saturation on the phosphophoryn template surface (denoted as the *cation-saturated state*). The approach which has been adopted for our investigations is to extend the preliminary global conformational studies of Lee *et al.*, (1977, 1983) and Cookson (1980), and examine phosphophoryn structure via conformational perturbation. Our work could then be refined more precisely at a later time when complete sequence data becomes available. The bovine isoform of phosphophoryn (bovine dentin phosphophoryn, hereafter abbreviated as **BDPP**) was chosen as the focus of these studies for the following reasons:

Figure 5: The Molecular Complementarity Theory of Template Macromolecules.

Taken from Mann (1988b). The molecular template here is represented as a surface of ligand atoms that may or may not possess charges (*right*). By adjusting the topology or contour of the surface (*left*), a higher ligand atom density is achieved. This results in shorter ligand atom-ligand atom distances per surface area, similar to that which may exist in a inorganic solid lattice. Cations, represented by spheres, then deposit in these high density domains to form the first metal atom array of the inorganic solid.



(1) This particular phosphophoryn isoform protein has thus far been studied in greater detail with regard to Ca^{2+} binding, partial amino acid sequence, amino acid composition, amino acid sequence, molecular weight, and solution structure [Veis, 1982, 1988; Sabsay *et al.*, 1991; Lee *et al.*, 1977, 1983; Rahima and Veis, 1988; Stetler-Stevenson and Veis, 1983, 1986, 1987].

(2) It exists as a single isoform in the mineral matrix and can be isolated to homogeneity in an intact form [Stetler-Stevenson and Veis, 1983],

(3) The tissue source can be acquired in sufficient quantities (i.e., milligrams) to permit the isolation of the protein in yields that are feasible for use in spectroscopy studies.

Using this particular PMMP protein, we wish to address the following issues regarding BDPP and PMMP macromolecules in general; *viz*,

1: Determine ligand group arrangement in divalent cation-depleted BDPP by peptide sequencing and multinuclear NMR spectroscopy.

2: What effect does pH, temperature, and solvent denaturation have on the conformation and folding of the Ca^{2+} binding regions in the cation depleted BDPP protein?

3: What solution conformation(s) does BDPP assume in the absence of Ca^{2+}, and what folding pathway(s) are utilized by the protein in the presence of Ca^{2+} to arrange the ligand donor groups three-dimensionally?

4: Based upon BDPP sequencing data, derive a simple peptide model that incorporates a Ca^{2+} binding motif, and determine the preferred secondary and tertiary structure of the peptide under conditions of counterion condensation by theoretical Monte-Carlo simulation.

Although most contemporary NMR structural determinations rely on sequence-specific assignments via homonuclear and heteronuclear shift correlation and cross-relaxation experiments, in our system, we are limited by the absence of complete primary amino acid sequence data for BDPP. Moreover, the inherent limitations of homogeneous and inhomogeneous broadening, restricted motion, spin relaxation, and spin diffusion that are extant with a large macromolecule like BDPP (155 kD, 1130 amino acid residues) restrict our NMR experiments to the identification of overlapping amino acid spin systems, and monitoring ensemble nuclear spin properties as a function of environmental perturbation to the BDPP molecule. In order to improve on these limitations, conformational minima search algorithms and force field computational studies were undertaken to define the secondary and tertiary folding of Ca^{2+} binding polyelectrolyte peptide sequences. These experiments will be discussed in detail in forthcoming chapters of this report.

REFERENCES

Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* **82** 4110-4114.

Addadi, L., and Weiner, S. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York pp. 133-152.

Addadi, L., Moradian, J., Shay, E., Maroudas, N.G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* **84** 2732-2736.

Addadi, L., Moradian-Oldak, J., and Weiner, S. (1991) *ACS Symposium: Surface Reactive Peptides and Polymers* American Chemical Society pp. 13-27.

Addadi, L., Moradian-Oldak, J., Furedi-Milhofer, H., and Weiner, S. (1992) *Conn. Tiss. Res.* **27** (2-3) 93.

Aoba, T., and Moreno, E.C. (1991) *ACS Symposium on Surface Reactive Polymer and Peptides* Chapter 7, pp. 85-106.

Aoba, T., Moreno, E.C., Kresak, M., and Tanabe, T. (1989) *J. Dent. Res.* **68** 1331-1336.

Arsenault, L.S. (1985) in *The Chemistry and Biology of Mineralized Tissues* (Butler, W.T., Ed.) Ebsco Media, Inc., Birmingham, Alabama pp. 364-367.

Berman, A., Addadi, L., and Weiner, S. (1988) *Nature* **331** 546-548.

Berman, A., Addadi, L., Kvick, A., Leiserowitz, L., Nelson, M., and Weiner, S. (1990) *Science* **250** 664-667.

Blakemore, R.P. (1975) *Science* **190** 377-379.

Blakemore, R.P., Frankel, R.B., and Kalmijn, A.J. (1981) *Nature* **286** 384-385.

Blakemore, R.P., Short, K.A., Bazylinski, D.A., Rosenblatt, C., and Frankel, R.B. (1985) *Geomicrobiology Journal* **4** 53-71.

Blumenthal, N.C., Posner, A.S., Silverman, L.D., and Rosenberg, L.C. (1979) *Calcif. Tiss. Int.* **27** 75-82.

Bonar, L.C., Mechanic, G.L., and Glimcher, M.J. (1965) *J. Ultra. Res.* **13** 296-307.

Bonar, L.C., Grynpas, M.D., and Glimcher, M.J. (1984) *J. Ultra Res.* **86** 93-99.

Bonucci, E. (1970) *Z. Zelleforsch Mikrosk Anat.* **103** 192-217.

Borbas, J.E., Wheeler, A.P., and Sikes, C.S. (1991) *J. Expt. Zool.* **258** 1-13.

Boskey, A.L., and Posner, A.S. (1973) *J. Phys. Chem.* **77** 2313-2316.

Bourdon, M.A., Krusius, T., Campbell, S., Schwartz, N.B., and Ruoslahti, E. (1987) *Proc. Natl., Acad. Sci USA* **84** 3194-3198.

Boyan, B.D., Schwartz, Z., Swain, L.D., and Khare, A. (1989) *Anat. Rev.* **224** 211-219.

Breitbart, R.E., Andreadis, A., and Nadal-Ginard, B. (1987) *Ann. Rev. Biochem.* **56** 467-495.

Brown, W.E. (1966) *Clin. Ortho. Rel. Res.* **44** 205-220.

Chen, C-C., and Boskey, A.L. (1985) *Calcif. Tiss. Int.* **37** 395-400.

Cookson, D.J., Levine, B.A., Williams, R.J.P., Jontell, M., Linde, A., and de Bernard, B. (1980) *Eur. J. Biochem.* **110** 273-278.

Crenshaw, M.A. (1972) *Biomineralization Res. Report* **6** 6-11.

Cuervo, L.A., Pita, J.C., Marquez, J.F., and Gatter, R.A. (1969) *J. Clin Invest.* **48** 630-641.

Deutsch, D., Palmon, A., Fisher, L.W., Kolodny, N., Termine, J.D., and Young, M.F. (1991) *J. Biol. Chem.* **266** (24) 16021-16028.

de Jong, E.W., Bosch, L., and Westbroek, P. (1976) *Eur. J. Biochem* **70** 611-621.

Donachy, J.E., Halloran, B.A., and Sikes, C.S. (1992) *Connect. Tiss. Res.* **27** (2-3) 110.

Dziewiatkowski, D.D. (1987) *Calcif. Tiss. Int.* **40** 265-269.

Dziewiatkowski, D.D., and Majznerski, L.L. (1985) *Calcif. Tiss. Int.* **37** 560-564.

Eanes, E.D., and Hailer, A.W. (1985) *Calcif. Tiss. Int.* **37** 390-394.

Eanes, E.D., and Hailer, A.W. (1987) *Calcif. Tiss. Int.* **40** 43-48.

Eanes, E.D., Hailer, A.W., and Costa, J.L. (1984) *Calcif. Tiss. Int.* **36** 421-430.

Elliot, J.C. (1973) *Clin. Ortho. Rel. Res.* **93** 313-345.

Elliot, S.R. (1991) *Nature* **354** 445-452.

Embery, G., and Rolla, G. (1980) *Acta Odentol. Scan.* **38** 105-108.

Evans, J. S., and Chan, S.I. (1991) in *Proteins: Structure, Dynamics, and Design* (Renugopalakrishnan, V., Carey, P.R., Smith, I.C.P., Huang, S.G., and Storer, A.C. Eds) ESCOM Press, Leiden, Netherlands, pp. 251-260.

Fincham, A.G., Belcourt, A.B., Termine, J.D., Butler, W.T., and Cothran, W.C. (1981) *Biosci. Repts.* **1** 771-778.

Fincham, A.G., Belcourt, A.B., Termine, J.D., Butler, W.T., and Cothran, W.C. (1983) *Biochem J.* **211** 149-154.

Fincham, A.G., Hu, Y., Pavlova, Z., Slavkin, H.C., and Snead, M.L. (1989) *Calcif. Tiss. Int.* **45** 243-250.

Francis, M.D., and Webb, N.C. (1971) *Calcif. Tiss. Res.* **6** 335-342.

Fujisawa, R., and Kuboki, Y. (1990) in *Mechanisms and phylogeny of mineralization in biological systems*. (eds S. Suga and Nakahara, H.) Tokyo: Springer-Verlag, pp. 107-110.

Genge, B.R., Sauer, G.R., Wu, L.N.Y., McLean, F.M., and Wuthier, R.E. (1988) *J. Biol. Chem.* **263** (34) 18513-18519.

Genge, B.R., Wu, L.N.Y., and Wuthier, R.E. (1990) *J. Biol. Chem.* **265** (8) 4703-4710.

George, A., Simonian, P., Tylzanowski, P., and Veis, A. (1992) *Connect. Tiss. Res.* **27** (2-3) 107.

Gibson, C., Golub, E., Ding, W., Shimokawa, H., Young, M., Termine, J., and Rosenbloom, J. (1991a) *Biochem. Biophys. Res. Commun.* **174** (3) 1306-1312.

Gibson, C., Golub, E., Herold, R., Risser, M., Ding, W., Shimokawa, H., Young, M., Termine, J., and Rosenbloom, J. (1991b) *Biochemistry* **30** 1075-1079.

Glimcher, M.J., and Krane, S.M. (1968) *in Treatise on Collagen, Volume IIB* (Ramachandran, G.N., and Gould, B.S., Eds.) Academic Press, New York, pp. 68-251.

Gotoh, Y., Pierschbacher, M.D., Grzesiak, J.J., Gerstenfeld, L., and Glimcher, M.J. (1990a) *Biochem. Biophys. Res. Comm.* **173** (1) 471-479

Gotoh, Y., Gerstenfeld, L.C., and Glimcher, M.J. (1990b) *Eur. J. Biochem.* **187** 49-58.

Iijima, M., Kamemizu, H., Wakamatsu, N., Goto, T., Doi, Y., and Moriwaki, Y. (1991) *J. Crystal Growth* **112** 467-473.

Hardingham, T., and Muir, H. (1973) *Biochem. Biophys. Acta* **279** 401-405.

Hascall, V.C., and Heinegard, D. (1974) *J. Biol. Chem.* **249** 4232-4241.

Hassel, J.R., Kimura, J.H., and Hascall, V.C. (1986) *Ann. Rev. Biochem.* **55** 539-567.

Hay, D.I., Bennick, A., Schlesinger, D.H., Minaguchi, K., Madapallimattam, G., and Schluckebier, S.K. (1988) *Biochem. J.* **255** 15-21.

Hay, D.I., Moreno, E.C., and Schlesinger, D.H. (1979) *Inorganic Perspect. Biol. Med.* **2** 271-285.

Heughhebaert, J.C., Zawacki, S.J., and Nancollas, G.H. (1990a) *J. Colloid Interface Sci* **135** (1) 20-32.

Heughhebaert, J.C., Zawacki, S.J., and Nancollas, G.H. (1990b) *J. Colloid Interface Sci* **135** (1) 33-44.

Heywood, B.R., Rajam, S., and Mann, S. (1991) *J. Chem. Soc. Faraday Trans.* **87** 735-743.

Howell, D.S., and Pita, J.C. (1976) *Clin Orth. Rel. Res.* **118** 208-229.

Howell, D.S., Pita, J.C., Marquez, J.F., and Gatter, R.A. (1969) *J. Clin Invest.* **48** 630-641.

Hoyer, J.R., and Daikhin, E. (1992) *Connect. Tiss. Res.* **27** (2-3) 188.

Hunter, G.K. (1987) *Conn. Tiss. Res.* **16** 111-120.

Hunter, G.K., Grynpas, M.D., Cheng, P-T, and Pritzker, K.P.H. (1987) *Calcif. Tiss. Int.* **41** 164-170.

Jodaikin, A., Weiner, S., Perl-Treves, D., Traub, W., and Termine, J.D. (1986) *Arch. Oral Biol.* **31** (10) 685-689.

Krampitz, G., Drolshagen, H., Hausle, J., and Irmscher, E. (1976) in *Biomineralization and Biological Metal Accumulation* (Eds. P. Westbroek and E.W. de Jong) Reidel, Dordrecht, pp. 231-247.

Krueger, R.C., Fields, T.A., Hildreth, J., and Schwartz, N.B. (1990) *J. Biol. Chem.* **265** (20) 12075-12087.

Kuboki, Y., Fujisawa, R., Tsuzaki, M., Liu, C.F., and Sasaki, S. (1984) *Calcif. Tiss. Int.* **36** 126-128.

Lechner, J.H., Veis, A., and Sabsay, B. (1981) in *The Chemistry and Biology of Mineralized Connective Tissues.* (A. Veis, Ed.) Elsevier North Holland, N.Y. pp. 395-398.

Lee, S.L., and Glimcher, M.J. (1981) *Calcif. Tiss. Int.* **33** 385-394.

Lee, S.L., Glonek, T., and Glimcher, M.J. (1983) *Calcif. Tiss. Int.* **35** 815-818.

Lee, S.L., and Veis, A. (1980) *Int. J. Pept. Prot. Res.* **16** 231-240.

Lee, S.L., Veis, A., and Glonek, T. (1977) *Biochemistry* **16** (13) 2971-2979.

Lee, S.L., and Veis, A. (1980) *Int. J. Pept. Protein Res.* **16** 231-240.

LeGeros, R.Z., Daculsi, G., Orly, I., Abergas, T., and Torres, W. (1989) *Scanning Microscopy* **3** (1) 129-138.

Lerner, E., Azoury, R., and Sarig, S. (1989) *J. Crystal Growth* **97** 725-730.

Lerner, L., and Torchia, D.A. (1986) *J. Biol. Chem.* **261** 12706-12710.

Linde, A., Lussi, A., and Crenshaw, M.A. (1989) *Calcif. Tiss. Int.* **44** 286-295.

Lohmander, L.S., DeLuca, S., Nilsson, B., Hascall, V.C., Caputo, C.B., Kimura, J.H., and Heinegard, D. (1980) *J. Biol. Chem.* **255** 6084-6091.

Lowenstam, H.A., and Weiner, S. (1989) in *On Biomineralization* Oxford University Press, London, pp. 1-240.

Lowenstam, H.A. (1981) *Science* **211** 1126-1131.

Lowenstam, H.A., and Rossman, G.R. (1975) *Chemical Geology* **15** 15-51.

Lussi, A., Crenshaw, M.A., and Linde, A. (1988) *J. Dent. Res.* **67** 180.

Macaskie, L.E., Empson, R.M., Cheetham, A.K., Grey, C.P., and Skarnulis, A.J. (1992) *Science* **257** 782-784.

Maier, G.D., Lechner, J.H., and Veis, A. (1983) *J. Biol. Chem.* **258** 1450-1455.

Mann, S. (1988a) *Nature* **332** 119-124.

Mann, S. (1988b) *in Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 35-62.

Mann, S., Hannington, J.P., and Williams, R.J.P. (1986) *Nature* **324** 565-567.

Mann, S., Heywood, B.R., Rajam, S., and Birchall, J.D. (1988) *Nature* **334** 692-695.

Mann, S., Heywood, B.R., Rajam, S., and Walker, J.B.A. (1991) in *ACS Symposium : Surface Reactive Peptides and Polymers.* pp 28-41.

Mann, S., Heywood, B.R., Rajam, S., and Walker, J.B.A. (1991) *J. Phys. D: Appl. Phys.* **24** 154-164

Mann, S., Kime, M.J., Ratcliffe, R.G.B., and Williams, R.J.P. (1983) *JCS Dalton Trans.* 771-774.

Mann, S., and Meldrum, F.C. (1991) *Adv. Materials* **3** 316-318.

Mann, S., and Williams, R.J.P. (1983) *JCS Dalton Trans.* 311-316.

Maroudas, A. (1979) in *Adult Articular Cartilage* (Freeman, M., Ed.) Pitman Medical, Kent, England pp. 215-290.

Marsh, M.E. (1989a) *Biochemistry* **28** 339-345.

Marsh, M.E. (1989b) *Biochemistry* **28** 346-352.

McLean, F.M., Keller, P.J., Genge, B.R., Walters, S.A., and Wuthier, R.E. (1987) *J. Biol. Chem.* **262** (22) 10481-10488.

Meldrum, F.C., Wade, V.J., Nimmo, D.L., Heywood, B.R., and Mann, S. (1991) *Nature* **349** 684-687.

Menanteau, J., Mitre, D., and Daculsi, G. (1984) *Calcif. Tiss. Int.* **36** 677-681.

Meyer, J.L., and Eanes, E.D. (1978) *Calcif. Tiss. Int.* **25** 59-68.

Mikuni-Takagaki,Y., and Glimcher, M.J. (1990) *Biochem. J.* **268**

Misra, D.N. (1986) *Calcif. Tiss. Int.* **38** 333-338.

Mooney, R.W., and Aia, M.A. (1961) *Chem. Rev.* **61** 433-448.

Moradian-Oldak, J., Berman, A., Weiner, S., and Addadi, L.(1991) *J. Inorg. Biochem.* **25** 667.

Moradian-Oldak, J., Frolow, F., Addadi, L., and Weiner, S. (1992a) *Proc. Royal Soc. Lond.* B 247 47-55.

Moradian-Oldak, J., Furedi-Milhofer, H., Veis, A., Addadi, L., and Weiner, S. (1992b) *Connect. Tiss. Res.* **27** (2-3) 189.

Moreno, E.C., Kresak, M., and Hay, D.I. (1982) *J. Biol. Chem.* **257** (6) 2981-2989.

Muir, H. (1980) *in The Joints and Synovial Fluid Volume 2.* (Sokoloff, L., Ed.) Academic Press, New York pp. 27-94.

Nancollas, G.H. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 157-182.

Narwot, C.F., Campbell, D.J., Schroeder, J.K., and Van Valkenburg, M. (1976) *Biochemistry* **15** (16) 3445-3449.

Nelson, D.G.A., and McLean, J.D. (1984) *Calcif. Tiss. Int.* **36** 219-232.

Nilsson, B., DeLuca, S., Lohmander, S., and Hascall, V.C. (1982) *J. Biol. Chem.* **257** 10920-10927.

Poole, A.R., Matsui, Y., and Lee, E.R. (1989) *Anat. Record* **224** 167-179.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* **42** 104-112.

Rajam, S., Heywood, B.R., Walker, J.B.A., Mann, S., Davey, R.J., and Birchall, J.D. (1991) *J. Chem. Soc. Faraday Trans.* **87** (5) 727-734.

Renugopalakrishnan, V., Strawich, E.S., Horowitz, P.M., and Glimcher, M.J. (1986) *Biochemistry* **25** 4879-4891.

Renugopalakrishnan, V., Pattabiraman, N., Langridge, R., Strawich, E., Glimcher, M.J., Horowitz, P.M., Zheng, S., Tu, A.T., Huang, S-G., Prabhakaran, M., Duzgunes, N., and Rapaka, R.S. (1988) in *Recent Progress in Chemistry and Biology of Centrally Acting Peptides* (eds. B.N. Dhawan and R.S. Rapaka) Central Drug Institute, Lucknow, India pp. 87-107.

Renugopalakrishnan, V., Pattabiraman, N., Prabhakaran, M., Strawich, E., and Glimcher, M.J. (1989) *Biopolymers* **28** 297-303.

Rusenko, K., Donachy, J.E., and Wheeler, A.P. (1991) *in Surface Reactive Peptides and Polymers: ACS Symposia* (American Chemical Society) pp. 107-124.

Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) *Biochem. J.* **276** 699-707.

Sajdera, S.W., and Hascall, V.C. (1969) *J. Biol. Chem.* **244** 77-87.

Schlesinger, D.H., and Hay, D.I. (1979) *in Peptides: Structure and Biological Function* (Gross, E., and Meienhoffer, J., Eds.) Pierce Chemical Company, Rockford, Ill. pp. 133-136.

Schlesinger, D.H., Hay, D.I., and Levine, M.J. (1989) *Int. J. Pept. Protein Res.* **34** 374-380.

Schwartz, N.B., Habib, G.M., Campbell, S.C., D'Elvlyn, D., Gartner, M., Krueger, R., Olsen, C., and Philipson, L. (1985) *Federation Proc.* **44** 369-372.

Shapiro, I.M., Wuthier, R.E., and Irving, J.T. (1966) *Arch. Oral. Biol.* **11** 501-512.

Shimokawa, H., Sobel, M.E., Sasaki, M., and Termine, J.D. (1987) *J. Biol. Chem.* **262** 4042-4047.

Shiraga, H., Min, W., Van Dusen, W.J., Clayman, M.D., Miner, D., Terrell, C.H., Sherbotie, J.R., Foreman, J.W., Przysiecki, C., Neilson, E.G., and Hoyer, J.R. (1992) *Proc. Natl. Acad. Sci. USA* **89** 426-430.

Sikes, C.S. (1991) *in ACS Symposium on Surface Reactive Polymer and Peptides* Abstract, pg. 95.

Sikes, C.S., and Wheeler, A.P. (1988a) in *Chemical Aspects of Regulation of Mineralization* (Eds. C.S. Sikes, A.P. Wheeler) University of South Alabama Publication Services, pp.15-20.

Sikes, C.S., and Wheeler, A.P. (1988b) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 95-131.

Skiles, D.D. (1985) in *Magnetite Biomineralization and Magnetoreception in Organisms* (Eds J.L. Kirschvink, D.S. Jones, and B.J. MacFadden) Plenum Press, New York, New York, pp. 43-98.

Snead, M.L., Lau, E.C., Zeichner-David, M., Fincham, A.G., Woo, S.L.C., and Slavkin, H.C. (1985) *Biochem. Biophys. Res. Commun.* **129** (3) 812-818.

Stetler-Stevenson, W.G., and Veis, A. (1983) *Biochemistry* **22** 4326-4335.

Stetler-Stevenson, W.G., and Veis, A. (1986) *Calcif. Tiss. Int.* **38** 135-141.

Stetler-Stevenson, W.G., and Veis, A. (1987) *Calcif. Tiss. Int.* **40** 97-102.

Tomazic, B.B., Tung, M.S., Gregory, T.M., and Brown, W.E. (1989) *Scanning Microscopy* **3** (1) 119-127.

Tsay, T-G., and Veis, A. (1985) *Biochemistry* **24** 6363-6369.

Uchiyama, A., Suzuki, M., Lefteriou, B., and Glimcher, M.J. (1986) *Biochemistry* **25** 7572-7583.

Veis, A. (1982) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Publishers, New York, pp.189-223.

Veis, A. (1988) in *Cell and Molecular Biology of Hard Tissues* (Evered, D., and Harnett, S., Eds.) CIBA Symposium 136, John Wiley and Sons, Chichester, U.K., pp. 161-177.

Veis, D.J., Albinger, T.M., Clohisy, J., Rahima, M., Sabsay, B., and Veis, A. (1986)*J. Expt. Zool.* **240** 35-46.

Veis, A., Spector, A.R., and Zamoscianyk, H. (1972) *Biochem. Biophys. Acta* **257** 404-413.

Wade, V.J., Levi, S., Arosio, P., Treffry, A., Harrison, P.M., and Mann, S. (1991) *J. Mol. Biol.* **221** 1443-1452.

Weiner, S. (1983) *Biochemistry* **22** 4139-4145.

Weiner, S., and Hood, L. (1975) *Science* **190** 987-989.

Weiner, S., and Traub, W. (1984) *Phil. Trans. R. Soc. London Ser. B* **304** 421-438.

Weissbuch, I., Addadi, L., Lahav, M., and Leiserowitz, L. (1991) *Nature* **253** 637-645.

Weissbuch, I., Frolow, F., Addadi, L., Lahav, M., and Leiserowitz, L. (1990) *J. Am. Chem. Soc.* **112** 7718-7724.

Wheeler, A.P., George, J.W., and Evans, C.A. (1981) *Science* **212** 1397-1398.

Wheeler, A.P., and Sikes, C.S. (1984) *Am Zool.* **24** 933-944.

White, S.W., Hulmes, D.J.S., Miller, A., and Timmins, P.A. (1979) *Nature* **266** 241.

Wilbur, K.M., and Simkiss, K. (1989) *Biomineralization* (Academic Press, New York, New York) pp. 3-325.

Williams, R.J.P. (1988) in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P.Eds.) VCH Publishers, New York, New York pp. 1-34.

Williams, R.J.P. (1984) *Phil. Trans. R. Soc. Lond.* **B304** 411-424.

Worms, D., and Weiner, S. (1986) *J. Exptl. Zool.* **237** 11-20.

Wu, L.N.Y., Genge, B.R., and Wuthier, R.E. (1991a) *J. Biol. Chem.* **266** (2) 1187-1194

Wu, L.N.Y., Genge, B.R., Lloyd, G.C., and Wuthier, R.E. (1991b) *J. Biol. Chem.* **266** (2) 1195-1203.

Yorke, E.D. (1985) in *Magnetite Biomineralization and Magnetoreception in Organisms* (Eds J.L. Kirschvink, D.S. Jones, and B.J. MacFadden) Plenum Press, New York, New York, pp. 233-41.

Young, M.F., Shimokawa, H.S., Sobel, M.E., and Termine, J.D. (1987) *Adv. Dental Res.* **1** 289-292.

Young, R.A. (1975) *Clin. Ortho. Rel. Res.* **113** 249-262.

Young, R.A., and Brown, W.E. (1982) in *Biological Mineralization and Demineralization* (Nancollas, G.H., ed) Springer-Verlag, Berlin pp. 101-112.

Zanetti, M., de Bernard, B., Jontell, M., and Linde, A. (1981) *Eur. J. Biochem.* **113** 541-545.

Zhang, Q., Domenicucci, C., Goldberg, H.A., Wrana, J.L., and Sodek, J. (1990) *J. Biol. Chem.* **265** (13) 7583-7589.

Zheng, S., Tu, A.T., Renugopalakrishnan, V., Strawich, E., and Glimcher, M.J. (1987) *Biopolymers* **26** 1809-1813.

Zierenberg, R.A., and Schiffman, P., (1990) *Nature* **348** 155-157.

Two

The N-Terminal Fragment of Bovine Phosphophoryn, an Extracellular Mineral Matrix Protein, Shares Sequence Homology with Viral, Bacterial, and Eukaryotic Transcriptional and Post-translational Regulatory Proteins.

(Note: This chapter has been published in *Proteins: Structure Dynamics and Design.* (Renugopalakrishnan, V., Carey, P.R., Smith, I.C.P., Huang, S.G., and Storer, A.C. Eds) ESCOM Press, Leiden, Netherlands. pp.251-260.

Abstract

We report here the N-terminal amino acid sequence for the first 23 residues of purified Bovine Dentine Phosphophoryn (BDPP), a polyelectrolyte mineral matrix protein (PMMP), as determined by Edman liquid/gas-phase automated sequencing:

S D P N S X D E D N G D A D A N D S D X N S D...

Uncertainties (denoted as X) exist at position 6 and 20 which are not resolvable by repeated sequencing runs. These positions may either be occupied by Asp or Ser. The 23 residue terminus possesses short stretches of Asp, Glu or Ser flanked by non-anionic amino acids Gly, Ala, Val, or Asn. This structural motif has been revealed by earlier $^{1}H/^{31}P$ NMR observations on BDPP in solution (Evans and Chan, 1991a,b).
A database search of existing protein sequences (NBRF, EMBL, GENBANK, NEW ENZYME LIST) has uncovered that the 23 amino acid BDPP sequence shares homologies to peptide fragments obtained from (a) Herpes Simplex Virus Type I unique transcriptional activator Vmw175 protein (52% homology), (b) DNA-directed RNA polymerase, σ subunit (*Samonella typhimurium*)(43% homology), and (c) the secretory gene product (*sec 7*) of *Saccharomyces cerevisiae* (43% homology). Homology of 39% is also observed between BDPP and cauliflower mosaic virus 55 kD coat protein. Chou-Fasman and Kyte-Doolittle prediction analyses of the homologous sequences indicate that the structural motifs of these domains are similar. The domains are primarily hydrophilic and may contain stretches of β-sheet, β-turn, or coil structure, but very little α-helix structure. These results are discussed in light of what is currently known about the evolution of template-mediated biomineral development in organisms.

Introduction

The formation of organism-mediated biominerals involving Ca^{2+}, Mg^{2+}, Fe (III), and other metal ions has presented an intriguing problem with regard to the underlying mechanism. In some systems, notably Mollusca,[1,2] vertebrate bone matrix,[3,4] and dentine tooth matrix,[5,6] a class of polyelectrolyte mineral matrix proteins (PMMP) which possess significant (>50%) amounts of anionic amino acids such as Asp, Glu, and o-phosphoserine (Pser), have been isolated and characterized. These proteins have been shown to either associate with a calcium mineral phase or mediate its formation from solution.[1-6] It is not known whether these proteins catalyze and/or inhibit the epitaxial growth of calcium mineral phases. It is believed that the anionic domains of the PMMP's act to chelate and coordinate divalent cations in solution and permit additional counter-anions to stabilize the PMMP-cation complex. The interest in these proteins not only centers on their role in mineral formation, but is also directed at their potential to participate in the formation of novel biocomposite materials, including both inorganic and organic polymers.

Of particular interest to the biologist is the evolution of the PMMP's as a class of extracellular proteins. The ability to form a mineral phase may have arisen as a result of evolutionary pressure for organism adaptation. It has been speculated that mineral phase formation may have evolved as a metal ion "reservoir"; this reservoir could then be tapped as the demand increased within the cell for metal ions critical to

metabolic, catalytic, and other cellular functions.[7,8] Similarly, the mineral phase may have also conveyed a structural improvement in the organism, permitting a cellular scaffolding to form which is thermoresistant as well as resistant to compressive and tensile forces. Such evolutionary mechanical adaptation would permit survival under compressive pressures (e.g., at ocean depths), temperature fluctuations, and allow for the support of multicellular structures capable of locomotion The evolution of the PMMP cation-binding protein with multiple ligand binding sites is a unique phenomenon. The ability of these PMMP's to mediate mineral phase formation suggests that there may be a common chemical mechanism to PMMP-mineral complexation.[9,10]

We have undertaken a study of one such PMMP, phosphophoryn, which is isolated from the dentine matrix of teeth and has been shown to be crucial to the formation of the inorganic calcium phosphate phase of this matrix. [5,6,11-13] The bovine isoform of phosphophoryn (abbreviated as BDPP; 155 kD) is comprised of 40% Asp and 40% Pser, with the remaining amino acids distributed as Gly, non-phosphorylated Ser, Glu, Ala, Val, Leu, Ile, His, and >1% of Phe and Tyr. [5] The protein possesses two types of Ca^{2+} binding sites (high and low affinity) per protein molecule,[14] and has been shown to precipitate Ca^{2+} and PO_4^{-3} from solution to form amorphous calcium phosphate,[12] a precursor of the apatitic calcium phosphate phase found in dentine matrix.

N-Terminal Sequence Analysis of Bovine Phosphophoryn.

In the course of our work, we have performed a N-terminal sequence analysis of bovine phosphophoryn (Figure 1), purified as described elsewhere,[15] using the automated Edman gas-liquid phase sequencing technique. This method resulted in a 23 amino acid sequence with uncertainties at only two positions: 6 and 20. We were unsuccessful in resolving these ambiguities by repeated sequencing runs: thus the X position is either Asp or Ser. Since the o-phosphoserine residue is unstable under conditions of acid hydrolysis, the sites of phosphorylation cannot be determined in this analysis. After 23 cycles on the automated Edman sequencer, equimolar amounts of Asp and Ser were released and determination of the actual sequence beyond position 23 was not possible. A preliminary report on the N-terminal sequence of rat incisor phosphoprotein[16] revealed a tetramer of either Asp-Asp-Asp-Asn or Asp-Asp-Pro-Asn. In either instance, Asp is assigned to the first position, whereas the Ser is assigned to this position in the bovine sequence. This discrepancy in the sequence may be due to species divergence in sequence, and the presence of other phosphoproteins in rat dentine.[6,15]

The 23 residue terminus possesses short stretches of Asp, Glu or Ser flanked by non-anionic amino acids Gly, Ala, Val, or Asn. In a recent paper,[16] we have described evidence for similar anionic domains of $(Asp)_n$ and $(Pser\text{-}Asp)_n$ flanked by intervening domains containing Gly,

Ala, Val, Glu, Ser, and Leu. It is shown that these intervening amino acids respond to changes in the protonation state of the anionic domains.[17] It is interesting that these intervening amino acids also occur in the 23 residue N-terminal sequence (Figure 1) in exactly the same manner as the rest of the protein. Additionally, the presence of a contiguous Ser-Asp sequence in the N-terminal domain lends support to our earlier NMR findings of the existence of heteropolymer repeats of $(Pser\text{-}Asp)_n$ in BDPP;[16] however, at this time we cannot specify the sites of serine phosphorylation on the 23 mer.

Sequence Homologies Between BDPP N-terminal Sequence and Other Known Protein Sequences.

With the BDPP N-terminal sequence data, we have performed a protein sequence databank search of GenBank, EMBL, NBRF Protein, and New Enzyme List, to ascertain if there exist homologies between BDPP N-terminal sequence and other known protein sequences. In particular, we are interested in the existence of stretches of Asp, Glu or Ser flanked by intervening non-ionic amino acids. As shown in Figure 1, the search yielded 5 matches with varying homology (39% to 52%) per 23 residues. On the basis of this analysis (Figure 1), we found that: (1) BDPP showed homology (>40%) to the anionic/intervening domains found in the following proteins: (a) Herpes Simplex Virus Type I (HSV-1) 175 kD unique transcriptional activator protein (Vmw175; also known as ICP-4 protein);[18-20] (b) DNA-directed RNA polymerase, σ subunit (*Samonella typhimurium*);[21] and (c) the secretory gene product 7(*sec 7*) of the

temperature-sensitive mutant of *Saccharomyces Cerevisiae*[22] (Figure 2); and (2) homologies of less than 40% were found between BDPP and Asp-Glu rich C-terminal region of Cauliflower mosaic virus (CaMV) coat protein (55 kD precursor of open region IV)[23](Figure 1). By substituting Asp or Ser at the 6 and 20 position of the BDPP, the homology could increase to 57% for HSV-1 transcriptional activator (position 6, Asp), to 52% for *sec 7* (position 20, Ser) and 43% for Cauliflower mosaic virus (position 20, Ser). These sequence fragments all possess anionic domains containing Asp or Glu, and non-ionic intervening amino acids such as Ala, Val, Asn, Pro, Gly, Thr, and Ser.

Peptide Structure Predictions.

Further analysis of the peptide sequences was performed to determine what other similarities exist at the level of peptide structure prediction (Chou-Fasman[24-26]) including hydrophobicity and hydrophilicity (Kyte and Doolittle[27]). It has been established that three-dimensional structures are more conserved in evolution than primary sequences.[28-30] A two-dimensional squiggle plot[31] detailing helical, beta-sheet, turn, and coil regions and hydrophilicity index predictions (Kyte-Doolittle; >1.3) is presented in Figure 2 for BDPP N-terminus and the corresponding fragments from HSV-1 Vm175, CaMV coat protein, DNA directed RNA polymerase σ subunit, and *sec* 7 protein. For BDPP, we substituted either Ser or Asp at positions 6 and 20; the presence of either amino acid did not affect the outcome of the Chou-Fasman prediction. In the case of the hydrophilicity plot, the presence of Ser at positions 6

and 20 reduced the hydrophilic moment at those positions by 50% as compared to Asp. The hydrophilicity plots reveal that all of the homologous fragments possess a large percentage of residues (>70% composition) that would be predicted to be hydrophilic. In addition, for each fragment the Chou-Fasman plots predicts the existence of turn domains (ranging from 6 to 9 turns per 23 residues), with BDPP possessing the largest number of turn regions (15 turns per 23 amino acids) (Figure 2). Note that the positioning of the turns varies with the protein sequence. Each protein fragment would be predicted also to have beta-sheet and coil (random) structure, with BDPP containing the highest content of both. The presence of beta-sheet structure has been shown experimentally to be characteristic of anionic charge domains in PMMP-divalent metal ion or PMMP-crystal complexes. [32-34] Another prediction is that the 23 residue N-terminus of BDPP would contain little if any helix structure, whereas the *sec* 7 fragment would contain two helical regions (residue 1-9, 12-18) separated by turn and sheet regions. For the σ subunit and HSV-1 Vm175 protein fragments, each is predicted to possess a single helical region of varying residue length, flanked by turn regions on either side of the helix (Figure 2). Thus, although the detailed theoretical arrangement of secondary structure in each fragment is different, there exist similarities in hydrophilicity, beta-sheet, coil, and turn content within the 23 amino acid residue fragments.

Structure-Function Considerations

What the above sequence data and peptide structure predictions suggest is that the N-terminal fragment of BDPP shares a number of sequence and domain motif similarities with other proteins which function intracellularly and are not involved in extracellular mineralization. At this time we are not able to determine what other parts of the BDPP sequence beyond the N-terminus may possess sequence homology with these or other proteins. On the other hand, we have come across other interesting similarities that exist between BDPP and these other proteins. 1) Regulation. Previous studies have established that several of these proteins are involved in regulatory functions of one kind or another, either at the level of transcription (HSV-1 activator Vm175;[18-20,35] DNA-directed RNA polymerase σ subunit[21]) or intracellular processing (*sec 7*).[22] BDPP itself has been postulated to be involved in the regulation of calcium phosphate mineralization. However, based on present information, none of these other proteins are involved directly in Ca^{2+} regulation. We note that none of the database searches has resulted in the detection of any sequence homology between the 23 amino acid BDPP N-terminal fragment and known Ca^{2+} regulatory proteins, such as calcitonin, calmodulin, intestinal calcium binding protein, or calbindin,[36,37] namely, the EF hand or helix-loop-helix class of proteins. Unfortunately, we do not know if this 23 residue peptide domain is involved in divalent cation binding in BDPP, or if the sequences in the other proteins participate in the formation of intra- or intermolecular peptide salt bridges as has been demonstrated in the

case of BDPP (Evans and Chan, 1991a). 2) Phosphorylation. From the existing body of data available to us, it is evident that most of the proteins in this study are phosphorylated. It is known that *sec 7* does bind Ca^{2+}, and is phosphorylated at a number of Ser residues within its sequence,[22] a trait shared by BDPP. Similarly, phosphorylation also occurs in CaMV coat protein (primarily as Pser (84%), PThr (14%) and PTyr (2%))[38,39] and in the HSV-1 Vmw 175 transcriptional activator protein which has a serine-rich domain.[18,40]

Insights into the Evolution of Polyelectrolyte Mineral Matrix Proteins.

Does this sequence homology data shed any insights into the evolution of mineral matrix proteins or PMMPs? For the sake of discussion, let us assume that these regulatory proteins, critical for cellular or viral function, appeared prior to the development of mineralization and the appearance of the PMMP. Two mechanisms by which the extracellular PMMP may have evolved are as follows: (1) It might be that existing intracellular regulatory proteins became activated by extracellular stress signals (see below), and were somehow exported outside the cell, where their sequence allowed them to function fortuitously as regulators of available cation concentrations. (2) Conversely, a "new" protein might have been generated from mutation or rearrangement of existing genomic fragments endowed with the ability to bind divalent cations to insure the survival of the cell under conditions of low $(M^{2+})_{free}$. In either (1) or (2), the function of the "new" or exported protein might not be to induce mineralization, but rather to bind divalent cations and regulate

the $(M^{2+})_{free}$ outside the cell. The mechanism by which this "new" protein modulated solution-to-solid transformation might have evolved at a later time.

In earlier reports it was speculated that environmental pressure may have played a role in the evolution of mineralization.[7,8] We would like to offer two possible conditions under which the appearance of the PMMP might have taken place: (1) The availability of free cations outside the cell. Changes in the free cation concentration might have acted as a "shock" stimulus, inducing the transcription of proteins to stabilize the cell. (2) Temperature change. An interesting trait which three of these diverse non-BDPP proteins possess is their ability to regulate cellular activity in response to alterations in temperature. Temperature-sensitive mutants exist either at the level of transcription (RNA polymerase σ subunit;[21] HSV-1 Vm175 transcriptional activator[18-20]), or post-translational secretion (*sec 7* [22]). It is conceivable that temperature shift might have been one of the signals which induced gene expression of proteins which later developed into PMMPs.

The concept of intracellular calcium mechanisms evolving prior to the appearance of calcareous skeleton development has previously been proposed by Lowenstam and Margulis.[8] In their hypothesis, early organisms, most likely bacteria, evolved a means of exporting Ca^{2+} to protect Mg^{2+}-based enzymes and prevent the precipitation of calcium phosphates intracellularly. The development of intracellular calcium binding proteins was postulated as one response to this cellular demand

for Mg^{2+} homeostasis. However, from the evidence available to us, there appear to be very little if any existing sequence homologies between intracellular Ca^{2+} EF hand proteins, such as calmodulin, and BDPP. We suggest that at the same time or later in evolution to divalent cation export, certain template mineralization molecules, distinct from the EF hand Ca^{2+} proteins, made an appearance; and that this appearance was stimulated by events occurring external to the cell. These molecules included proteins which bind Ca^{2+} and transport or sequester it. This reasoning is consistent with the hypothesis put forth by Lowenstam,[41,42] in which mineralization developed from "pin-point" mineral deposition into defined mineral phases that exhibit contiguous, ordered structures. This progression from random foci to ordered phases most likely paralleled the evolution of the PMMP and other organic templates which could exert control over the precipitation and growth of crystal foci. Admittedly, the exact stimulus for this development of Ca^{2+} export and precipitation regulation has not yet been established. However, our present findings pose an interesting evolutionary question which hopefully will be answered by further research: How did mineralization templates, and mineralization, evolve in response to environmental stress?

Acknowledgements

We thank Professor Heinz Lowenstam for his reading of the manuscript and helpful comments. JSE is a recipient of a National Research Service Award from the NIH (DE-05445-02).

REFERENCES

1. Weiner, S. (1983) *Biochemistry* 22 4139-4145.

2. Addadi, L., Moradian, J., Shay, E., Maroudas, N.G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* 84 2732-2736.

3. Uchiyama, A., Suzuki, M., Lefteriou, B., and Glimcher, M.J. (1986) *Biochemistry* 25 7572-7583.

4. Gotoh, Y., Sakamoto, M., Sakamoto, S., and Glimcher, M.J. (1983) *FEBS Letters* 154 116-120.

5. Stetler-Stevenson, W.G., and Veis, A. (1983) *Biochemistry* 22 4326-4335.

6. Tsay, T., and Veis, A. (1985) *Biochemistry* 24 6363-6369.

7. Lowenstam, H.A., and Weiner, S. (1989) in *On Biomineralization* Oxford University Press, New York, New York pp. 260-271.

8. Lowenstam, H.A., and Margulis, L. (1980) *Biosystems* 12 27-41.

9. Mann, S. (1988) *Nature* 332 119-124.

10. Addadi, L.., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* 82 4110-4114.

11. Narwot, C.F., Campbell, D.J., Schroeder, J.K., and Valkenburg, M.V. (1976) *Biochemistry* 15 (16) 3445-3449.

12. Marsh, M.E. (1989) *Biochemistry* 28 339-345.

13. Fujisawa, R., Kuboki, Y., and Sasaki, S. (1987) *Calcif. Tiss. Int.* 41 44-47.

14. Stetler-Stevenson, W.G., and Veis, A. (1987) *Calcif. Tiss. Int.* 40 97-102.

15. Linde, A., Bhown, M., and Butler, W.T. (1980) *J. Biol. Chem.* 255 (12) 5931-5942.

16. Evans, J.S., and Chan, S.I. (1991a) *Biochemistry*, in press.

17. Evans, J.S., and Chan, S.I. (1991b) *Biochemistry*, in press.

18. McGeoch, D.J., Dolan, A., Donald, S., and Brauer, D.H.K. (1986) *Nucl. Acids Res.* 14 (4) 1727-1745.

19. O'Hare, P., and Hayward, G.S. (1985) *J. Virology* 53 (3) 751-760.

20. Dixon, R.A.F., and Schaffer, P.A. (1980) *J. Virology* 36 (1) 189-203.

21. Erickson, B.D., Burton, Z.F., Watanabe, K.K., and Burgess, R.R. (1985) *Gene* 40 67-78.

22. Achsteter, T., Franzusoff, A., Field, C., and Schekman, R. (1988) *J. Biol. Chem.* 263 (24) 11711-11717.

23. Franck, A., Guilley, H., Jonard, G., Richards, K., and Hirth, L. (1980) *Cell* 21 285-294.

24. Chou, P.Y., and Fasman, G.D. (1974) *Biochemistry* 13 (2) 211-221.

25. Chou, P.Y., and Fasman, G.D. (1974) *Biochemistry* 13 (2) 222-245.

26. Chou, P.Y., and Fasman, G.D. (1977) *J. Mol. Biol.* 115 135-175.

27. Kyte, J., and Doolittle, R.F. (1982) *J. Mol. Biol.* 157 105-132.

28. Bajaj, M., and Blundell, T. (1984) *Ann. Rev. Biophys. Bioeng.* 13 453-492.

29. Johnson, M.S., Sutcliffe, M.J., and Blundell, T.L. (1990) *J. Mol. Evol.* 30 43-59.

30. Chothia, C., and Lesk, A.M. (1986) *EMBO J.* 5 (4) 823-826.

31. Jameson, J.S., and Wolf, E.G. (1988) *CABIOS* 4 (1) 181-186.

32. Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* 82 4110-4114.

33. Lee, S.L., Veis, A., and Glonek, T. (1977) *Biochemistry* 16 (13) 2971-2979.

34. Weiner, S., and Traub, W. (1984) *Phil. Trans. R. Soc. London* B 304 425-434.

35. Everett, R.D. (1988) *J. Mol. Biol.* 203 739-751.

36. Levine, B.A., and Dalgarno, D.C. (1983) *Biochem. Biophys. Acta* 726 187-204.

37. Strynadka, N.C.J., and James, M.N.G. (1989) *Ann. Rev. Biochem.* 58 951-998.

38. Martinez-Izquierdo, J., and Hohn, T. (1987) *Proc. Natl. Acad. Sci USA* 84 1824-1828.

39. Hahn, P., and Shepherd, R.J. (1982) *Virology* 116 480-488.

40. Wilcox, K.W., Kohn, A., Skylanskaya, E., and Roizman, B. (1980) *J. Virology* 33 167-182.

41. Lowenstam, H.A. (1978) in *Biogeochemistry of Amino Acids* (Eds. P.E. Hare, T.C. Hoering, and K. King, Jr.) John Wiley and Sons, New York, New York, pp. 3-16.

42. Lowenstam, H.A. (1984) *Palaeontology* 2 79-95.

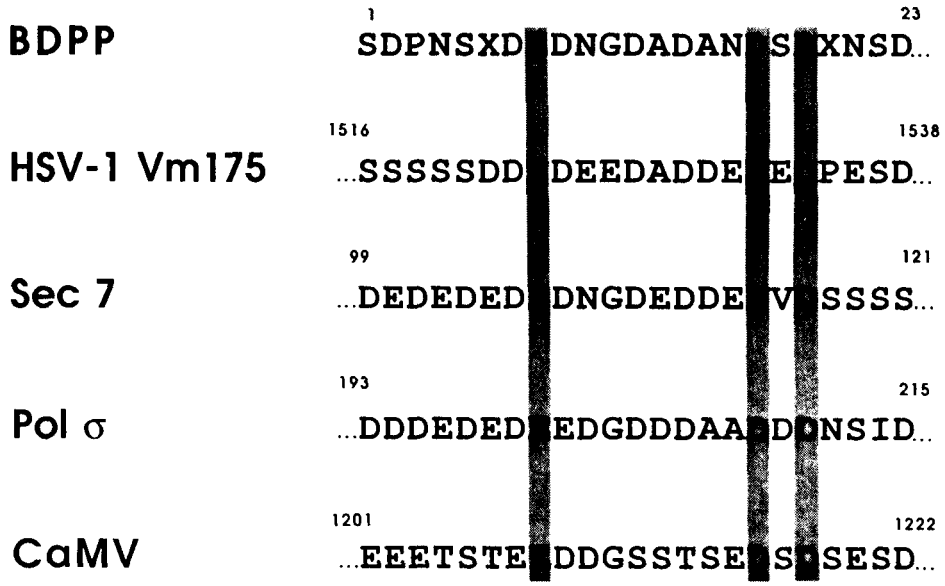
43. Pearson, W.R. (1990) *Methods in Enzymology* 183 63-98.

44. Doolittle, R.F. (1990) *Methods in Enzymology* 183 99-110.

Figure 1: *N-terminal 23 Residue Amino Acid Sequence of BDPP and its Comparison with Database Sequences.*

The N-terminal 23-residue amino acid sequence of BDPP was analyzed by the automated Edman gas-liquid phase sequencing technique. BDPP protein, (10 µl sample, 1 µg) purified as described (Evans and Chan, 1991a)(>95% homogeneity), was applied on a polybrene coated glass filter and analyzed on a Applied Biosystems 477A Liquid-Gas Phase Sequencer and PTH analyzer. A total of 35 cycles were performed on a single sample. Cycles were manually analyzed and compared to computer analysis of the cycles. Three runs on parallel samples were performed. Amino acids are denoted by the single-letter code. Data represent the best fit between the three determinations. **X** denotes uncertainty in that particular cycle; possible amino acids are either Asp or Ser.

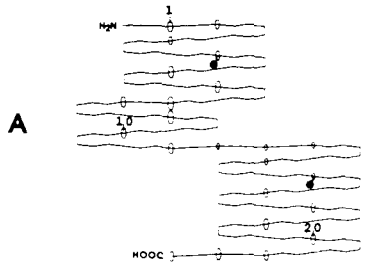
The 23 residue N-terminal BDPP sequence was compared to known protein sequences in the GenBank, NBRF, EMBL, and New Enzyme List sequence databanks, using the interactive Genetics Computer Group software programs WORDSEARCH[43] and/or TFASTA[44] on a Digital VAX/VMS 5.7A operating system. The search, which compared two-word databits, included two ambiguities (X) in the BDPP N-terminal sequence at positions 6 and 20, and resulted in the following matches (>40% homology) to known sequences. Shaded columns outline the consensus sequence of the five proteins. Numbers indicate the residue position in the overall primary amino acid sequence. ‡Per 23 amino acids. n/a: not applicable.

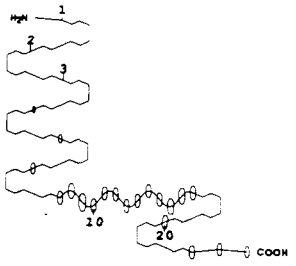


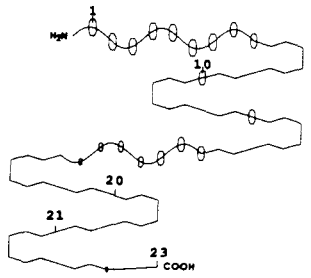
	% Homology	Anionic Residues†
BDPP	n/a	9
HSV-1 Vm175	52	15
Sec 7	43	16
Pol σ	43	17
CaMV	39	12

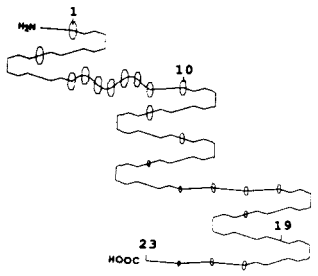
Figure 2: *Chou-Fasman/Kyte-Doolittle Squiggle Plots of Predicted Secondary Structures for BDPP and Homologous Protein Fragments.*

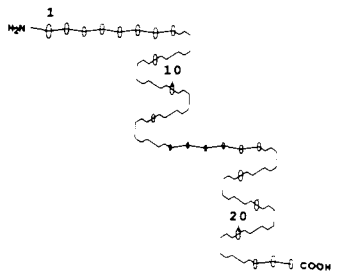
Squiggle plots have been constructed for each of the sequence fragments in Figure 1 using the Genetics Computer Group software programs PEPTIDESTRUCTURE/PLOTSTRUCTURE.[31] For comparison, sequences are numbered 1 through 23; the actual sequence positions are given in Figure 1. Chou-Fasman secondary structures are represented as different wave forms: **helix:** sine wave; **β-sheet:** sharp saw-tooth wave; **turn:** 180° turn; **coil:** dull saw-tooth wave. Kyte-Doolittle hydropathy/hydrophilicity predictions are identified at each residue by either a hexagon (hydrophilic) or a diamond (hydrophobic), with the size of the symbol proportional to the value of hydrophobicity or hydrophilicity. The Kyte-Doolittle threshold for both values has been set to a value greater than or equal to 1.3, as noted at the bottom of the Figure. Potential sites of glycosylation, denoted by a solid black circle, are predicted for sites where the residues have the composition NXT or NXS. When X = D, W, or P residue, then the site is assumed to be a weak glycosylation site; otherwise, it is a strong glycosylation site. In this plot, the unidentified positions 6 and 20 (X) of the BDPP sequence are filled by Ser. Substitution of either Ser or Asp at position 6 and 20 results in no change in the Chou-Fasman prediction, and only a small decrease in hydrophilicity for the Ser residue. Legend to Figures: **A**: BDPP 23 residue N-terminal sequence. **B:** HSV Vmw 175 transcriptional activator peptide fragment. **C:** Sec 7 secretory protein fragment. **D:** DNA-directed RNA polymerase, σ subunit fragment. **E:** Cauliflower Mosaic Virus 55 kD coat protein fragment.











◊Kyte-Doolittle Hydrophilicity >= 1.3
◊Kyte-Doolittle Hydrophobicity >= 1.3

Three

Preliminary Evidence for Novel Chemical Groups in Bovine Dentine Phosphophoryn, a Polyelectrolyte Mineral Matrix Protein.

Abstract: A unique pair of ^{1}H NMR resonances have been identified in purified Bovine Dentine Phosphophoryn (BDPP), a polyelectrolyte mineral matrix protein of tooth dentine. These resonances, occurring at 0.032 and 0.155 ppm from TSS, represent 3.6% of the total proton spins in the BDPP protein. Since BDPP possesses a low content of aromatic amino acids (1% or less of the total amino acid composition), it is unlikely that these peaks arise entirely from amino acid methyl sidechain protons in the protein. However, it is known that silicon atoms can induce through-bond deshielding effects on protons. An ICP quantitative analysis reveals the presence of Si in the BDPP protein (13-14 atoms/molecule BDPP), which could account for 50% of the proton spins resonating at this high field. The unusual finding of Si bound to BDPP is supported by earlier in vivo observations in which Si levels were identified at the mineralization front in bone [Carlisle, 1981], and, the recent finding that silicate-associated oleic acid molecules exist in bone extracellular matrix at the time of mineralization onset [Kashiwa *et al.*, 1992]. Thus, we believe that BDPP may possess novel chemical modifications, perhaps related to its role in calcium phosphate biomineralization..

Introduction

A unique family of proteins, termed polyelectrolyte mineral matrix proteins (PMMP), have been identified on the basis of their unusual content (60% or greater) and sequence arrangement of negatively charged amino acids [1-7]. This superfamily of proteins utilizes the "template" principle to actively bind and create a three-dimensional array of metal ions on the protein surface [8]. In biology, this process of utilizing organic molecule templates to direct the synthesis of inorganic complexes in biology is termed biomineralization, and its occurrence has been identified in all phyla [9,10].

The chemistry of the template-metal ion binding phase of mineralization, termed the nucleation process, is poorly understood with regard to the sequence of events that initiate the metal ion sequestration by the PMMP template. One approach to unraveling this puzzle is the study of the PMMP conformation in the absence and presence of cations that are specific for the forming mineral phase. We have examined the conformation of one particular PMMP protein, bovine dentine phosphophoryn (abbreviated BDPP), an Asp, o-phosphoserine (PSer) rich 155 kD polypeptide that is responsible for the nucleation of amorphous calcium phosphate [11,12] and the formation of its mature crystalline phase, hydroxyapatite (HAP)[13,14]. During the course of our ^{1}H NMR spectroscopy studies on the conformation of the purified BDPP protein [15], we noted the presence of proton spins resonating near 0.00 ppm. This observation was considered unusual, since it has been established that phosphophoryn and other PMMP proteins contain few aromatic amino acids (<1% of the total amino acid composition), and therefore the occurrence of aromatic ring current shift effects on sidechain methyl protons in BDPP cannot account for all these spins resonating at higher field. In this preliminary report, we examine the nature of this ^{1}H NMR deshielding effect, and provide preliminary evidence for the presence of silicon (Si) atoms in the BDPP molecule.

Materials and Methods

Isolation and Purification of BDPP. All water utilized in this study was of the distilled doubly deionized form generated by a Barnstead NanoPure II Water Purification System. A full description of the purification protocol will be given in a subsequent article [15]. In general, the procedure for the isolation of BDPP from unerupted mandibular bovine third molars (taken from steers, 1 - 3 years old) follows the general scheme outlined by Stetler-Stevenson and Veis [2], with the substitution of 0.1 M EGTA/0.5 M EDTA, pH 7.50 solution, for the resolubilization of the $CaCl_2$ precipitate, and, for the cation depletion step after ion exchange and molecular sieve chromatography. Chelex-100 resin (Bio-Rad) was also employed on the purified BDPP protein prior to NMR and ICP analysis to create a divalent cation depleted form of BDPP ($[Ca^{2+}] < 1$ ppm/μg as determined by ICP analysis). Capillary zone electrophoresis (CZE) was utilized to assess the purity of the BDPP protein; details of the CZE procedure can be found in Figure 3.

Elemental Analysis of BDPP. BDPP was assayed for Ca^{2+} and Si via inductively coupled plasma atomic emission spectrometry (ICP), using a Plasma 200 Spectrometer fitted with a double monochrometer. All samples, including blanks, were handled in plastic containers. The purified BDPP sample (50 μg/mL in DD_{water}; 3 mL total volume) was analyzed along with a DD_{water} blank, and a 20 μg/mL bovine serum albumin sample (in DD_{water}) as a control. Two separate determinations on each sample were made. Quantitation of elemental levels were conducted on solution standards of Si and $Ca(NO_3)_2$.

^1H and ^{29}Si NMR Spectroscopy. The purified BDPP sample was dissolved in 400 μL of D_2O (99.96% atom-*d*)(final concentration: 12 mg/mL), and the pH was adjusted to 7.1 with NaOD solution (98.6% atom-*d*), uncorrected for deuterium isotope effects. One-dimensional ^1H NMR (500.138 MHz) spectra were recorded at 300 K and 363 K, and two-dimensional ^1H-^1H COSY-90 spectra (using the Bruker microprogram COSYHG.AUR) were recorded at 300 K on a Bruker AMX-500 Spectrospin spectrometer, using an Aspect X32 computer for data acquisition and processing. ^{29}Si NMR (99.35 MHz) spectra, with and without composite-pulse proton-decoupling, were also acquired on the same sample, using a 5 mm inverse broadband probehead. NMR acquisition and processing parameters are given in the

appropriate figure legends. Proton quantitation was performed using an external $CHCl_3$ sample run under identical acquisition conditions; integration of the BDPP ^1H NMR spectrum was conducted upfield and downfield of the HOD solvent peak. 1-D ^1H NMR spectra were also acquired for a sample containing 5 μL lyophilizer pump oil (in 400 μL D_2O, and 400 μL D_2O containing 1 M NaCl.)

Results and Discussion

Evidence for Modified Groups in BDPP. Unique to the native BDPP protein are a set of proton spins centered at 0.032 and 0.155 ppm (Peaks 1,2) that appear as narrow resonances ($\Delta\nu_{1/2}$ = 12-20 Hz) which are superimposed upon a broad baseline resonance component (50 Hz or greater)(Bottom spectrum, Figure 1A). These two peaks represent 3.6 % of the total observable proton spins in the protein. It should be noted that the resonance frequencies for Peaks 1 and 2 are 0.5 to 1.0 ppm upfield from values reported for methyl proton resonances of Ala, Val, Ile, and Leu sidechains in free amino acids or in peptides (16,17). These high-field resonances were not observed in earlier ^1H NMR studies on the 30 kD rat phosphophoryn fragment [18]. Additionally, Peaks 1 and 2 do not possess detectable *J*-scalar couplings with any other resonances along the diagonal of the COSY-90 contour plot [Figure 2]. Since phosphophoryn possesses a low content of aromatic π-bonding ring systems (Tyr, Phe)[2,19,20], the potential for ring-current shift effects on BDPP sidechain methyl protons is low, and therefore this deshielding effect could only account for approximately 1-1.5% of the total spin population in BDPP. Given the stringency in the purification scheme utilized in the isolation of BDPP, it is unlikely that impurities, such as lipids, would be associated with the protein; such contaminants would have been separated and detected by CZE at 200 nm (Figure 3). In addition, lyophilizer pump oil, a possible contaminant, possesses a broad methyl resonance at 0.8 75 ppm and a methylene resonance at 1.292 ppm when suspended in D_2O (5 μL oil/400 μL D_2O). This was not observed in our BDPP ^1H NMR spectrum. That these peaks represent chemical groups in BDPP is further demonstrated in Figure 1A, where increasing the ionic strength at constant pH resulted in the reduction in area and intensity not only for Peaks 1 and 2, but also for other peaks which represent aliphatic amino acid sidechain methyl groups in the protein [Figure 1A: Top, Bottom spectrum, arrows]. The cooperative ionic strength effect would not be observed for a non-covalently-

associated impurity. In this same figure, denaturation by temperature results in changes in peak intensity and area for Peaks 1 and 2 in BDPP. Again, the change in intensities for Peaks 1 and 2 as a function of temperature would be atypical for a small impurity. Since denaturation would be expected to disrupt any aromatic ring current shift interactions with clustered methyl protons in the hydrophobic region (which, incidentally, has been found to exist in PMMP proteins [21,22]), it is surprising to find residual proton signals at 0.032 and 0.155 ppm after extreme thermal denaturation and slow cooling. Thus, although a small portion of the proton spins representing Peaks 1 and 2 (e.g., the broad component) might be attributable to deshielded amino acid sidechain methyl protons in the BDPP protein , the remaining methyl proton spins (narrow component) must be originating from *other chemical groups* in the BDPP protein.

In addition to ring current shift effects, other effects can give rise to proton deshielding. One possibility is the presence of silane or siloxane groups, where Si induced proton chemical shifts typically resonate near 0.0 ppm [23]. The presence of Si in Nature is not unusual; historically, silicon has been identified in complexes with cholesterol, lipids, and galactose residues [24]. A recent report has identified the presence of silicate-carboxylate complexes in mineral-inducing fatty acids in vertebrate bone extracellular matrix [25]. These complexes could only be dissociated by strong acid (e.g., HF). Significant levels of silicate or Si have also been historically identified in the early stages of calcium phosphate mineralization in vertebrate bone matrix [26], These levels of elemental Si, representing 0.08 - 1.00% by weight of the total extracellular matrix, are found in concentrated regions were mineralization is taking place [26]. Coincidentally, phosphophoryn is also localized at the mineralization front in the dentine extracellular matrix [20], and represents approximately 0.5% or less by weight of the total organic matrix [19, 20]. Given its role as a template macromolecule in nucleation, it is possible that BDPP may possess Si-bound complexes, perhaps involving specific attachment to amino acid sidechain or carbohydrate residues. These Si complexes may have remained with the protein during purification. To pursue this line of reasoning, we performed a quantitative Si elemental analysis on purified BDPP using ICP spectrometry (Table I). Compared to bovine serum albumin in DD_{water} and DD_{water} blank controls, it is evident that BDPP contains significant amounts of

Si: ≈ 13 moles Si per mole BDPP. If we assume a 1:1 association of Si to amino acid sidechains, this would figure to ≈ 1% of the total amino acid composition (i.e., ≈ 10-15 residues, based on a total number of 1130 amino acids in the protein)[2], which correlates with our proton quantitation findings for Peaks 1 and 2. ^{29}Si NMR experiments failed to detect the presence of any ^{29}Si species in the protein, even with a 20 second relaxation delay to allow for long T_1's usually associated with ^{29}Si, and signal accumulation in excess of 30,000 scans. However, it should be noted that the absence of a detectable ^{29}Si signal may be due to the low abundance of ^{29}Si spins in the BDPP sample. Thus, although it is tempting at this point to suggest that BDPP possesses Si atoms, it is not possible at this time to rigorously confirm this using ^{29}Si NMR.

Thus, it is probable that some percentage of the ^{1}H NMR deshielding effect observed for Peaks 1 and 2 is manifested by some type of chemical group in BDPP that is not an amino acid methyl group. The presence of a protein-bound Si species indeed would be unusual and interesting, in light of the fact that previous ^{29}Si NMR studies of biological silicates failed to demonstrate the presence of a Si-C or Si-O-C induced shift effect arising from a template protein-silica linkage [27,28]. Unfortunately, because of the low concentration of Si and associated protons in BDPP, we have not yet been able to correlate the protons in Peaks 1 and 2 with the ^{29}Si species, or establish any J_{SiH} with any other proton resonances in the ^{1}H NMR BDPP spectrum via two-dimensional NMR heteronuclear correlation techniques. For the time being, we are <u>tentatively</u> assigning Peaks 1 and 2 in the ^{1}H NMR BDPP spectrum to (a) a small population of Leu, Ile, Ala, and Val methyl sidechain protons in BDPP, and, (b) methyl protons arising from other groups associated with the protein.

Summary. From our ^{1}H NMR experiments (Figure 1) and the ICP analysis (Table I), there is reason to believe that a certain percentage of BDPP contains modified groups that are not attributable to amino acid sidechains. This unique observation of proton resonances near 0.00 ppm was not reported in the earlier ^{1}H NMR studies on the 30 kD phosphophoryn fragment [18], perhaps due to the loss of specific peptide fragments during isolation. Earlier biochemical studies on the intact bovine, rat, and other phosphophoryn protein isoforms failed to perform an elemental analysis of the protein [2,19,29], and thus it is possible that such data went

unnoticed. However, from the mounting evidence that Si performs a crucial role in the early stages of calcium phosphate nucleation in connective tissues such as bone [25, 26], it is possible that a certain percentage of the Si detected in these studies may be protein-bound. The observation that fatty acid acyl esters of silica exist at the mineralization front and induce calcium phosphate precipitation from metastable solutions [25] gives strong support to the notion that a protein template macromolecule might also possess bound silicate or siloxane groups. Further work is underway to identify the chemical form of Si in BDPP, and, the localization of these Si-containing moieties on the BDPP molecule.

Acknowledgments

We thank Mr. Gordon R. Bradford of the Citrus Research Center, University of California, Riverside, for performing the ICP analysis. JSE acknowledges a Post-doctoral National Research Service Award from the National Institute of Dental Research (DE-05445-05), and a fellowship award from AMGEN Pharmaceuticals.

References

1. Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) Domain structure and sequence distribution in dentin phosphophoryn. *Biochem. J.* 276 699-707.

2. Stetler-Stevenson, W. G., and Veis, A. (1983) Bovine dentine phosphophoryn: Composition and molecular weight. *Biochemistry* 22 4326-4335.

3. Weiner, S. (1983) Mollusc shell formation: Isolation of two organic matrix proteins associated with calcite deposition in the bivalve *mytilus californianus*. *Biochemistry* 22 4139-4145.

4. Weiner, S., and Hood, L. (1975) Soluble protein of the organic matrix of mollusk shells: A potential template for shell formation. *Science* 190 987-989.

5. Uchiyama, A., Suzuki, M., Lefteriou, B., and Glimcher, M.J. (1986) Isolation and chemical characterization of the phosphoproteins of chicken bone matrix: Heterogeneity in molecular weight and composition. *Biochemistry* 25 7572-7583.

6. Glimcher, M.J., and Lefteriou, B. (1989) Soluble glycosylated phosphoproteins of cementum. *Calcif. Tiss. Int.* 45 165-172.

7. Shiraga, H., Min, W., Van Dusen, W.J., Clayman, M.D., Miner, D., Terrell, C.H., Sherbotie, J.R., Foreman, J.W., Przysiecki, C., Neilson, E.G., and Hoyer, J.R. (1992) Inhibition of calcium oxalate crystal growth *in vitro* by uropontin: Another member of the aspartic acid-rich protein superfamily. *Proc. Natl. Acad. Sci. USA* 89 426-430.

8. Mann, S. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 35-62.

9. Lowenstam, H.A., and Weiner, S. (1989) in *On Biomineralization.* Oxford University Press, London, pp. 1-240.

10. Williams, R.J.P. (1988) in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P.Eds.) VCH Publishers, New York, New York pp. 1-34.

11. Marsh, M.E. (1989) Self-association of calcium and magnesium complexes of dentin phosphophoryn. *Biochemistry* 28 339-345.

12. Marsh, M.E. (1989) Binding of calcium and phosphate ions to dentin phosphophoryn. *Biochemistry* 28 346-352.

13. Narwot, C.F., Campbell, D.J., Schroeder, J.K., and Van Valkenburg, M. (1976) Dental phosphoprotein-induced formation of hydroxylapatite during in vitro synthesis of amorphous calcium phosphate. *Biochemistry* 15 (16) 3445-3449.

14. Moradian-Oldak, J., Furedi-Milhofer, H., Veis, A., Addadi, L., and Weiner, S. (1992) Acidic macromolecule-inorganic calcium phosphate crystal interactions. *Connect. Tiss. Res.* 27 (2-3) 189.

15. Evans, J.S., and Chan, S.I. (1992) *submitted.*

16. Grob, K-H., and Kalbitzer, H. R. (1988) Distribution of chemical shifts in ^{1}H nuclear magnetic resonance spectra of proteins. *J. Mag. Resonance* 76 87-99.

17. Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology* (eds. by Jardetzky and Roberts) Academic Press, New York, New York, pp. 143-186, 227-231, 234-270, 448-465.

18. Cookson, D. J., Levine, B. A., Williams, R. J. P., Jontell, M., Linde, A., and de Bernard, B. (1980) Cation binding by the rat-incisor dentine phosphophoryn. *Eur. J. Biochem.* 110 273-278.

19. Rahima, M., and Veis, A. (1988) Two classes of dentin phosphophoryns, from a wide range of species, contain immunologically cross-reactive epitope regions. *Calcif. Tiss. Int.* 42 104-112.

20. Veis, A. (1982) Biochemical studies of vertebrate tooth mineralization. In *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Publishers, New York, New York, pp.189-223.

21. George, A., Simonian, P., Tylzanowski, P., and Veis, A. (1992) Identification of a novel acidic tooth protein. *Connect. Tiss. Res.* 27 (2-3) 107.

22. Rusenko, K., Donachy, J.E., and Wheeler, A.P. (1991) in *Surface Reactive Peptides and Polymers: ACS Symposia* (American Chemical Society) pp. 107-124.

23. Schraml, J. (1990) ^{29}Si NMR spectroscopy of trimethylsilyl tags. *Prog. NMR Spectroscopy* 22 289-348.

24. Iler, R.K. (1955) *The Colloid Chemistry of Silica and Silicates*. Cornell University Press, Ithaca, New York, pp. 18-86; 294-296.

25. Kashiwa, H.K., Thiersch, N.J., Pratum, T.K., and Roe, J. (1992) Oleic acid-silica complex that calcifies bone. *Connect. Tiss. Res.* 27 (2-3) 93.

26. Carlisle, E.M. (1981) Silicon in bone formation. In *Silicon and Siliceous Structures in Biological Systems.* (Simpson, T.L., and Volcani, B.E., Eds.) Springer-Verlag, New York, New York, pp. 69-93.

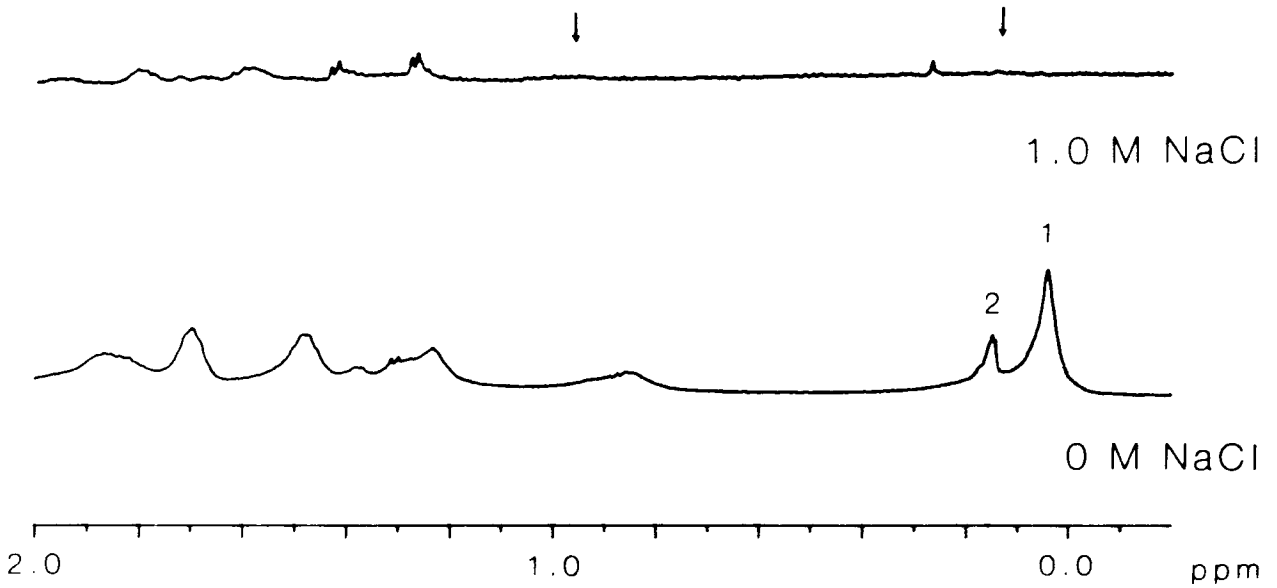
27. Perry, C.C. (1988) Chemical Studies of Biogenic Silica. In *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Publishers, New York, New York, pp. 233-256.

28. Fyfe, C.A., Thomas, J.M., Klinowski, J., and Gobbi, G.C. (1983) Magic-angle spinning NMR (MAS-NMR) spectroscopy and the structure of zeolites. *Angewandte Chemie* 22 (4) 259-336.

29. Lee, S.L., Veis, A., and Glonek, T. (1977) Dentin phosphoprotein: An extracellular calcium-binding protein. *Biochemistry* 16 (13) 2971-2979.

Figure 1: *One-Dimensional ^1H NMR 500 MHz Spectra of BDPP in D_2O, Under Varying Conditions of Ionic strength and Temperature.*

The high field aliphatic region of the BDPP NMR spectra is shown. Acquisition and processing parameters include a spectral window of 10 kHz, a pulse angle of 45°, receiver gain of 20 dB, recovery delay of 1 sec, 32 K data points (0.492 Hz/pt), a lorentzian apodization function of 1 Hz, and 445 scans. A homonuclear gated presaturation pulse (0.5 watts, decoupler offset centered on the HOD peak) is applied during the recovery phase (utilizing 4 dummy scans) to achieve solvent suppression. Proton chemical shifts are referenced from 0.1 M 2,2,2-trimethyl 2-silapentane sulfonic acid (TSS) in D_2O. Legend to Figures is as follows:
A: Bottom spectrum, 12 mg/mL BDPP in D_2O, no NaCl added; top spectrum, same concentration of BDPP in 1 M NaCl, pH 6.94, in D_2O. Both spectra recorded at 300 K. B: Bottom spectrum: 12 mg/mL BDPP in D_2O, 300 K; middle spectrum, same sample, at 363 K (90° C); top spectrum, same sample, slowly cooled to room temperature from 90° C over a 4 hr period, then recorded at 300 K.


1.0 M NaCl
1
2
0 M NaCl
2.0
1.0
0.0
ppm

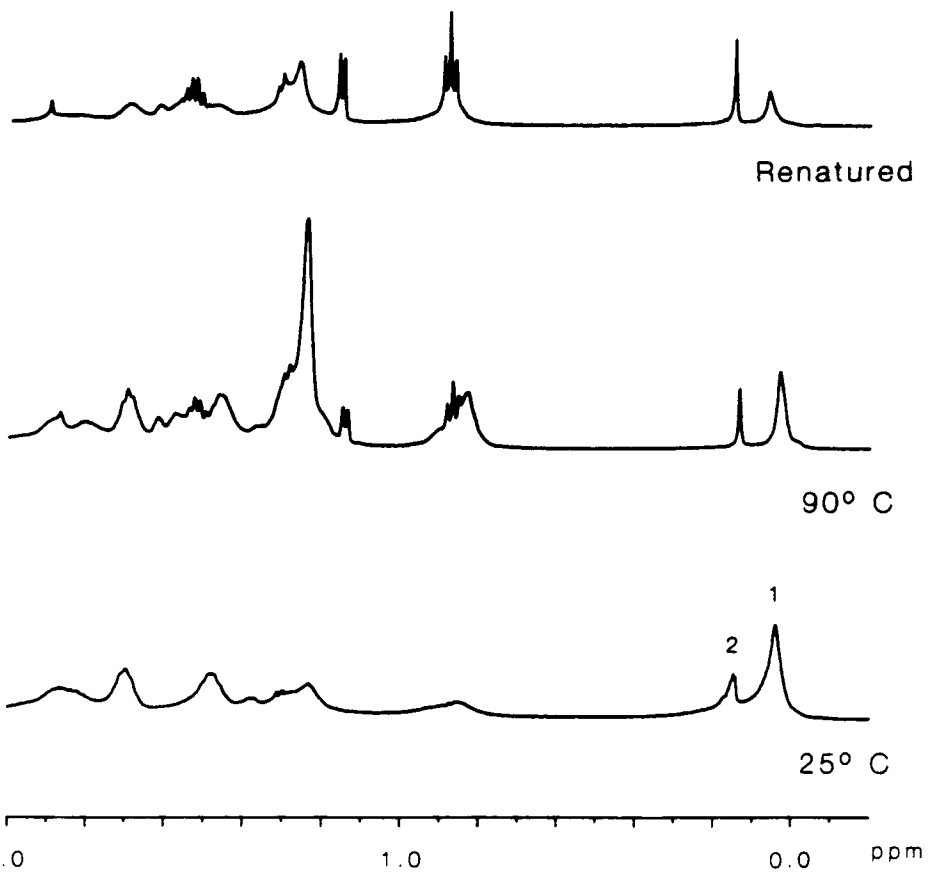
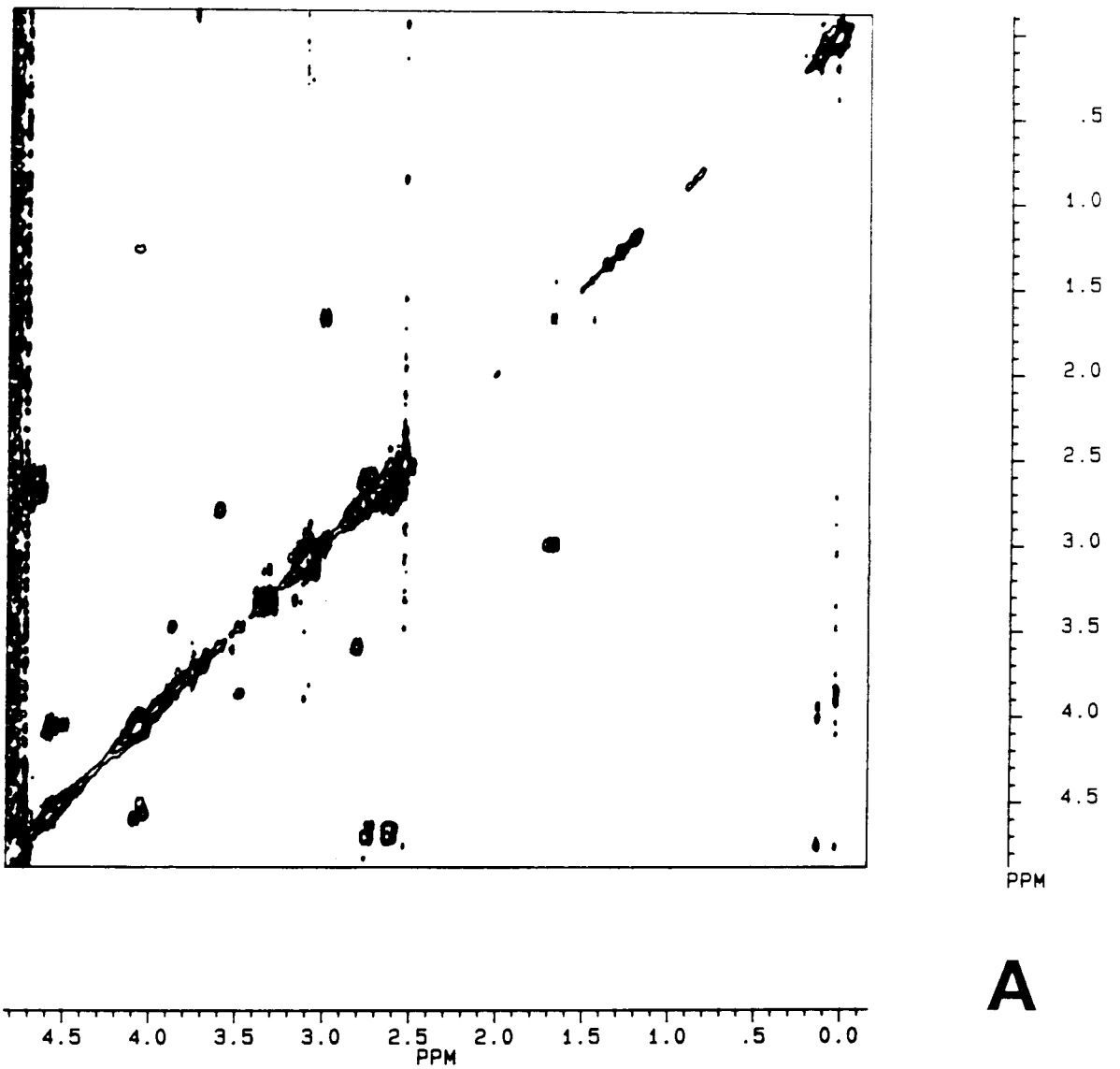

Renatured
90° C
1
2
25° C
2.0
1.0
0.0
ppm

Figure 2: *Two-Dimensional ^1H NMR 500.138 MHz COSY-90 Spectrum of BDPP in D_2O, pH 7.1, 300 K.*

Contour plot (unsymmetrized) of the aliphatic region of the BDPP spectrum is presented in the magnitude (positive) mode. Acquisition and processing parameters include a recovery delay of 1 sec., a 90° pulse of 9.7 μsec duration, a t_1 increment of 3 μsec, a spectral window of 5 kHz in both the F1 and F2 dimensions, a receiver gain of 40 dB, 128 scans per experiment, with a total of 512 experiments. A homonuclear gated presaturation pulse (0.5 watts, decoupler offset centered on the HOD peak) is applied during the recovery phase (utilizing 4 dummy scans) to achieve solvent suppression. A 2 K x 2 K data matrix (4.883 Hz/pt in both dimensions) is created with zero filling in the F2 dimension, using a squared sinebell window function with no shift.



A

Figure 3: *Capillary Zone Electrophoresis of BDPP.*

A 25 μg/mL and 500 μg/mL BDPP sample in DD_{water} were utilized in electrophoretic runs at (**A**): pH 2.5 [20 mM sodium phosphate buffer (NaP), normal polarity mode] and (**B**): at pH 11.0 [20 mM sodium 3-cyclohexylamino-1-propane-sulfonic acid (CAPS) buffer, reverse polarity mode], using an Applied Biosystems 270A Capillary Electrophoresis System running at 30° C, 25 kV constant applied voltage, electrokinetic loading, with monitoring of eluted components via absorbance at 200 nm. Capillaries utilized for separation were made of fused silica glass, and possessed dimensions of 72 cm total length, and an inner diameter of 75 mm. Applied sample volume ranged from 10 nL (pH 2.5 run; sample concentration = 25 μg/mL) to 42 nL (pH 11.0 run; sample concentration = 500 μg/mL).

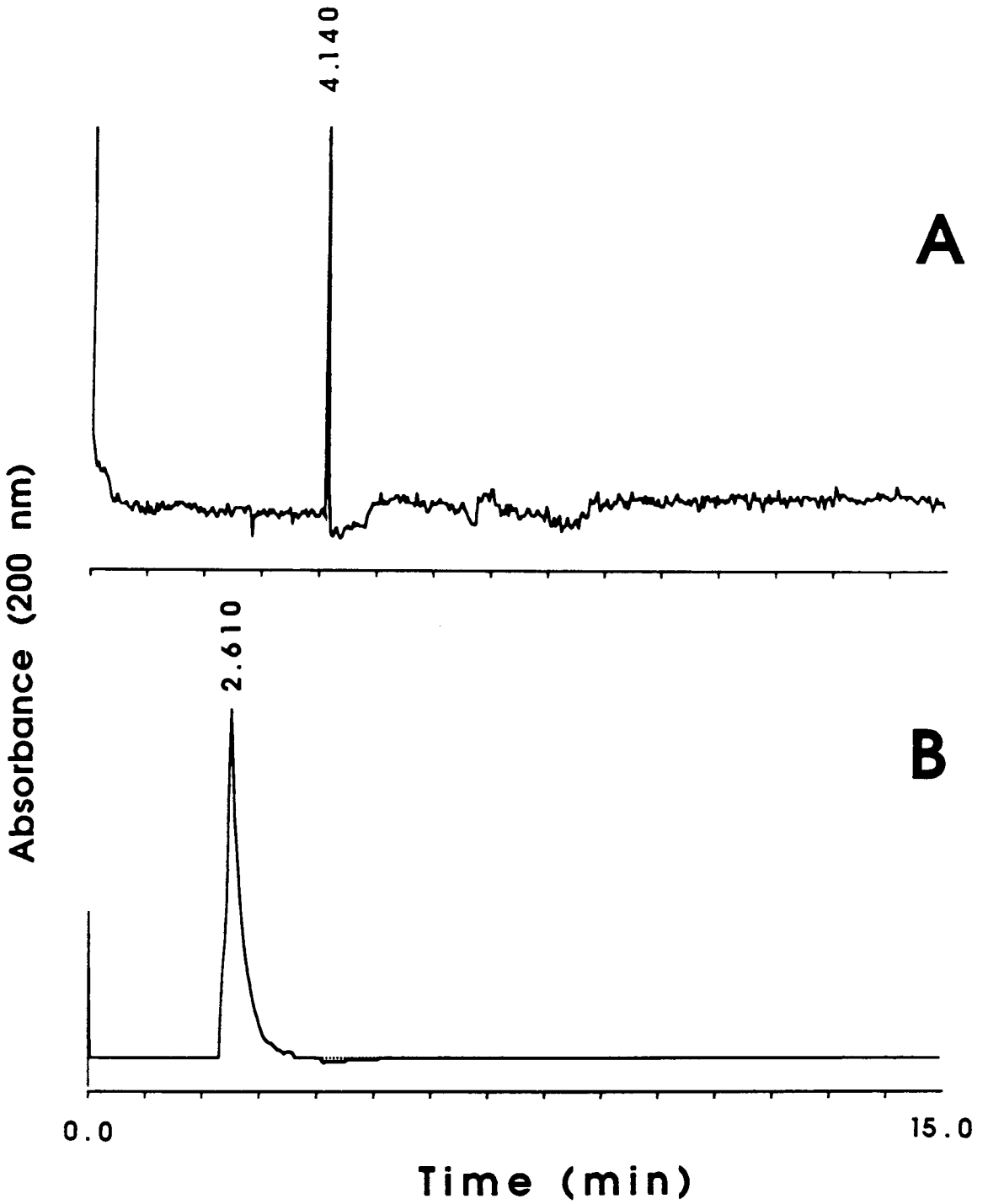

4.140
A
2.610
B
Absorbance (200 nm)
0.0
15.0
Time (min)

Table I: *Elemental Analysis of Si in BDPP.*

ICP analysis was performed on a 50 μg/mL BDPP sample (total volume: 3 mL in DD_{water}). Control samples included a DD_{water} blank, and bovine serum albumin (2 μg/mL in DD_{water}). Mean ppm values (N = 3) are reported in the Table. Mole ratio of Si/BDPP protein were based upon a BDPP MW of 155,000 [2].

Table I

ICP Analysis of Si Content

Sample	**Si Content, Uncorrected (ppm)**	**Adjusted Si Content (mole Si/mole protein)**
DD_{water}	0.0	N/A
DD_{water}/BSA	0.0	0.0
DD_{water}/BDPP	0.042	13.47

Four

NMR Studies on Bovine Phosphophoryn, a Polyelectrolyte Mineral Matrix Protein.

pH-Dependent Folding Experiments and Evidence for Polyelectrolyte Clustering of Phosphoserine and Aspartic Acid Residues.

Abstract

The protein folding behavior of a polyelectrolyte protein, Bovine Dentine Phosphophoryn (BDPP), in the pH range of 1.82 to 11.0 has been investigated. One- and two-dimensional $^{1}H/^{31}P$ NMR spectroscopy has been utilized to obtain proton spin assignments for amino acid residues in D_2O. 1-D ^{31}P NMR and 2-D $^{1}H/^{31}P$ HMQC NMR experiments verify the existence of three separate classes of PSer resonances in BDPP (α,β,χ), representing three distinct PSer residue populations at pH 6.94. By means of pH titration and ^{1}H NMR, six populations of Asp residues can be identified. Three of these populations show a broad titration profile and pK_a's for the γ–COOH group that are shifted to higher pH (1-2 pH units) than that of the free Asp amino acid. The other three Asp residue populations also possess higher pK_a's, but exhibit secondary inflection points on their pH titration curves that correspond to an observed pK_a of 6.17-6.95. The presence or absence of secondary inflection points for the Asp residues and the ^{31}P NMR chemical shift dispersion for the three PSer residue populations indicate that BDPP may be comprised of homologous $(Asp\text{-}Asp)_n$, $(PSer\text{-}PSer)_n$ and heterologous $(Asp\text{-}PSer)_n$ sequences arranged into polyelectrolyte cluster regions. The ε-CH_2 protons for Lys residues of BDPP do not show any definitive inflection point that would be characteristic of a primary amine, a behavior that may be indicative of salt-bridge formation with anionic sidechains.

INTRODUCTION

Biomineralization is a widespread phenomenon in nature that involves an inorganic solution-to-solid transformation under conditions that do not require high temperatures and pressures. The resulting solid phases, which range from amorphous to crystalline structures, act as a structural unit for biomechanical support and possibly act as a repository for essential ions [Lowenstam, 1981]. There is substantial evidence that this inorganic solution-to-solid transformation is mediated by template macromolecules [Lowenstam, 1981, Addadi *et al.*, 1987, Weiner, 1983, Mann, 1988]. A particular class of template macromolecules are proteins which are comprised of Asp, Ser, and Glu amino acids. Because these proteins possess a high negative charge density, they are often termed "acidic" proteins [Weiner, 1983, Mann, 1988]. One example is phosphophoryn, a highly phosphorylated high molecular weight polypeptide isolated from the tooth dentine of several species, including man [Stetler-Stevenson and Veis, 1983; Veis *et. al.*, 1972, Rahima and Veis, 1988]. Phosphophoryn is comprised of (1) Asp and PSer residues (80% of the total amino acid composition) in a nearly 1:1 stoichiometry [Stetler-Stevenson and Veis, 1983; Rahima and Veis, 1988]; (2) less abundant (<1-5%) amino acids such as non-phosphorylated Ser and Thr, Glu, and Pro; (3) basic amino acids Lys, His, and Arg; (4) aliphatic amino acids Ala, Val, Ile and Leu; and (5) aromatic amino acids (Tyr, Phe) also occur in the polypeptide sequence.

Of interest to the study of biomineralization is the structure of phosphophoryn in relationship to its function as a calcium phosphate nucleation activator. Previous investigations have shown that this protein possesses high and low affinity binding sites for Ca^{2+} [Lee *et al.*, 1977; Zanetti *et al.*, 1981]. Spectroscopic studies, including circular dichroism (CD) and infrared spectroscopy (IR) [Lee *et al.*, 1977], ^{1}H and ^{31}P NMR [Cookson *et al.*, 1980, Lee *et al.*, 1977, Lee *et al.*, 1983], transmission electron spectroscopy [Marsh, 1989a], and

viscometry studies [Stetler-Stevenson and Veis, 1987] have shown that a conformational change is induced in phosphophoryn upon binding of Ca^{2+}. However, the earlier spectroscopy studies did not fully address the conformation and solution structure of phosphophoryn in the absence of Ca^{2+} ion, which is a necessary prerequisite if a complete understanding of conformational change in the presence of Ca^{2+} is to be obtained. More importantly, earlier spectroscopic studies were conducted on phosphophoryn molecules which were not full-length in sequence [Lee *et al.*, 1977; Cookson *et al.*, 1980]. Thus, a re-examination of the the structure and behavior of phosphophoryn in the presence of Ca^{2+} needs to be undertaken.

Since the initial studies conducted in phosphophoryn, an intriguing hypothesis concerning the role of template proteins in the biomineralization process has emerged. Mann [1988] postulated that if template controlled crystal growth involves the arrangement of template anionic atoms into a precise stereochemical array for metal atom coordination, then the template must somehow increase its ligand atom density to approach that found in the complementary anion array in the crystal. Thus, in order to understand the mechanism by which phosphophoryn and other similar matrix proteins mediate nucleation, *it is essential to understand how a polyelectrolyte protein can fold into this complementary template surface for cation deposition.* One approach is to examine the protein folding transition(s) that accompany the cation loading of a metal ion-depleted protein template. In the divalent cation depleted state, repulsive charge-charge interactions between anionic sidechains will dictate the solution conformation of the protein. The cation-depleted protein conformer can then be conceived as a "denatured" state, which in the presence of divalent cations or other neutralizing species, achieves a "native" template state. Therefore, in a unique departure from earlier studies, we have applied homonuclear and heteronuclear one- and two-dimensional NMR spectroscopy ($^{1}H/^{31}P$) to the study of pH-induced folding effects on full-length cation-depleted adult bovine dentine phosphophoryn (BDPP). Because of the

polyelectrolyte behavior and and the significant backbone and sidechain motion exhibited by this large protein, we are able to successfully apply 2-D NMR correlation spectroscopy to identify the spin systems arising from amino acid sidechains which undergo rapid reorientation in solution. In this first series of investigations, we report the amino acid spin system assignments for the BDPP protein, and establish that the conformational folding of this protein is dominated by electrostatic effects from heterogeneous Asp and PSer containing sequence regions in the BDPP protein.

MATERIALS AND METHODS

Materials. The reagents used in this study were obtained from the following manufacturers: EDTA (tetrasodium salt) and $CaCl_2$ from US Biochemicals (Cincinnati, OH); EGTA (disodium salt), benzamidine-HCl, ε-aminocaproic acid, PMSF, and NEM from Fluka Chemicals (NY). Tris-HCl from BRL (Bethesda, MD); and ammonium bicarbonate (NH_4HCO_3) from J.T. Baker. Deuterium oxide (D_2O, 99.96% atom), deuterium chloride (DCl, 20% v/v solution, 98% atom), sodium deuteroxide (NaOD, 30% v/v solution, 98% atom), TSS, and EHDP from Aldrich Chemicals (Milwaukee, WI). Chelex-100 resin was purchased from Bio-Rad (Richmond, CA). All buffers and protein solutions were prepared using 0.2 μm filtered deionized distilled water (DD_{water}); all glassware used in NMR sample preparation and analysis was first treated with EDTA/EGTA to remove unwanted divalent cations.

Isolation of BDPP. The procedure for obtaining BDPP followed the protocol of Stetler-Stevenson and Veis (1983), with modifications as noted below. Unerupted mandibular third molars of 1-2 year old steers were obtained fresh from a local slaughterhouse (Hines Packing, Paramount, CA), packed on ice, and processed through the EDTA demineralization step as described (Stetler-Stevenson and Veis, 1983, Veis *et al.*, 1986). EDTA extract supernatant was dialyzed against

deionized distilled water (DD_{water}) at 4°C for two weeks, using Spectra/Por 4 dialysis tubing [molecular weight cut-off (MWCO) = 12-14 kD] and changing the outer solution for fresh DD_{water} every day. The resulting dialysate was concentrated [Amicon stir-cell, XM-50 membrane (MWCO=50 kD)], lyophilized, and stored at -80° C until needed.

Crude BDPP was recovered from the EDTA lyophilized extract via $CaCl_2$ precipitation (Stetler-Stevenson and Veis, 1983, Kuboki *et al.*, 1979). The precipitate, after 3 X washes with 1 M $CaCl_2$/20 mM Tris-HCl, pH 8.2, was solubilized with 0.5 M EDTA/0.1 M EGTA, pH 7.5. The mixture was stirred for 8 hrs at 4°C, and the soluble supernatant was clarified via centrifugation (RC-5B, SS34 rotor, 30 min) at 10,000 rpm, 4°C. The supernatant was dialyzed against DD_{water} for 2 weeks (Spectra/Por 4, MWCO 12-14 kD), lyophilized, and stored at -80°C.

Purification of BDPP. Purification of BDPP was performed as described elsewhere (Stetler-Stevenson and Veis, 1983) with the following modifications. Lyophilized EDTA/EGTA solubilized BDPP $CaCl_2$ precipitate was reconstituted in DD_{water} and chromatographed on a preparative FPLC Mono Q column (Pharmacia, Uppsala, Sweden), with eluant monitoring at 254 nm. Starting conditions for chromatography involved a mobile phase of 20 mM Tris, pH 7.5, with a constant flow rate of 0.5 mL/min and a pre-gradient run of 10 min following sample application. A linear gradient of 0 to 1.5 M NaCl in mobile phase buffer was applied (pH 7.5) for 50 min. A peak which eluted at approximately 1.0 M NaCl (retention time, 43 min) was collected, dialyzed against DD_{water} (Spectrapor 7, MWCO=50 kD), and lyophilized. This fraction was then chromatographed over a Pharmacia FPLC Sepharose 12 column (30 cm), using a mobile phase of 100 mM NH_4HCO_3/10 mM Tris-HCl, pH 7.5 at a flow rate of 0.5 mL/min and monitoring the eluant at 254 nm. This column had been calibrated using Bio-Rad Chromatography globular Protein Standards under identical solvent conditions. A single species eluting at a retention

time of 24 min, which corresponds to a molecular weight of approximately 155 kD as reported elsewhere (Stetler-Stevenson and Veis, 1983), was collected and dialyzed against DD_{water} as before. This 155 kD species gave rise to a single band (Stains-all/Coomassie) on 5-10% linear gradient SDS-PAGE gels [Stetler-Stevenson and Veis, 1983; Laemmli, 1970]. Purity was also analyzed by capillary zone electrophoresis as described in the previous manuscript [Evans *et al.*, 1992]. Typically, 120 third molars yielded approximately10 mg of pure BDPP.

Preparation of BDPP for NMR. The calcium-depleted form of the 155 kD BDPP protein was prepared by dissolving the protein in 0.5 M EDTA/0.1 M EGTA, pH 7.5 (1 mg/ml final protein concentration) and dialyzing against EDTA/EGTA for two days at 4°C (Spectra/Por 7, MWCO=50 kD). The dialysate was then transferred and dialyzed against DD_{water} (volume ratio: 1:2000) for 12 days at 4°C, with outer solution changes made every day. After this step, the dialysate was dialyzed against DD_{water} containing 100 g/L Chelex-100 (the pH was adjusted to 7.4) for two days. The dialysate was then lyophilized, reconstituted in D_2O, re-lyophilized twice from D_2O to exchange away labile protons and then studied by NMR spectroscopy (final concentration of protein in D_2O: 12 mg/ml).

Elemental Analysis of BDPP. BDPP was assayed for Na^+ and Ca^{2+} via ICP spectrometry, using a Beckmann SpectraSpan Model V6 ICP spectrometer. The purified BDPP sample (50 μg/mL in DD_{water}; 3 mL total volume) was analyzed along with a DD_{water} blank, and a 2 μg/mL bovine serum albumin sample (in DD_{water}) as a control. The Ca^{2+} content in the purified BDPP sample was found to be < 1 ppm per μg protein, and 5-7% by weight Na^+.

pH Titration of BDPP. BDPP was dissolved in 99.96% D_2O (12 mg/mL). The pH was adjusted with either DCl or NaOD in D_2O, and no buffer solution was utilized. All pH units are reported without correction for deuterium isotope effects or sodium ion effects at high

pH. pH titrations were obtained by adding known volumes of NaOD/DCl in 99.96% D_2O to the protein solution (μl volumes and single pass titration were employed to minimize ionic strength effects) and measuring the resulting pH before and after the NMR experiments using an Ingold 6030-02 glass microelectrode and Radiometer PHM83 Autocal pH meter. pK_a values were determined by linear least squares fitting of the data to one or two single ionizations, assuming fast exchange between the conjugate acids and bases. *O*-phospho-L-serine and the dipeptide L-Asp-L-Phe-OCH_3 (both 0.1 M in D_2O) were also titrated under identical conditions. In the context of pH-dependent chemical shift effects, an upfield shift is defined by the symbol $+\Delta\delta$, whereas a downfield shift defined as $-\Delta\delta$.

$^1H/^{31}P$ NMR Spectroscopy. All NMR spectra were recorded at 298° K on a Bruker AM-500 or AMX-500 Spectrometer at a field strength of 11.7 T, with data processing performed on an Aspect 3000 or Aspect X32 computer. One-dimensional ^1H (500.138 MHz) and ^{31}P (202.458 MHz) NMR spectra were typically obtained utilizing 16 K and 8 K data points, respectively; specific acquisition parameters and data handling are described in the Figure legends for each of the spectra. For 1-D ^1H NMR experiments, solvent suppression was accomplished by using presaturation during the relaxation delay. ^{31}P and ^1H NMR spin-lattice T_1 measurements were acquired utilizing the technique of inversion-recovery (180° - τ - 90°-Acquire, using the Bruker INVREC.AUR microprogram). Spin- spin T_2 measurements were obtained using the Carr-Purcell-Meiboom-Gill (CPMG) spin-echo technique [90° - (τ - 180° - τ)$_n$- Acquire, using the Bruker microprogram CPMG.AUR] to average out *J*-modulation and diffusion effects. Both T_1 and T_2 values were then obtained from exponential fitting of the data, using the Bruker T_1 program. For 2-D NMR, ^1H-^1H correlation spectroscopy was obtained using "clean" HOHAHA spin-locking techniques (Bax and Davis, 1985a,b; Griesinger *et al.*, 1988) via the Bruker program CLMLEVPRTP.AUR; ^{31}P-^1H correlation spectroscopy was recorded using the HMBC heteronuclear multiple quantum correlation technique (Bax and Summers, 1986; Live and

Edmondson, 1988) with the Bruker program INVDR2LP.AUR. In the HOHAHA 2-D experiments, 2 K data points were accumulated, with zero-filling in the F2 dimension to create a 2 K x 2 K data matrix using a phase-sensitive (positive and negative levels) squared sine-bell window function in both dimensions with no shift. In the $^{31}P/^{1}H$ HMBC experiment, 4K data points were accumulated and zero-filled to create a 4 K x 4 K data matrix utilizing the same window function as for HOHAHA. Other parameters utilized in the 2-D NMR spectroscopy are described in the Figure legends. Proton chemical shifts are reported relative to external 2,2,2-trimethyl 2-silapentane sulfonic acid (TSS, 0.1 M in D_2O); phosphorus chemical shifts are reported from external EHDP (2-hydroxyethyldiphosphonate) (0.5 M in D_2O, pH 7.0, with the ^{31}P centerpeak resonances at 10.55 and 10.48 ppm relative to 85% H_3PO_4 at 0.0 ppm).

RESULTS

We have studied the conformation of BDPP in D_2O in the absence of Ca^{2+} ion, with emphasis on the amino acid sidechains of Glu, Ser, PSer, and Asp, the likely candidates of coordination sites for Ca^{2+}. With some modifications in existing BDPP purification protocols (Stetler-Stevenson and Veis, 1983; Veis *et al.*, 1986), we have obtained a 155 kD BDPP fraction that gives rise to a single band on SDS-PAGE gels. When this prep is treated with EDTA/EGTA and Chelex-100 to remove Ca^{2+} ions which remain bound to BDPP through the isolation procedure, we obtain following the deionization procedure a preparation of BDPP with a Ca^{2+} level below 1 ppm/μg protein, as determined by ICP spectroscopy. This BDPP protein preparation yields ^{1}H and ^{31}P NMR spectra with narrower linewidths (20-40%) and additional resonances which are not observed in a BDPP protein treated with EDTA alone (data not shown).

^{1}H-NMR Spectroscopy of BDPP. At pH 6.94, the ^{1}H NMR spectra of BDPP (Figure 1) reveals numerous aliphatic resonances in the chemical shift range of 0.0 to 5.0 ppm. Downfield from the solvent

peak, we can identify low-field aromatic resonances in the range of 5.5 to 8.5 ppm whose signal intensities are at least 64-fold weaker than the upfield aliphatic signals. Given that the content of Tyr, Phe, and His in BDPP are very low (1-5 residues/1000 amino acids), this is not unexpected. From the known amino acid content of BDPP, most of the ^{1}H NMR aliphatic resonances originate from 452 Asp and 518 Ser amino acid residues (representing 88% of the total amino acid content)[Veis, 1988]. Other contributions to the aliphatic portion of the ^{1}H NMR spectra are expected to arise from Glu, Gly, Pro, Lys, Thr, Ala, Leu, Ile, and Val (Stetler-Stevenson and Veis, 1983, Rahima and Veis, 1988). Although one would expect a 155 kD protein to generate a ^{1}H NMR spectrum with broad lines resulting from restricted motional fluctuations of the sidechains about the peptide bond in solution, the BDPP protein possesses resonance linewidths at half-height ($\Delta\nu_{1/2}$) as narrow as 5-20 Hz (Table I); however, some signals have $\Delta\nu_{1/2}$ values that are several-fold greater (30 Hz or larger, Table I).

A tentative assignment of the non-exchangeable proton resonances is possible from a comparison of relative peak areas and chemical shifts with the reported amino acid composition of BDPP [Stetler-Stevenson and Veis, 1983; Rahima and Veis, 1988] and known proton chemical shifts for amino acid sidechains [Grob and Kalbitzer, 1988; Jardetzky and Roberts, 1981]. In phosphophoryn, one would expect that Asp and Ser resonances should predominate in terms of relative ^{1}H NMR peak area, and that all other peak areas representing non-Asp and Ser amino acid sidechain resonances should be present in low abundance relative to Asp and Ser. This has been observed in our BDPP samples (Table I). Verification of chemical shift assignments can be made via a 2-D NMR HOHAHA experiment (Figure 2a, b, c) which yield information on coupled spin systems (*J*-scalar couplings) characteristic of each of the various amino acid sidechains. Peak assignments derived in this manner are given in Table I. Because of the potential for significant chemical shift overlap and the lack of an amino acid sequence, it is not possible to give sequence-

specific chemical shift assignments at this time. However, the bulk of the observed chemical shift values are consistent with an extended global conformation for BDPP at this pH. The majority of BDPP Ser residues are phosphorylated (with the exception of 10%)[Stetler-Stevenson and Veis, 1983; Lee *et al.*, 1977]; accordingly, we have identified the peak at 4.087 ppm (Peak N) to be predominantly that of PSer β-CH_2 (see below). The PSer α-CH protons exist within a resonance frequency range of 4.45 to 4.7 ppm [Figure 2a]. The Asp β-CH_2 protons of BDPP exist as multiple resonances from 2.542 to 2.817 ppm (Peaks B-F). Multiple connectivities within these peaks B-F (observed close to the diagonal in Figure 2b) and between peaks B-F and peaks 12-16 result from the *J*-coupling between β and β' diastereotopic protons on the Asp sidechains. A set of α-CH resonances are also observed near 5.70 ppm (Figure 1, Table I), which, owing to their weak intensities, are unassignable at this time.

The protons of Lys residues [$A_2(F_2T_2)$MPX spin system][Basus, 1989] appear on the HOHAHA contour plot as four resonances (Peaks 11, 9, 8, and 7) which exhibit relayed connectivities between all methylene protons and α-CH protons (Figure 2c). The α-CH protons for Lys residues resonate near 4.4 - 4.5 ppm (Figure 2c).

Superimposed on the major resonances for Asp sidechains are weaker signals arising from non-exchangeable sidechain protons of less abundant amino acids (Peaks A, B, Table I, Figure 2a, b). From 4.0 to 3.0 ppm the proton assignments become more ambiguous and complex due to the presence of numerous overlapping spin systems for Asp β, β', Ser β,β'-CH_2 and α-CH (AMX spins), Pro δ-CH_2 ($A_2(T_2)$MPX spin system), and (Gly α, α'-CH protons (AX spin system)(Figure 2a,b). However, based on the connectivity diagram in the HOHAHA contour plot, we can assign predominant spin system(s) (Table I) for each peak [Basus, 1989]. Peaks 21-24 represent α-CH protons for various other amino acid sidechains including Lys, Ser, PSer, and Asp (Table I, Figure 2a,b). In general, the α-CH shifts for Ser and Gly residues in BDPP fall into two categories: those which are

typical of helical regions (resonating upfield of 4.0 ppm) or of β-strand regions (resonating downfield of 4.0 ppm)[Williamson, 1990; Pastore and Saudek, 1990; Wishart *et al.*, 1991].

In the upfield portion of the aliphatic region we have identified resonances belonging to Pro β,β', γ,γ' CH_2 [$A_2(T_2)MPX$ spin system] and Ala β CH_3(Table I , Figure 2c). Multiple connectivities were observed for Peak 4a, 4b, 5, and 6 (Figure 2a,c; Table I); we have assigned these resonances to the Ala β-methyl protons, and the β, γ methylene ring protons of Pro, which in a large polypeptide (100 kD or greater) would be expected to be relatively broadened on account of the restricted motion of the imino ring in the peptide backbone. The α-CH chemical shift values for Pro and Ala residues in BDPP compare well with NMR database values obtained for β-strand conformation [Wishart *et al.*, 1991]. We were unable to identify the particular amino acid sidechains contributing to the resonances at 2.284 (peak 10) and 0.837 ppm (Peaks 3a,b). Peak 10 may arise from the β-methylene protons of Leu or Ile; or conceivably, may arise from a population of Glu γ-CH_2 protons that resonate upfield. Peaks 3a,3b, and 10 appear as broad signals of low intensity ($\Delta\nu_{1/2} \cong 40$–60 Hz), and lack detectable *J*-scalar connectivities. We have presented tentative assignments for Peaks 3 and 10 based solely on the ^1H chemical shift and the known amino acid composition of BDPP (Table I). Peaks 1 and 2 were previously assigned to a small population of ring-current shifted methyl protons of Val, Leu, and Ile, as well as unidentified covalently bound methyl groups [Evans *et al.*, 1992]. It should be noted that the ratio of peak intensities for Peaks 1 and 2 change, depending on the pH of the sample.

Because of the low abundance of aromatic amino acids in BDPP, we are unable at this time to establish definitive assignments for the proton resonances appearing in the region 6.50 to 8.5 ppm (Figure 1, Peaks O through X). These probably arise from His (C_2H, C_4H), Tyr (C_2H, C_3H, C_5H, C_6H), and Phe ring protons. We were also unable to distinguish methylene protons arising from Arg sidechains, perhaps

due to the low abundance of this particular amino acid, and the potential for chemical shift overlap of these resonances with the methylene proton resonances of Lys. Similarly, resonances from the less abundant residues Met, Cys, and Ile (less than 5 residues/1000 amino acids) were not observed in either the ^{1}H NMR 1-D or the HOHAHA experiments.

^{31}P NMR Spectroscopy of BDPP. The ^{31}P NMR spectrum of BDPP (Figure 3) at pH 6.94 appears as a major peak (α) at 2.929 ppm with two shoulders centered 3.124 ppm (β) and 3.333 ppm (χ) downfield to the major peak. Proton decoupling did not improve the lineshape significantly; only a small reduction in linewidth (2-4%) was observed with composite pulse decoupling centered on the PSer β-CH_2 proton frequency; or with broad-band decoupling (3% reduction)(data not shown). The total phosphorus content of BDPP is known to arise from phosphoserine exclusively; there have been no reports of other phosphorylated amino acids such as *o*-phosphotyrosine (PTyr) or *o*-phosphothreonine (PThr)(Stetler-Stevenson and Veis, 1983; Rahima and Veis, 1988). The observed chemical shift of the three BDPP ^{31}P resonances (α,β,χ, Figure 3) lies in the frequency range for PSer at pH 5-8 [Matheis and Whitaker, 1984]. Quantitation of the phosphorus peak area yielded 854.6 moles P_i per mole of BDPP. From the amino acid composition, 685 moles P_i/BDPP would have been anticipated, a discrepancy of +20%. The ratio of peak areas (χ : β : α) is 1 : 1.3 : 2.8 (or 22%/23%/55%), corresponding to 168 : 217 : 478 moles P_i. The $\Delta\nu_{1/2}$ of the ^{31}P NMR signal (α peak) is 34 Hz, which is about 37% greater than the $\Delta\nu_{1/2}$ of the ^{31}P NMR resonance reported in earlier NMR studies of rat 30 kD phosphophoryn (at 36.43 MHz ^{31}P; Lee *et al.*, 1977) and 74% greater than that reported for fetal bovine 94 kD phosphophoryn (at 80.99 MHz ^{31}P; Lee *et al.*, 1983). In these earlier studies, the ^{31}P NMR lineshape for phosphophoryn was noted to be non-Lorentzian, with an unresolved shoulder 0.2 ppm downfield of the main peak.

To verify the amino acid phosphorylated, a $^{1}H/^{31}P$ heteronuclear multiple bond correlation experiment (HMBC) was performed. This experiment measures long-range *J*-scalar coupling through more than one bond (in this case, $^{3}J_{POCH}$). As illustrated in Figure 4, the $^{31}P/^{1}H$ correlation is localized to protons in the chemical shift region of 4.07 ppm, which corresponds to the major PSer β-CH_2 species, peak N (in the ^{1}H spectra, Table I , Figures 1, 2a,b) as expected. We were not able to detect any $^{3}J_{POCH}$ correlations between the phosphate moiety and protons from amino acid sidechains other than Ser.

T_1 and T_2 Relaxation Measurements. Non-selective spin-lattice (T_1) and spin-spin (T_2) relaxation times were obtained for several ^{1}H and ^{31}P NMR BDPP resonances at pH 6.94, 298 K, in D_2O. These values appear in Table I . Although we treated the BDPP sample and all glassware with chelators to remove unwanted paramagnetic ions, we did not purge paramagnetic oxygen (O_2) from our BDPP NMR samples. Thus, the T_1 values in Table I reflect a lower limit due to the presence of dissolved O_2 in in the sample solution. In spite of this, a comparison of the spin-spin and spin-relaxation values for both nuclei reveals several interesting points: (1) For each resonance, T_1 and T_2 values are non-equivalent. For the proton resonances measured, in most cases the T_1 values are 2 to 20 times greater than the corresponding T_2 values obtained (Table I). Similarly, values obtained for ^{31}P T_1 relaxation are 30-50 times greater than the corresponding ^{31}P T_2 relaxation times. Thus, for BDPP sidechains, $T_1 \neq T_2$, which is consistent for a large macromolecule. (2) Proton T_2 relaxation values obtained for the various BDPP resonances show a broad range of variation (15 ± 8 msec to 144.6 ± 7 msec). However, within experimental error, the values for phosphorus T_2 relaxation times are nearly identical (Table I). (3) Given the relationship between linewidth and spin-spin relaxation ($\Delta\nu_{1/2} = \frac{1}{\pi T_2}$), the proton T_2 values obtained in our experiments would give rise to theoretical linewidths that are smaller (by a factor of 2-10) than those observed in

Figure 2. A similar observation is noted for the linewidth and T_2 relaxation of Peak α. Collectively, observations (2) and (3) suggest that the BDPP protein molecule at pH 6.94 possesses amino acid sidechains which have different mobilities.

pH Titration Effects on the ^1H NMR Spectrum of BDPP. A significant number (80% or more) of BDPP amino acid sidechains (Asp, PSer, Glu, His, Lys) are ionizable. Of these, Asp, Glu, and PSer are likely candidates as Ca (II) ion and/or solid phase calcium phosphate coordination sites. His and Lys may also be involved in possible salt-bridging within the protein structure (intrachain), or between BDPP molecules in solution (interchain). In addition, at neutral pH, the sidechains of Asp, Glu, and PSer could exist as either fully or partially deprotonated species, and charge repulsion could contribute to the secondary and tertiary folding of the Ca^{2+}-free protein in solution. To arrive at an understanding of these complex effects, we have constructed a pH profile of ^1H NMR chemical shifts for the PSer, Asp and Lys amino acid residues (Figure 5a,b), from which pK_a values were derived (Table II). Because of chemical shift overlap we are not able at this time to obtain pK_a's for individual amino acid sidechains in BDPP; however, we can determine pK_a's for groups of amino acids that contribute to a given peak. The β-CH_2 protons of PSer (peak N) exhibit a downfield shift (-Δδ) of 0.16 ppm) over a pH range of 11 units with inflection points at 2.51 and 6.88 (Figure 5a), reflecting the two pK_a's of the monophosphate ester phosphorus to which these protons are $^3J_{POCH}$ scalar coupled (Table II). For comparison, a similar titration of *o*-phospho-L-serine over the pH range from 1.82 to 9.55 (Table II) reveals inflection points at pH 2.53 (occurring over a pH range of 1 unit) and 5.82 (occurring over a pH range of 2 units) for the β-methylene protons. pK_a's previously been reported for *o*-phospho-L-serine and PSer residues in phosphophoryn are given in Table II [Lee *et al.*, 1977; Cookson *et al.*, 1980; Data for Biochemical Research, 1969; Vogel, 1989]

The variation of the spectrum in the region of the Asp β-CH_2 resonances with pH is quite complex (peaks B-F, Figure 5a). At pH 1.82, the Asp residues of BDPP exist in two major spin populations (Figure 5a) centered at 2.901 ppm and 3.009 ppm. These residues separate into at least 6 distinct spin populations above pH 6.53 (Figure 5a) . With increasing pH, these Asp resonances shift upfield. The chemical shift-pH profile reveals quite different pK_a's for various groupings of β-CH_2 resonances (Table II). In addition, peaks B, C, and E show an additional inflection in their titration curves (Figure 5a, Table II). It should be recognized that part of the spin population of Peak B is comprised of Ser β and Gly α protons as well. For comparison, we also studied the pH titration behavior of the dipeptide L-Asp-L-Phe-OCH_3. Here, the Asp β, β'-CH_2 resonances titrate with a pK_a of 2.75 due to the ionization of the γ-carboxyl group. The pK_a's reported for L-aspartic acid, Asp residues in phosphophoryn, and poly-L-Asp are given in Table II.

The ε-CH_2 protons of Lys exhibit only a small pH-dependent chemical shift (Figure 5b) in the pH range over which the deprotonation of the primary amine nitrogen is expected (pH 10-11). This reduced pH effect on chemical shift may be indicative of salt-bridging between the positively- charged ε primary amine nitrogen and negatively charged residues, e.g., PSer, Asp, Glu, within the polypeptide. A pH dependence of the β-CH_2 protons of Lys was also observed. Here, two distinct β-CH_2 populations (a and b) exist below pH 7.2, but they become equivalent above this pH. (Figure 5b). Lys β-CH_2 peak a experiences an upfield shift ($\Delta\delta$ = +0.015 ppm) over a pH range from 4.34 to 6.88. Concurrent with this, Lys β-CH_2 peak b shifts downfield ($\Delta\delta$ = -0.017 ppm) over the same pH range. The chemical shift differences between these populations is maximal at pH 1.82 ($\Delta\delta$ = 0.06 ppm). The δ-CH_2 proton resonances of Lys does not show a pH dependence over the pH range examined.

pH Titration Effects on the ^{31}P NMR Spectrum of BDPP. The effect of pH on the ^{31}P chemical shift of the PSer residues in BDPP is shown

in Figure 6. The three ^{31}P PSer resonances (peaks α, β, and χ) give three slightly separate titration curves, each of which exhibits a downfield shift ($\Delta\delta$ = -4.2 ppm) as pH increases from 4 to 9; a major inflection point occurs in the pH region of 6.88 to 7.1. Below pH 6.0 and above pH 8.2, the three PSer ^{31}P resonances become chemical shift equivalent. A second inflection point (weak) is noted at pH 2.52; through this transition point, PSer phosphorus resonances experience a downfield shift of 0.1 ppm. The ^{31}P NMR pH titration of *o*-phospho-L-serine (Table II) reveals two inflection points: a minor one at pH 2.62 and the major inflection at 6.01. The pK_a's for the three PSer resonances of BDPP and *o*-phospho-L-serine are given in Table II, along with previously reported pK_a values [Lee *et al.*, 1977; Cookson *et al.*, 1980; Vogel. 1989] for phosphophoryn PSer sidechains and *o*-phospho-L-serine.

DISCUSSION

From the ^{1}H 500 MHz and ^{31}P 202.45 MHz NMR data compiled in this investigation, it appears that the conformation of BDPP is very sensitive to changes in the ionization of its amino acid sidechains. In previous $^{1}H/^{31}P$ NMR studies of the 30 kD rat phosphophoryn conducted at lower field [Cookson *et al.*, 1980] and the initial CD/IR/NMR experiments conducted on the 30 kD BDPP fragment [Lee *et al.*, 1977], the general consensus has been that phosphophoryn assumes an extended, random coil conformation in solution, with aspartate and phosphoserine peptide residues experiencing similar environments and exhibiting a non-complex titration behavior. A similar conclusion was drawn by Lee *et al.* (1983) from ^{31}P NMR studies conducted at lower magnetic field on the 94 kD fetal BDPP protein. With the improvement in chemical shift resolution afforded by this present study at higher field, and the application of two-dimensional NMR experiments, we believe that our data points to a structure and behavior of the polyelectrolyte regions in Ca^{2+}-free BDPP that is more complex than was previously recognized.

Electrostatic Effects on Asp and PSer Sidechains . The pH titration of BDPP at low ionic strength reveals at least six distinct Asp resonances at neutral pH (peaks B-F, Figures 1, 2a, b, 5) and three distinct phosphoserine resonances (Figure 3). We were unable to directly examine pH titration effects for Glu residues, primarily due to the chemical shift overlap of Glu β,γ-CH_2 with the predominant Asp β-CH_2 resonances. The chemical shifts of the β, β-CH_2 protons for Asp and PSer and the ^{31}P resonance of PSer are sensitive to short-range electrostatic effects resulting from protonation and deprotonation of carboxyl and monophosphate groups. In addition, long range electrostatic interactions will also influence the position of these resonances due to effects from charged groups with different ionization states that are in close proximity to one another (i.e., through space). Long range influences on pK_a's are usually manifested in one of two ways: 1) by a shift in the observed pK_a of the affected charged group towards a value representing the pK_a of the influencing group if the differences in pK_a's are less than two units, or, 2) by the presence of an additional inflection point in the titration curve for one of the ionizing groups if their pK_a's differ by at least two units [Jardetzky and Roberts, 1981].

From our pH titration data (Table II, Figure 5a), it is clear from the pK_a's of the Asp residues that polyelectrolyte effects are present. The observed pK_a's of these sidechains are all shifted to higher values than expected for an isolated sidechain group. The pH titration profile in the region of the inflection point is also broad, spanning 2-4 pH units (Figure 5a, Peaks B-F). Superimposed upon this general pK_a shift for Asp residues is the appearance of secondary titration inflection points on the protonation and deprotonation of select Asp residues (Peaks B, C, and E Table II, Figure 5a); these values, ranging from 6.17 to 6.95, are within ± 0.8 units of the second pK_a for PSer (Table II). We also note that those Asp residues that exhibit secondary inflection points also possess broader titration profiles (peaks B,C,E, inflection points occurring over a pH range of 4-5 units, nearly 100% greater than peaks D,F, Figure 5a) than those Asp

residues which have only a single inflection point. The presence of secondary inflection points was also noted in the ^{1}H NMR study on rat 30 kD phosphophoryn, in which 80% of the total Asp residues titrated with a pK_a of 4.2, and the remaining 20% with a pK_a of 3.2 [Cookson *et al.*, 1980]. Our conclusion is that the observed pK_a shift and broad titration profile of the Asp residues reflects electrostatic interactions of the Asp carboxylate group with other negatively charged sidechains (either Asp, PSer or Glu), and that the presence of a secondary inflection point for > 50% of the Asp residues results from the second ionization of neighboring PSer groups; this implies that the remaining Asp residues are only influenced by carboxyl group amino acid sidechains.

For the PSer residues, the ^{31}P NMR titration data revealed three distinct pK_a's of 6.97, 6.95, and 6.93 for the second ionization of the monophosphate group (Figure 6, Table II). This phenomenon was not observed in the ^{1}H NMR pH titration data (Figure 5a), presumably due to the β,β'-methylene protons being three-bond distant from the phosphorus nucleus and hence lower sensitivity of the ^{1}H shift compared to the ^{31}P NMR towards ionization and conformational change. Both the ^{1}H and ^{31}P NMR titration profile for PSer residues in BDPP are broader than that obtained for *o*-phospho-L-serine (^{1}H titration by 1-2 pH units; ^{31}P NMR titration by 1 unit). At high pH (pH > 8) all three PSer resonances (α,β,χ) become equivalent (Figure 6). Long-range electrostatic effects from neighboring charged sidechains would be minimal at high pH due to charge repulsion forcing sidechains away from one another. At low pH (pH < 5) such charge effects should be minimal, due to partial hydrogen ion neutralization of charge on carboxyl groups and monophosphate esters. Thus, ^{31}P chemical shift equivalency among the various PSer sidechains is again expected (Figure 6). In the pH range of 5.5-7.5, however, carboxyl groups on BDPP become fully deprotonated (Figure 5a), whereas monophosphate esters are either monobasic or dibasic in form. Accordingly, the protonation state and ^{31}P chemical shift of these PSer sidechains can still be affected electrostatically by these

neighboring charged groups. This would account for ^{31}P chemical shift inequivalency for 45% of the PSer residues in BDPP in this pH range. On this basis, at neutral pH we would assign peak α (Figure 3) to a population of PSer residues (55%) in BDPP which are localized in regions that do not experience long-range electrostatic interactions. Peaks β and χ, however, would be expected to arise from PSer residues (45%) that are located in regions which are polyelectrolyte clusters, consisting of Asp, Glu, and/or PSer residues. It is possible that PSer residues which contribute to peak χ (22%) experience a greater negative charge density than those PSer residues that comprise peak β (23%).

Arrangement of Asp and PSer Sidechains in BDPP. From the data presented above, we conclude that BDPP contains polyelectrolyte clusters that are homologous (carboxylate or monophosphate ester) and heterologous (carboxylate and monophosphate ester) in their anionic sidechain composition (*vide infra*). These regions may represent primary amino acid sequences that selectively cluster certain residues together and allow for the heterogeneity in Asp pK_a's and in the ^{31}P PSer chemical shifts at neutral pH in BDPP. Asp and PSer residues may exist in sequences that are either repeating homopolymer regions [$(Asp)_n$ or $(PSer)_n$], or may exist as separate amino acids that are interdispersed among non-ionizing amino acids in the peptide sequence. Similarly, there are PSer and Asp residues which reside in heteropolymer clusters. These heteropolymer cluster regions may vary in sequence (e.g., Asp-PSer-Asp, PSer-Asp-Asp-PSer, etc.), and this heterogeneity in composition may well explain the broad pH titration profile (Figure 5a) and variation in pK_a shift (Table II) observed.

Evidence for Conformational Heterogeneity BDPP. Our HOHAHA experiments at neutral pH (Figure 2a) indicate that both Ser and Gly residues in BDPP exhibit α-CH shifts that are typical of either helical regions or b-strand regions (Figure 2a,b)[Williamson, 1990; Pastore and Saudek, 1990; Wishart *et al.*, 1991]. These observations suggest

that in the divalent cation-depleted state BDPP is not entirely "random coil," but may actually possess defined secondary structures in localized regions of the protein sequence. A further exploration of the secondary and tertiary structure of BDPP polyelectrolyte regions will be presented in the third manuscript of this series [Evans and Chan, 1992c].

The spin-spin and spin-lattice relaxation data (Table I) indicate that the BDPP molecule possesses heterogeneous sidechain or backbone motions, as witnessed by the variation in T_2 and T_1 values for proton NMR resonances. However, the similarity of the ^{31}P NMR relaxation values (within experimental error) indicates that in different regions of BDPP the PSer sidechain motions (for resonances α,β,χ) may be very similar. The Asp β, β'-CH_2 proton T_1 and T_2 measurements are also within experimental error (Table I), implying that Asp sidechains (resonances B-F) experience similar sidechain motions in solution. Of particular interest are the T_1 and T_2 values obtained for neutral amino acid sidechain or backbone protons (Table I, Peaks 1-10; Peaks 18-H). These values suggest that the BDPP protein molecule possesses regions which differ from the polyelectrolyte cluster regions in peptide backbone and/or sidechain mobility. This observation is consistent with a protein molecule that may be conformationally heterogeneous.

Evidence for Salt-Bridging in BDPP. In BDPP, the pH titration of Lys residues indicates that the ε-CH_2 protons, which normally would show a titration behavior reflecting the primary NH_2 deprotonation around pH 10-11, do not exhibit this behavior (Figure 5b). It is conceivable that Lys ε-NH_3^+ forms ion pairs with some of the anionic sidechains of BDPP in solution; it has been shown that salt bridging of this type in proteins results in a loss of titration behavior for these ε-CH_2 protons [Jardetzky and Roberts, 1981] due to the protonation stabilization from a negatively charged species such as Asp, Glu, or PSer. The functional significance of Lys residues in BDPP and their location in the peptide sequence is not known at present.

Summary

This study provides NMR data in support of a structure for BDPP that is comprised of both homologous and heterologous cluster regions of Asp and PSer residues. This result is hardly surprising, given the predominance of Asp and PSer in the polypeptide sequence. The repeat structure $(\text{-Asp-PSer-Asp-PSer-})_n$ was originally proposed by Lee *et al.* (1977) for 30 kD BDPP based on the given amino acid stoichiometry between Asp and PSer. Cookson *et al.* [1980] also postulated the existence of anionic clusters in phosphophoryn (rat), based on the pH titration cooperativity between PSer and Asp residues. Recent experimental data on the peptide sequence for BDPP and rat incisor α-phosphosphoryn (90 kD) has revealed that repeating blocks of $(\text{-Asp-PSer-})_n$, $(\text{-Asp-Asp-})_n$, and $(\text{-PSer-Ser-})_n$ comprise a significant portion of the primary structure (Sabsay *et al.*, 1991). The Asp-Ser and Asp-Asp motif has also been found in the N-terminal 23 amino acid sequence of BDPP (Evans and Chan, 1991). The sequence length of these regions or the number of heterologous and homologous anionic clusters that comprise BDPP or other phosphophoryns are not known at present. However, as shown in the following manuscript, these anionic clusters influence the conformational folding of the BDPP molecule in solution.

Acknowledgements

We thank Dr. Arthur Veis for discussions regarding the preliminary amino acid sequence of rat α-phosphophoryn, and Dr. Heinz Lowenstam for helpful discussions and reading of the manuscript. The technical expertise of John Racs in CZE separation is also acknowledged.

References.

Addadi, L., Moradian, J., Shay, E., Maroudas, N. G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* 84 2732-2736.

Basus, V. J. (1989) *Methods in Enzymology* 177 132-141.

Bax, A., and Davis, D. G. (1985a) *J. Am. Chem. Soc.* 107 2820-2821.

Bax, A., and Davis, D. G. (1985b) *J. Mag. Resonance* 65 355-360.

Bax, A., and Summers, M. F. (1986) *J. Am. Chem. Soc.* 108 2093-2094.

Berger, A., and Katchalski, E. (1951) *J. Am. Chem. Soc.* 73 4084-4088.

Cookson, D. J., Levine, B. A., Williams, R. J. P., Jontell, M., Linde, A., and de Bernard, B. (1980)*Eur. J. Biochem.* 110 273-278.

Data for Biochemical Research (1969) (Eds. Dawson, R. M. C., Elliot, D. C., Elliot, W. H., and Jones, K. M.) Oxford University Press, New York, 2nd edition, pp. 13, 120.

Evans, J. S., and Chan, S. I. (1991) in *Frontiers in Peptide and Protein Structure/Whistler Symposia,* in press.

Griesinger, C., Otting, G., Wuthrich, K., and Ernst, R.R. (1988) *J. Am. Chem. Soc.* 110 7870-7872.

Grob, K-H., and Kalbitzer, H. R. (1988) *J. Mag. Resonance* 76 87-99.

Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology* (eds. by Jardetzky and Roberts) Academic Press, New York, New York, pp. 143-186, 227-231, 234-270, 448-465.

Kuboki, Y., Fujisawa, R., Aoyama, K., Sasaki, S. (1979) *J. Dent. Res.* 58 1926-1932.

Laemmli, U. K. (1970) *Nature* (London) 227 680-685

Lee, S. L., Glonek, T., and Glimcher, M. J. (1983) *Calcif. Tiss. Int.* 35 815-818.

Lee, S. L., Veis, A., and Glonek, T. (1977) *Biochemistry* 16 (13) 2971-2979.

Live, D. H., and Edmondson, D. E. (1988) *J. Am. Chem. Soc.* 110 4468-4470.

Lowenstam, H. A. (1981) *Science* 211 1126-1131.

Lowenstam, H. A., and Weiner, S. (1985) *Science* 227 51-53.

Mann, S. (1988) *Nature* 332 119-124.

Marsh, M. E. (1989) *Biochemistry* 28 339-345.

Marsh, M. E., and Sass, R. L. (1984) *Biochemistry* 23 1448-1456.

Matheis, G., and Whitaker, J. R. (1984) *Int. J. Biochemistry* 16 (8) 867-873.

Pastore, A., and Saudek, V. (1990) *J. Mag. Res.* 90 165-176.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* 42 104-112.

Sabsay, B., Lechner, J., Stetler-Stevenson, W. G., and Veis, A. (1991) *Biochem. J.* 276, *in press.*

Stetler-Stevenson, W. G., and Veis, A. (1983) *Biochemistry* 22 4326-4335.

Stetler-Stevenson, W. G., and Veis, A. (1986) *Calcif. Tiss. Int.* 38 135-141.

Stetler-Stevenson, W. G., and Veis, A. (1987) *Calcif. Tiss. Int.* 40 97-102.

Veis, A. (1988) in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P., eds) VCH Publishers, New York, New York 199-216.

Veis, D. J., Albinger, T. M., Clohisy, J., Rahima, M., Sabsay, B., and Veis, A. (1986) *J. Exptl. Zool.* 240 35-46.

Veis, A., Spector, A. R.,. and Zamoscianyk, H. Z. (1972) *Biochem. Biophys. Acta* 257 404-413.

Vogel, H. J., (1989) *Methods in Enzymology* 177 263-282.

Weiner, S. (1983) *Biochemistry* 22 4139-4145.

Williamson, M.P. (1990) *Biopolymers* 29 1423-1431.

Wishart, D.S., Sykes, B.D., and Richards, F.M. (1991) *J. Mol. Biol.* 222 311-333.

Zanetti, M., de Bernard, B., Jontell, M., and Linde, A. (1981) *Eur. J. Biochem.* 113 541-545.

Figure 1. *^{1}H 500.138 MHz NMR Spectrum of BDPP (12 mg/ml) in 99.96% D_2O, pH 6.94.*

Acquisition parameters include a spectral window of 10 kHz, a pulse angle of 45^{o} (5 μsec), receiver gain of 20 dB, recovery delay of 1 sec, 32 K data points (0.492 Hz/pt), a Lorentzian apodization function of 1 Hz, and 445 scans. A homonuclear gated presaturation pulse (0.5 watts, decoupler offset centered on the HOD peak) is applied during the recovery phase (utilizing 4 dummy scans) to achieve solvent suppression. Low field region of spectrum is expansion plotted at a factor of 64 x that of the primary spectrum. Proton chemical shifts are referenced from TSS in D_2O.

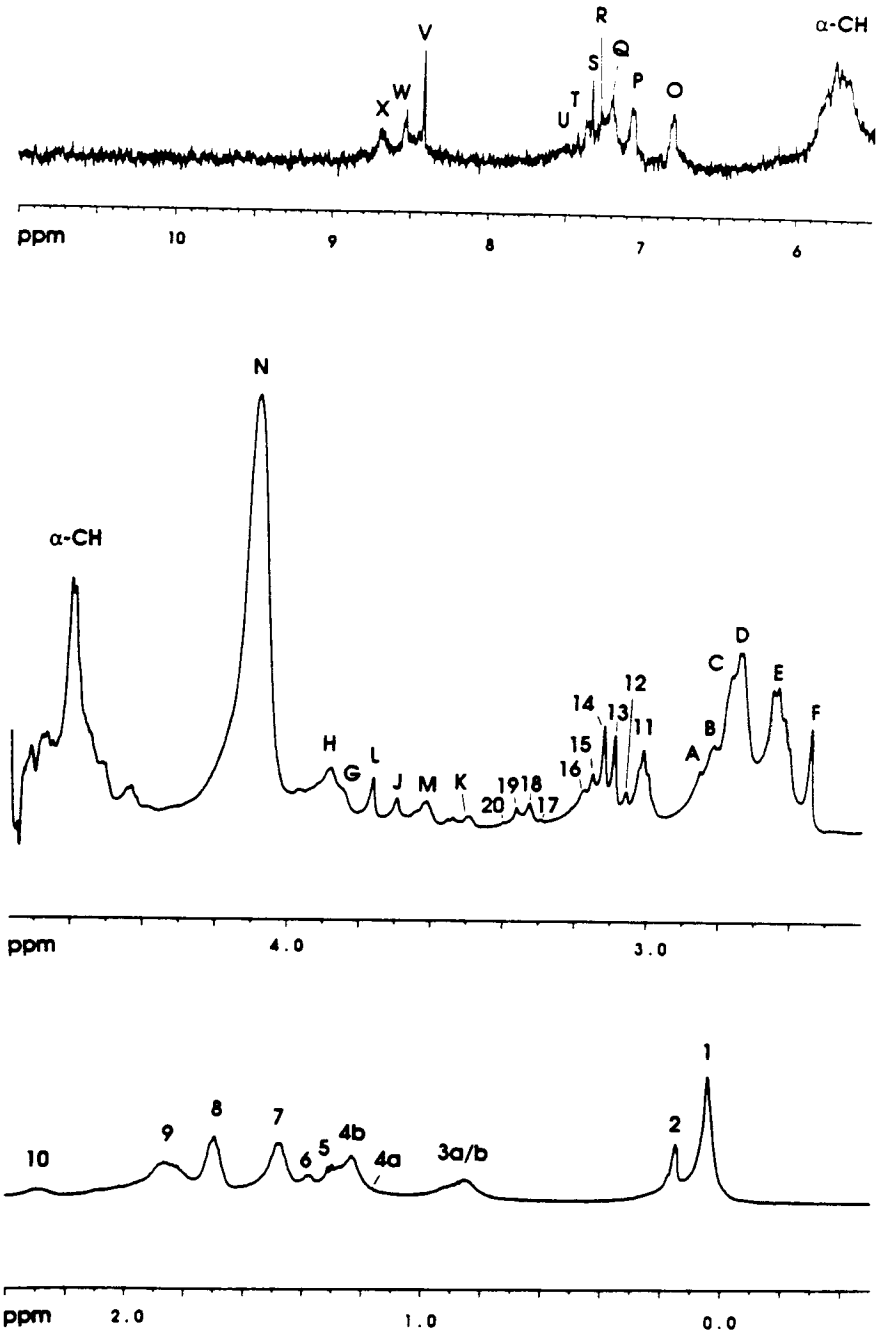

α-CH
R
V
S
Q
X
W
T
P
O
U
ppm
10
9
8
7
6
N
α-CH
D
C
12
E
14
13
F
11
B
15
H
L
A
16
G
J
M
K
19
18
20
17
ppm
4.0
3.0
1
9
8
7
2
10
4b
6
5
4a
3a/b
ppm
2.0
1.0
0.0

Figure 2: *^1H - ^1H"Clean" HOHAHA 2-D Homonuclear Spin-Lock Correlation Experiment (TPPI Phase Sensitive Contour Plot) of BDPP in 99.96% D_2O, pH 6.94.*

Details of the spin-lock correlation experiment are described in Materials and Methods. BDPP sample is identical to that used in Figure 2. The 2-D experiment utilizes 2 dummy scans and 240 scans per experiment, and the data represents the sum of 512 experiments. A 2 K x 2 K data matrix (4.883 Hz/pt in both dimensions) is created with zero filling in the F2 dimension, using a squared sinebell window function in the phase-sensitive mode. The 90° ^1H transmitter pulse width is 11 μsec; the decoupler 90° pulse width is 37 μsec and the 180° pulse width is 74 μsec (at a decoupler setting of 15 dB attenuation). Spectral window for F1 is 2.5 kHz, and 5.0 kHz for F2. A spin-lock time of 74 msec is used. **(a)**,**(b)**,**(c)**: Expansion of high field region (aliphatic) of HOHAHA contour plot. In spectrum **(b)**, the contours are plotted several levels above the noise to resolve the β, β' diastereotopic protons near the diagonal. ^1H chemical shifts are reported from 0.1 M TSS in D_2O.

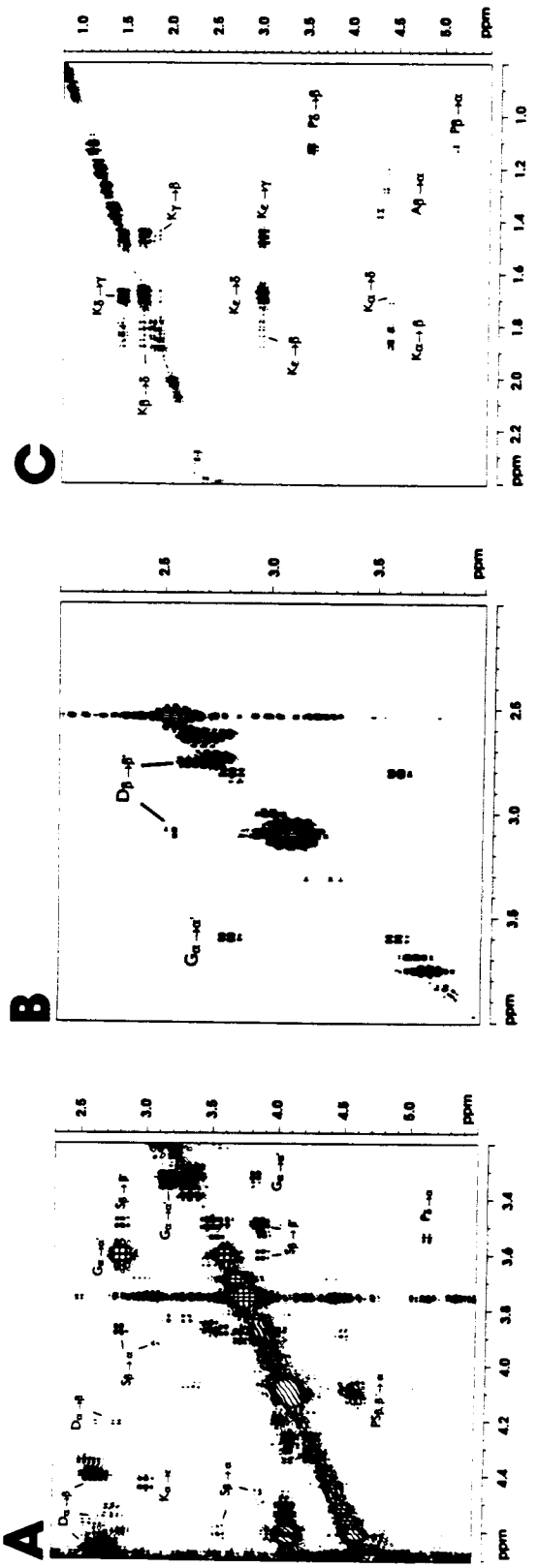

A
B
C
ppm

Table I : Peaks, identified by either a number or letter code in Figure 3a, b, are presented with the centerpeak chemical shift (relative to TSS for ^{1}H, EHDP for ^{31}P). *^{1}H NMR assignments are determined from the HOHAHA spectrum (Figure 2a,b,c) on the basis of proton chemical shifts and *J*-connectivities for various amino acid spin systems. A square bracket [ ], followed by a question mark (?), indicates that connectivities may or may not have been observed for these peaks, and that the assignments, based on chemical shift, are tentative. A parenthesis enclosing an assigned proton spin indicates that these protons are subordinate in number compared to the major spin system. Relaxation times were determined as described in Materials and Methods. Each value was obtained from exponential line fitting to 12 tau points. S.D.: standard deviation computed from best fit. ‡Peak areas are determined by first applying a baseline correction to the spectrum, then assigning an overall area value (A=1000 for ^{1}H, 100 for ^{31}P) for the entire BDPP spectrum ($^{1}H/^{31}P$ spectral window=10 kHz). Peak areas do not represent absolute quantitation of resonance intensities. [a]Small peaks are also noted at 2.018, 2.0923 ppm. [b]Peaks A-F include diastereotopic β, β' protons; the chemical shift distribution of these resonances have not been determined. [c]Peak area and linewidth include those of shoulder regions found at 2.623 and 2.652 ppm as well. [d]Peak area and linewidth values represent peaks A-C. [e]Shoulder regions at 2.998 and 3.019 ppm are also observed. [f]Peak area given is for the entire multiplet (peaks 12-16 inclusive). [g]Peak area given is for the entire multiplet (peaks 17-20). [h]Peaks G and H may include Peak I at this pH. The chemical shift of Peak I at this pH is not known. [i]In addition to major peaks 21-25, minor peaks or shoulder regions are also observed at 4.447, 4.585, 4.683, 4.657, and 4.713 ppm. [j]Linewidth includes peaks α, β, χ. N.D.: not determined. $\Delta\nu_{1/2}$: Linewidth at half-height, determined from either Figure 2 or Figure 4. †^{31}P assignments based on 2-D $^{31}P/^{1}H$ HMBC experiment.

Table I

^{1}H NMR Assignments and Data for BDPP at 298 K, pH 6.94*

Peak	δ, ppm	Assignment	T_1 (± SD)	T_2 (± SD)	$\Delta\nu_{1/2}$ (Hz)	Area‡
1	0.032	V, L, I -CH_3, -CH_3?	613 (76)	145 (7)	21.7	22.2
2	0.155	V, L, I -CH_3, -CH_3?	775 (36)	61 (10)	12.3	14.5
3a,b	0.837	L,V,I,A CH_3, P-CH_2 ?	259 (19)	39 (9)	65.0	14.4
4a	1.135	P β (γ?)	N.D.	N.D.	N.D.	N.D.
4b	1.225	P β (γ?)	258 (50)	41 (6)	65.0	21.9
5	1.289	A β	N.D.	N.D.	6.0	N.D.
6	1.385	A β	257 (62)	32 (24)	16.7	5.6
7	1.466	K γ	307 (71)	28 (14)	34.0	22.6
8	1.690	K δ	357 (63)	74 (20)	28.6	17.4
9[a]	1.849	K β	341 (64)	15 (8)	78.8	33.2
10	2.284	L, I, E -CH_2 ?	154 (29)	25 (16)	46.3	8.4
F[b]	2.542	D β, β'	261 (17)	64 (7)	6.9	7.2
E[c]	2.627	D β, β'	422 (9)	75 (6)	29.0	31.5
D	2.681	D β, β'	433 (14)	137 (29)		
C[d]	2.701	D β, β'	N.D.	N.D.	34.0	61.8
B	2.817	D β, β', S β, β', G α, α'	N.D.	N.D.		
A	2.855	S β, β' G, α, α'	N.D.	N.D.		
11[e]	3.006	K ε	487 (50)	61 (5)	23.1	15.5
12[f]	3.056	D β	286 (10)	117 (35)	5.4	34.9
13	3.085	D β, S β, G α	326 (48)	139 (19)	5.9	
14	3.114	D β, S β, G α	395 (50)	143 (22)	5.9	
15	3.144	D β, S β, G α	442 (58)	143 (22)	5.4	
16	3.178	D β, S β, G α	329 (30)	126 (23)	7.9	
17	3.295	S β, β'	N.D.	N.D.	N.D.	
18[g]	3.328	S β, β'	418 (29)	79 (17)	12.3	11.3
19	3.364	S β, β'; G α, α'	534 (54)	83 (16)	10.3	
20	3.399	S β, β'; G α, α'	N.D.	N.D.	N.D.	
K	3.521	P δ	420 (29)	60 (14)	15.3	4.2
M	3.612	G α	309 (34)	65 (21)	26.6	8.1
J	3.697	G α; S β, α	349 (45)	103 (29)	10.8	6.3
L	3.763	S β, β', α; G α	339 (42)	86 (18)	9.8	8.3
G[h]	3.851	S β, β'; G α; α–CH	N.D.	N.D.	N.D.	N.D.
H	3.881	S β, β'; α-CH	434 (66)	54 (17)	48.7	30.6
N	4.087	PSer β, β'	523 (75)	32 (15)	35.9	151
21[i]	4.436	α-CH	N.D.	N.D.	18.7	15.7
22	4.597	α-CH	500 (91)	35 (15)	19.7	66.0
23	4.670	α-CH	N.D.	N.D.	N.D.	17.2
24	4.701	α-CH	N.D.	N.D.	N.D.	N.D.
25	5.748	α-CH	N.D.	N.D.	N.D.	N.D.

Table I

^{31}P NMR Assignments and Data for BDPP at 298 K, pH 6.94†

Peak	δ, ppm	Assignment	T_1 (± SD)	T_2 (± SD)	$\Delta\nu_{1/2}$ (Hz)	Area
αJ	2.929	PSer P	1530 (10)	38 (5)	34	55.9
β	3.124	PSer P	1486 (30)	31 (7)	N.D.	25.4
χ	3.333	PSer P	1330 (38)	44 (10)	N.D.	19.7

Table II: pK_a's are determined by non-linear least squares fit to chemical shift vs. pH data obtained for BDPP sidechain resonances at low ionic strength. **Boldface** numbers indicate secondary inflection points on titration curves. **a**Determined from $^{1}H/^{31}P$ NMR pH titration experiments on *o*-phospho-L-serine in D_2O in this study. **b***o*-phospho-L-serine, Data for Biochemical Research, p. 122, 1969. **c**PSer residues in 30 kD rat phosphophoryn, determined by ^{1}H NMR, Cookson *et al.* 1980. **d** *o*-phospho-L-serine, Vogel, 1989. **e**PSer residues in 35 kD bovine phosphophoryn fragment, determined by ^{31}P NMR, Lee *et al*, 1977. **f**PSer residues in 30 kD rat phosphophoryn, determined by ^{31}P NMR, Cookson *et al.*, 1980. **g**L-Asp-L-Phe-OCH_3 Asp β-CH_2, determined by ^{1}H NMR (this study). **h**Asp residues in 30 kD rat phosphophoryn, determined by ^{1}H NMR, Cookson *et al.*, 1980. **i**L-aspartic acid, Data for Biochemical Research, p. 13, 1969. **j**Obtained for poly-L-aspartic acid in water, Berger *et al.*, 1951.

Table II

Titration Behavior of BDPP Residues in D_2O at 298 K

Peak	Sidechain	pK_a(± S.D.)	Reported pK_a	Peak	Sidechain	pK_a(± S.D.)	Reported pK_a
N	PSer β-CH_2	2.51 (0.21)	2.80[a]; 2.65[b]	**B**	Asp β-CH_2	4.14 (0.15)	2.75[g]; 3.2, 4.2[h]; 1.99[i]; 5.2[j]
		6.88 (0.18)	6.06[a]; 5.91[b]; 6.6[c]			**6.17** (0.27)	
α	PSer P	2.50 (0.10)	2.62[a]	**C**	Asp β-CH_2	5.23 (0.07)	
		6.97 (0.21)	6.01[a]; 5.8[d]; 6.8[e], 7.2[f]			**6.95** (0.18)	
β	PSer P	6.93 (0.17)		**D**	Asp β-CH_2	5.42 (0.20)	
χ	PSer P	6.95 (0.20)		**E**	Asp β-CH_2	5.18 (0.16) **6.95** (0.22)	
				F	Asp β-CH_2	4.13 (0.27)	

Figure 3: *^{31}P NMR 202.458 MHz NMR Spectrum of BDPP at pH 6.94 in 99.96% D_2O.*

Same BDPP sample as for Figures 2,3. Parameters for ^{31}P NMR experiment include a spectral window of 10 kHz, 2000 aquisitions, 16 K data points (1.221 Hz/pt), 45^0 pulse (8.3 μsec), receiver gain of 80 dB, a Lorentzian apodization function of 5 Hz. No proton decoupling is employed. ^{31}P chemical shifts are referenced from 0.5 M EHDP, pH 7.0 in 99.96% D_2O, with the EHDP ^{31}P centerpeak resonances at 10.55 and 10.48 ppm relative to 85% H_3PO_4 at 0.0 ppm).

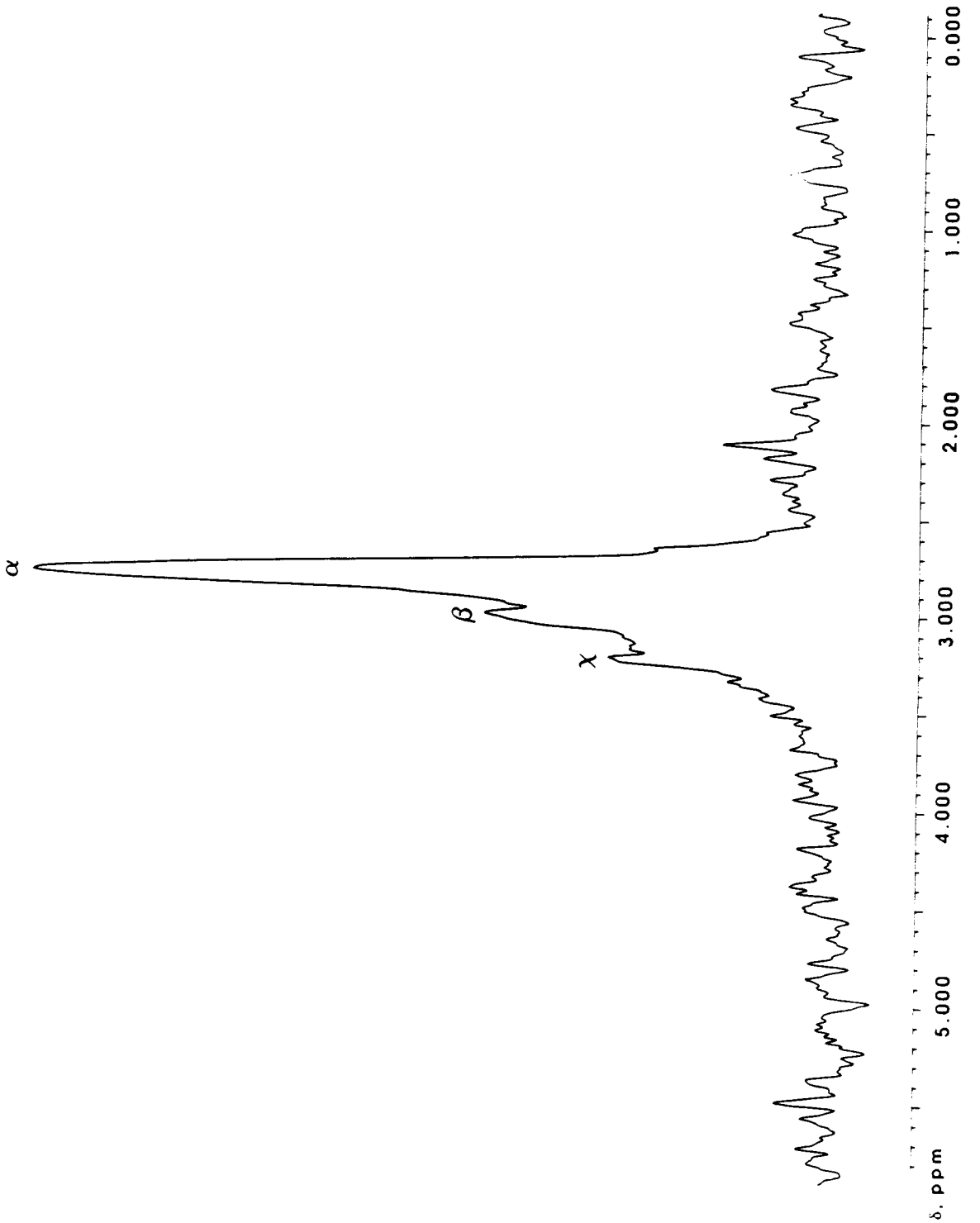

α
β
χ
0.000
1.000
2.000
3.000
4.000
5.000
δ, ppm

Figure 4: *$^{1}H/^{31}P$ Heteronuclear Multiple Quantum Coherence Correlation Experiment (Magnitude Mode Contour Plot) of BDPP in 99.96% D_2O, pH 6.94.*

BDPP sample is identical to that utilized in Figures 2, 3, and 4. Acquisition parameters include a relaxation delay of 2 sec, a ^{1}H 90^{o} and 180^{o} pulses of 13 μsec and 26 μsec, respectively (at a decoupler power setting of 10 w); a ^{31}P 90^{o} pulse of 16.5 μsec, a spectral window of 5 kHz (^{1}H) and 2.5 kHz (^{31}P), an incremental delay of 1 μsec, 4 dummy scans, and a *J*-pass filter delay of 83 msec. The data represent the sum of 512 experiments (2 K data points, 160 scans/experiment). A 4 K x 4 K data matrix was created, using a squared sine-bell window function in both dimensions, with zero-filling in the F2 dimension. ^{1}H chemical shifts (ppm) are referenced from TSS and ^{31}P chemical shifts are referenced from 0.5 M EHDP, pH 7.0 in 99.96% D_2O, with the EHDP ^{31}P centerpeak resonances at 10.55 and 10.48 ppm relative to 85% H_3PO_4 at 0.0 ppm).

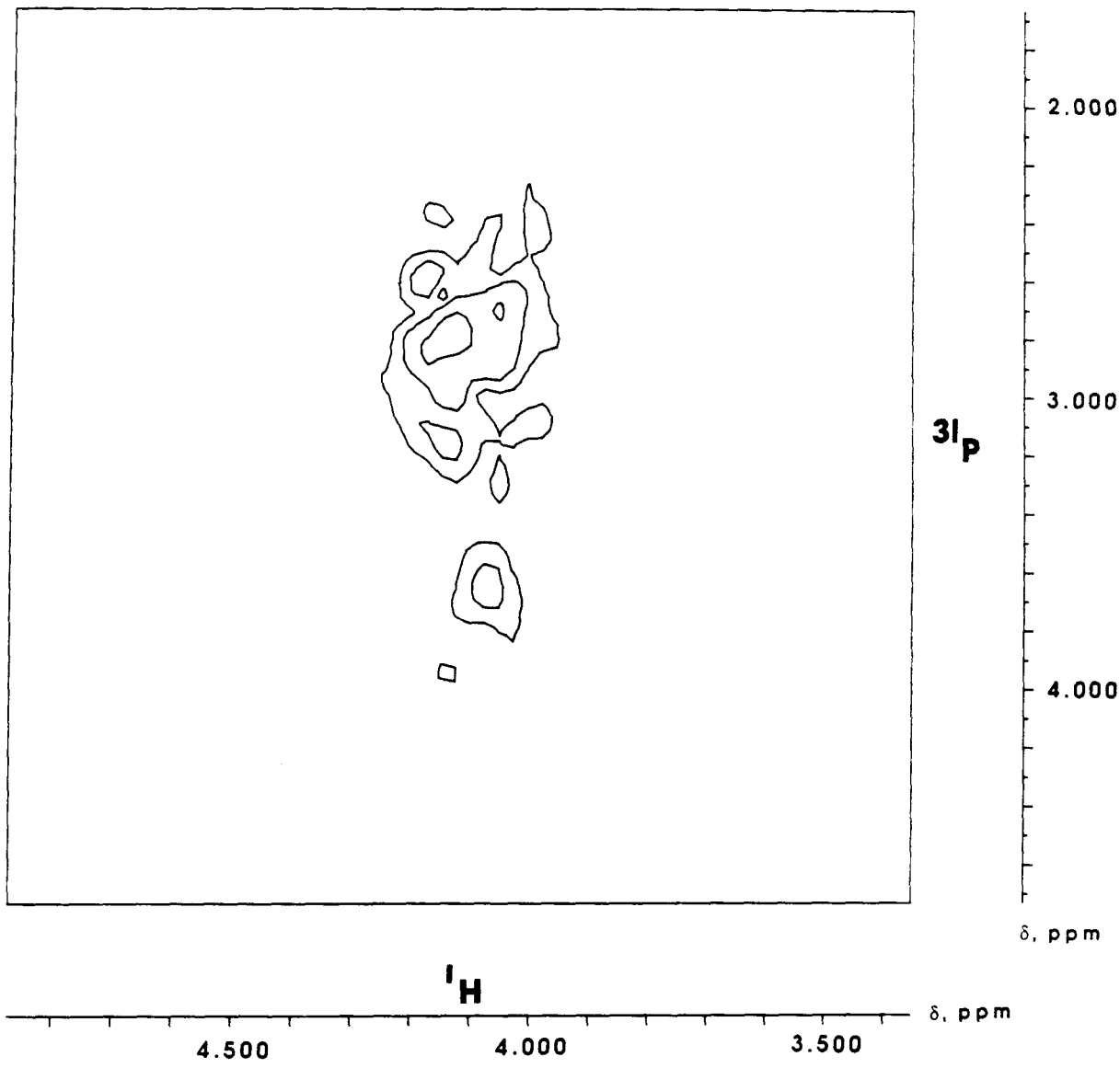
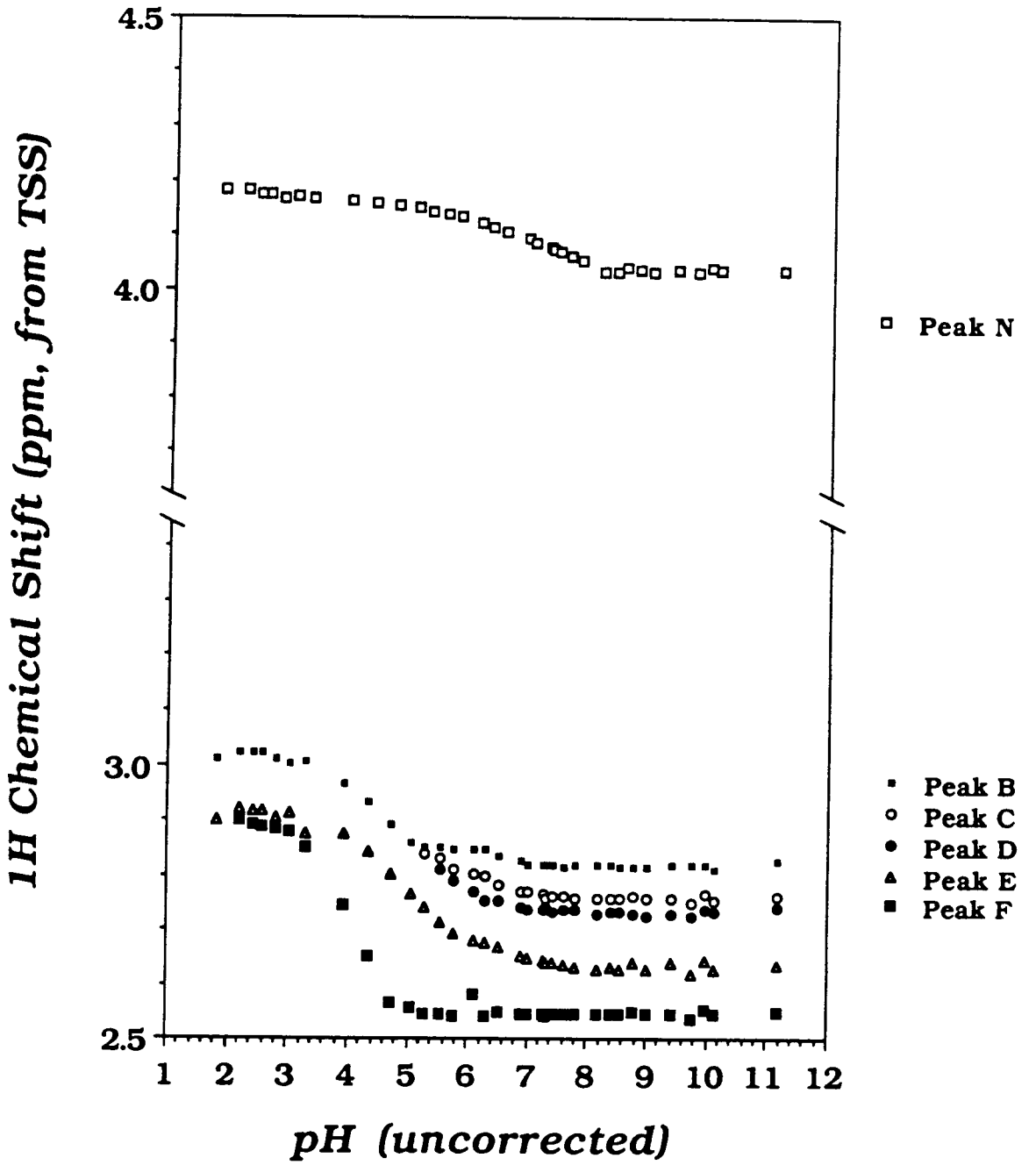

2.000
3.000
4.000
^{31}P
δ, ppm
^{1}H
δ, ppm
4.500
4.000
3.500

Figure 5: *^1H Chemical Shift vs. pH Titration Data for PSer, Asp, and Lys Residues in BDPP.*

The pH titrations are conducted in D_2O (99.96%) in the absence of buffer, using a BDPP concentration of 12 mg/ml. No correction has been made for deuterium isotope effects. ^1H Chemical shifts (ppm) are referenced from TSS (0.0 ppm).


4.5
4.0
3.0
2.5
1 2 3 4 5 6 7 8 9 10 11 12
1H Chemical Shift (ppm, from TSS)
pH (uncorrected)
Peak N
Peak B
Peak C
Peak D
Peak E
Peak F

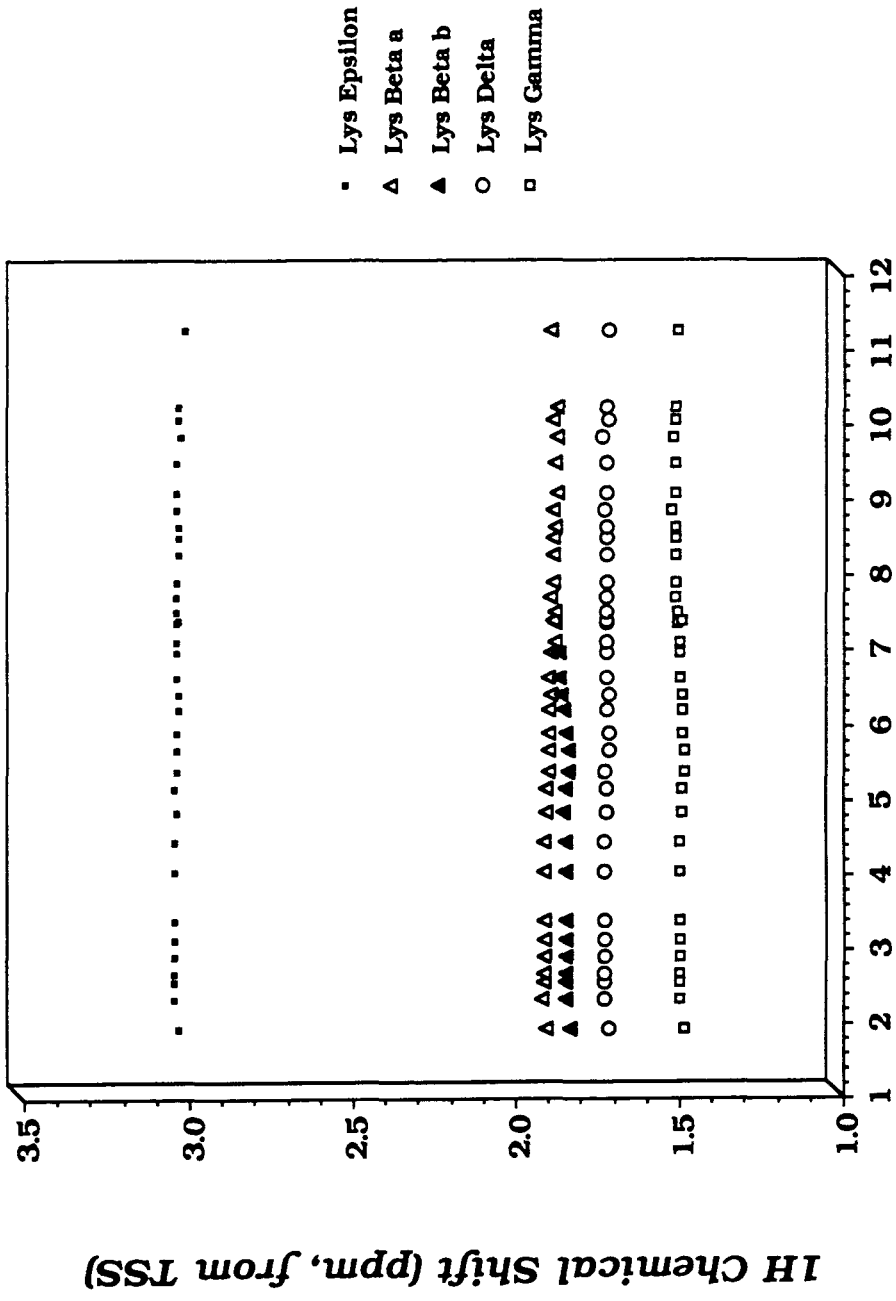
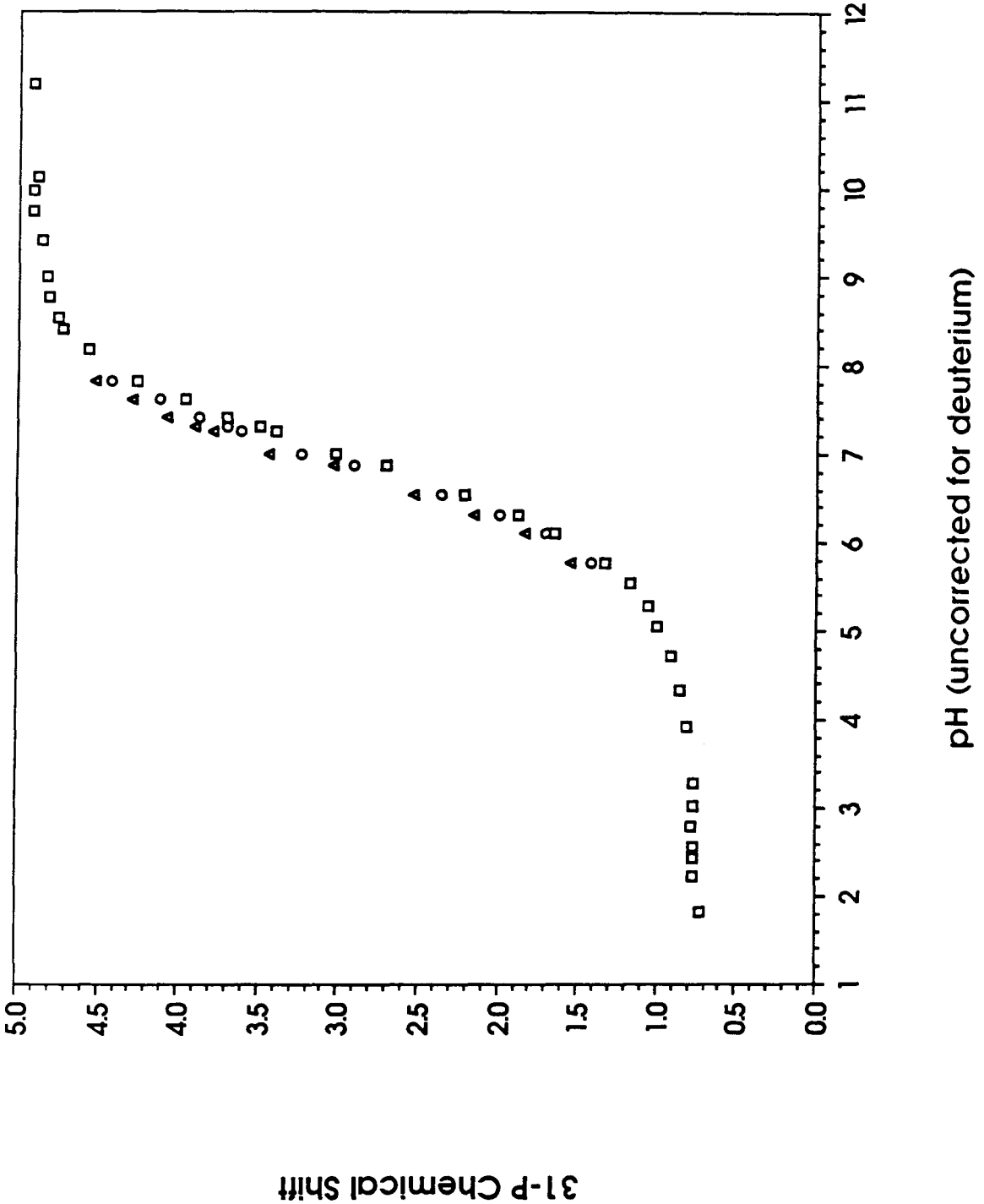

Lys Epsilon
Lys Beta a
Lys Beta b
Lys Delta
Lys Gamma
3.5
3.0
2.5
2.0
1.5
1.0
1 2 3 4 5 6 7 8 9 10 11 12
1H Chemical Shift (ppm, from TSS)
pH (uncorrected)

Figure 6: *^{31}P Chemical Shift vs pH Titration Data for BDPP PSer Residues.*

^{31}P NMR chemical shifts (ppm) are referenced from 0.5 M EHDP, pH 7.0 in 99.96% D_2O, with the EHDP ^{31}P centerpeak resonances at 10.55 and 10.48 ppm relative to 85% H_3PO_4 at 0.0 ppm). pH values are uncorrected for deuterium isotope effects. Legend: Peak α: squares; Peak β: circles; Peak χ: triangles.


31-P Chemical Shift
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10
11
12
pH (uncorrected for deuterium)

Five

NMR Studies on Bovine Dentine Phosphophoryn, a Polyelectrolyte Mineral Matrix Protein

pH-Dependent Conformational Change and Protein Folding are Mediated by Hinge Regions

Abstract

pH studies of Ca^{2+}-depleted Bovine Dentine Phosphophoryn (BDPP), a polyelectrolyte mineral matrix protein, have been undertaken in D_2O to determine how the protonation and deprotonation of Asp and Pser sidechains affect the folding of the protein molecule. Using ^1H NMR spectroscopy, we have observed perturbations in the resonance frequencies, linewidths, lineshapes, and intensities for Pser, Asp, Gly, Ala, Val, Leu/Ile, Pro, and Ser proton resonances as a function of pH. These studies reveal that certain populations of Ser, Gly α-CH and β-CH_2 and Pro δ, β, γ-CH_2 proton resonances in BDPP exhibit pH-dependent resonance frequency shifts. The "apparent" pK_a for the transition points of these frequency shifts corresponds to either the $pK_a{}^1$ of Pser monophosphate ester and/or the pK_a of Asp COOH group of BDPP polyelectrolyte regions. On the basis of these transition points, we can assign four types of Ser, Gly, or Pro-containing "intervening" regions in BDPP, based on their sensitivity to protonation and deprotonation events occurring at $(Asp)_n$, $(PSer)_n$, or $(PSer\text{-}Asp)_n$ anionic cluster regions which flank the intervening regions. Furthermore, ^1H NMR reveals that BDPP assumes a folded conformation at low pH. As pH increases, this conformation undergoes several unfolding transitions as the BDPP molecule assumes an open conformation in response to increased electrostatic repulsion between polyelectrolyte anionic regions in the protein. These folding-unfolding transitions are mediated by the intervening regions, which act as "hinges" to allow the polyelectrolyte regions to fold relative to one another. In addition, our present work reveals that BDPP possesses one or more hydrophobic cluster regions comprised of Ala, Val, Pro, Leu, and Ile residues. These sidechains exhibit broadened ^1H NMR linewidths (> 50 Hz), methyl group clustering effect, and are non-responsive to protonation/deprotonation events that occur at polyelectrolyte regions. Based on these observations, we propose a preliminary model for the BDPP molecule in solution, as well as for the folding-unfolding conformational transition that occurs in the protein molecule as a function of pH.

Introduction

One of the more significant findings described in the previous paper is that the Asp and Pser residues of bovine dentine phosphophoryn (BDPP) are organized into homologous and heterologous polyelectrolyte clusters within the protein. This result is evidenced by the titration profile exhibited by groups of these amino acids in the protein [Evans and Chan, 1992]. In this same study, we have also identified ^{1}H NMR resonances that originate from sidechain protons which lie close to the peptide backbone (within 1-2 bonds). These include Gly α,α' -CH, Asp α-CH and β, β'-CH_2, Ser α-CH and β-CH_2, Ala β-CH_3, Lys β-CH_2, and Pro α, β, γ, δ-CH_2. These resonances would be expected to be sensitive to changes in peptide torsional bond angles (Φ, Ψ, Ω). In addition, we have identified amino acid sidechain protons which are several bonds (3 or more) removed from the peptide backbone (Glu γ-CH_2, Leu δ, δ'-CH_3, Val γ, γ' CH_3, Pro γ CH_2, Lys γ, δ, ε-CH_2). These signals would be less sensitive to conformational events at the peptide backbone, but would be more sensitive to changes in sidechain non-bonding interactions (hydrophobic, electrostatic, hydrogen bonding). By varying the pH, we can therefore induce changes in the electrostatic interactions between Asp and Pser sidechains in BDPP and observe what alterations, if any, occur in the resonance frequencies, linewidths, peak areas, and peak intensities for "neutral" residues† in BDPP that lie close to or away from the peptide backbone. In effect, we can observe electrostatically-induced conformational changes in BDPP.

Specifically, we address the issue of charge neutralization and its effects upon the folding properties of BDPP by employing ^{1}H NMR spectroscopy to observe changes in the protein as a function of pH. This approach was chosen for two reasons. First, pH experiments can aid in the understanding of BDPP protein folding dynamics that occur in response to the neutralization of

---

† In this paper we are examining Glu residues as a separate case from Asp residues in terms of their participation in the secondary and tertiary structure of BDPP, due to the stoichiometry of the Glu residues relative to Asp (1:10 or smaller) in BDPP.

sidechain electrostatic charge. Protein folding that is induced by sidechain charge neutralization may play a role not only in the formation of a Ca^{2+}-protein complex that eventually forms a mineral phase, but may also influence the binding of polyelectrolyte mineral matrix proteins (PMMP) to mineral surfaces that results in controlled crystal growth. Second, NMR spectroscopy allows one to follow changes in the protein conformation not only at the anionic domain regions, but also in those regions which contain neutral amino acids. On the basis of this present study, we show that charge neutralization results in the folding of BDPP into a compact conformation, and, that this transition is mediated by conformational change involving the neutral amino acids in the BDPP sequence.

Materials and Methods

Materials. The reagents used in the present study were obtained from the following sources: EDTA [ethylenediaminetetraacetic acid, tetrasodium salt] and $CaCl_2$ from US Biochemicals (Cincinnati, OH); EGTA (ethylene glycol-O,O'-bis(2-aminoethyl)-N,N,N',N'-tetraacetic acid, disodium salt), benzamidine-HCl, ε-aminocaproic acid, PMSF, and NEM from Fluka Chemicals (NY); Tris-HCl from BRL (Bethesda, MD); D_2O (99.96% atom), deuterium chloride (DCl, 20% v/v solution, 98% atom), sodium deuteroxide (NaOD, 30% v/v solution, 98% atom) from Aldrich Chemicals (Milwaukee, WI). Chelex-100 resin was purchased from Bio-Rad (Richmond, CA). All buffers and protein solutions were prepared from 0.2 μm filtered deionized distilled water (DD_{water}). All glassware used in NMR sample preparation and analysis was first treated with EDTA/EGTA to remove unwanted divalent cations. The isolation and preparation of Ca^{2+} depleted BDPP from unerupted bovine third molars for the NMR experiments has been previously described [Evans and Chan, 1992].

NMR Experiments. ^1H NMR spectra were acquired as described elsewhere [Evans and Chan, 1992] on a Bruker AM-500. All proton chemical shifts are referenced to external 2,2,2-trimethyl-2-silapentane sulfonic acid (TSS, 0.1 M) in D_2O. pH experiments were performed as described previously [Evans and Chan, 1992]. Acquisition parameters for each particular experiment are presented in the Figure legends. The term $\underline{\Delta\delta}$ is used in this report to denote

upfield (positive) or downfield (negative) resonance frequency shift. pH values are uncorrected for deuterium isotope effects.

Homonuclear Hartmann-Hahn Correlation Spectroscopy. During the course of the pH titration, the appearance of previously unidentified aliphatic proton resonances were noted in the BDPP NMR spectrum. The assignment of these resonances was obtained via the use of ^1H-^1H HOHAHA experiments [Evans and Chan, 1992]. The sample was 12 mg/ml BDPP protein in 99.96% D_2O, pH 7.22 (uncorrected for deuterium isotope effects). The HOHAHA experiment utilized 2 K data points which were zero-filled in the F2 dimension to create a 2 K x 2 K data matrix and was obtained by a phase-sensitive squared sine-bell window function in both dimensions with no shift. Other acquisition parameters are given in the Figure legends.

Results

pH-Dependent Effects on the ^1H NMR Chemical Shifts of BDPP.

Previously, we established that electrostatic interactions occurred among the Asp and Pser residues in BDPP on the basis of pH-dependent ^1H chemical shifts observed for Asp and Pser sidechains [Evans and Chan, 1992]. We now examine the effects of pH on the ^1H chemical shifts of neutral amino acid protons that are near and distant to the peptide backbone. These protons provide convenient "reporters" on conformational changes in the BDPP protein in response to pH.

pH Titration Effects on the ^1H Chemical Shifts for Protons that are One and Two-Bond Distant from the Peptide Backbone. Figure 1 and Table 1 present the pH titration of several distinct populations of Ser (non-phosphorylated) β,β'-CH_2, α-CH and Gly α,α'-CH protons, and Pro δ-CH_2 protons, over the pH range of 1.82 to 11.17. Significant changes in the ^1H chemical shifts of these resonances are observed as the pH is altered. Specifically, Peak A (Gly α, α'; Ser β, β' protons) experiences a downfield chemical shift ($\Delta\delta$ = -0.158 ppm) over a pH range of 3 units, with a corresponding inflection point at pH 4.14. Peak G (Ser β,β'-CH_2, Gly α, α'-CH protons) shifts upfield ($\Delta\delta$ = +0.315) as the

pH is increased over the narrow range of 2 - 4. The sharp inflection point in the pH titration profile corresponds to an apparent pK_a of 2.53. Peak H (Ser β, β'-CH_2 , α-CH protons) experiences a smaller upfield shift ($\Delta\delta$ = +0.08 ppm) over a broader pH range of 2-5 units. We observe two inflection points for Peak H at pH 2.50 and 5.12. Peak I (Ser-β, β'-CH_2/α-CH and Gly α, α'-CH protons) experiences a downfield shift ($\Delta\delta$ = -0.025 ppm) from pH 1.82 to 3.03, with an apparent pK_a of 2.54. Peak J (Ser β, β'-CH_2/α-CH, Gly α-CH protons) also exhibits a downfield shift with increasing pH ($\Delta\delta$ = -0.04 ppm) with an inflection point at pH 2.51. Peak K (Pro δ-CH_2) experiences an upfield shift ($\Delta\delta$ = +0.1 ppm) corresponding to an apparent pK_a's of 2.50, and a second upfield shift ($\Delta\delta$ = +0.05 ppm) corresponding to apparent pK_a's of 5.17 and 7-8 (Table 1). Peak L (Gly α, α'-CH, Ser β, β'-CH_2/α-CH) exhibits a significant downfield shift ($\Delta\delta$= -0.3 ppm) over a pH range of 2 to 5.8, with inflection points observed at pH 2.55 and 5.17 (Table 1). Peak M (Gly α, α'-CH) does not exhibit any noticeable shift in resonance frequency as pH is altered (Figure 1). Peak 10, which represents the β, γ methylene protons of Leu, Ile, and/or the sidechain protons of Glu (γ-CH_2), also demonstrates a pH-dependent chemical shift (Figure 1). A single inflection point occurs for this peak at pH 5.17 (Table 1).

pH Titration Effects on the ^1H Chemical Shifts For Protons that are Two or More Bonds Distant From the Peptide Backbone. The proton chemical shifts for distinct populations of β, β', γ, γ'-CH_2 protons of Pro and a population of β-CH_3 protons of Ala (Peaks 6, 5, and 4), γ-CH_3 protons of Val, and δ-CH_3 protons of Leu/Ile residues (Peak 3a/b), and a population of diastereotopic methyl protons γ, γ' and δ, δ' of Val and Leu/Ile, respectively and other methyl groups (Peaks 1, 2), were observed to be independent of the solution pH in the range of pH 2.5 to 11.00 (data not shown) with the following notable exceptions. In the pH range of 6.0 to 9.0, Peak 5 exhibits an irregular pH profile. Peaks 3, 4 and 5 show slight variations in the pH profile over the range of 2.0 to 3.5 with no evidence of a inflection point. Peaks 1 and 2 reveal no effect of pH on the ^1H chemical shifts.

^{1}H NMR Spectral Broadening Effects as a Function of pH.

With alterations in solution pH, the ^{1}H NMR spectra also reveal variations in the linewidths ($\Delta\nu_{1/2}$), peak intensities and peak areas. Representative ^{1}H NMR spectra of BDPP at low and high pH are given in Figure 2 and 3, for various sidechain protons which lie near or distant to the peptide backbone.

Asp/Pser Residues Exhibit Conformational Heterogeneity With pH. Striking pH-dependent perturbations are noted for the proton resonances of Asp and Pser sidechains in BDPP. Chemical shift effects involving the diastereotopic β, β'-CH_2 protons of Asp and Pser sidechains have been observed as the pH is varied (Figure 2). For the Asp residues B-F, as pH increases from 1.82 to 5.04, the Asp β, β'-CH_2 resonances centered at 2.95 ppm appear as two major spin populations, D and E, with unresolved shoulder regions upfield and downfield of the center peak. As pH increases from 5.28 to 5.78, the Asp β-CH_2 protons experience an upfield shift and distinct β' CH_2 proton resonances are then observed as a broad peak ($\Delta\nu_{1/2}$=55 Hz) centered at 3.118 ppm. At the same time, the Asp β, β'-CH_2 proton resonances centered at 2.80 ppm become non-equivalent and separate into peaks C, D, E, and F. Above pH 5.78, the Asp β, β'-protons resolve into a series of multiplets (defined as Peaks 12-16)[Evans and Chan, 1991a], with an average $\Delta\nu_{1/2}$ of 6 Hz (at pH 8.55). The Asp β, β'-CH_2 protons (Peaks B-F) are further resolved above pH 5.78 as the individual signals become narrower. This transition from broad to resolvable proton resonances arises from changes in inhomogeneous broadening effects and changes in spin-spin relaxation as the protein undergoes conformational change and results in perturbations in the environment around the Asp residues. A similar effect is seen for Ser and Gly diastereotopic protons (Peaks 17-20, see below). Since Peaks B-F and 12-16 are comprised in part of β, β'-CH_2 Asp protons, and the chemical shift separation between these resonances extends from 2.54 to 3.18 ppm at pH 6.94, 25^o C (a range of 320 Hz), this would suggest that there may be different environments for Asp β, β' sidechain protons in BDPP at neutral pH.

The β, β'-CH_2 proton resonance for BDPP Pser residues (Peak N, Figure 2) also undergoes alterations in linewidth and lineshape as a function of pH. At

low pH (pH 1.82 to 5.28) peak N assumes an asymmetric lineshape with a downfield shoulder region, with an apparent $\Delta\nu_{1/2}$ of 47 Hz. From pH 5.55 to 6.53, the lineshape becomes symmetric, and the linewidth decreases to a minimum $\Delta\nu_{1/2}$ of 21.5 Hz at pH 6.53; at the same time, peak area and peak intensity are observed to increase by 100% or more. From pH 6.53 to pH 11.17, the linewidth increases exponentially to a maximum value of 75.5 Hz, with the lineshape again assuming a symmetric profile, and peak area remaining relatively constant. This symmetry/asymmetry lineshape effect reflects the change in the environment for Pser β, β'-CH_2 protons. At low pH, the diastereotopic protons become non-equivalent due to neutralization of the monophosphate ester. In the physiologic pH range, however, these diastereotopic protons become nearly equivalent.

Conformational Effects Observed for Non-Asp/Pser Proton Resonances that are One to Two Bonds Distant from the Peptide Backbone. Peaks G,H, I, and J, which represent populations of Ser β, β'-CH_2, α-CH, and Gly α,α'-CH protons of BDPP, initially appear as narrow resonances at low pH (average $\Delta\nu_{1/2}$=3-8 Hz at pH 1.82), but become nearly coincident as a broad peak ($\Delta\nu_{1/2}$=43 Hz) centered at 3.89 ppm (Figure 2) as pH increases to 6.94. Above pH 6.94, this group of broad resonances becomes obscured by the β-CH_2 Pser resonance that has shifted as the pH increases. From pH 1.82 to 3.03, Peak I experiences a decrease in signal intensity. At pH 3.03 and above this signal becomes equivalent with peaks G and H. Peak J also exhibits variations in signal intensity over the pH range investigated, with maximal intensity at pH 1.82 and minimal signal intensity at pH 11.17.

Peaks 17-20, which represent the diastereotopic β, β'-CH_2 and α, α'-CH protons of Ser and Gly, respectively, are not detectable at pH 1.82, but begin to appear on the ^{1}H NMR spectrum at pH 2.22 as a broad peak ($\Delta\nu_{1/2}$ = 20). And, as the pH increases to 6.30, this single peak resolves into a multiplet of four narrow resonances (average $\Delta\nu_{1/2}$ = 8-10 Hz, Figure 3). From pH 6.30 to pH 7.30, these peaks broaden (average $\Delta\nu_{1/2}$ = 15-20 Hz) with a concomitant decrease in signal intensity. Above pH 7.30, this multiplet sharpens again into a set of four narrow resonances (average $\Delta\nu_{1/2}$ = 8-10 Hz) and reaches maximum intensity at pH 8.41. At higher pH, the signal intensity decreases 20%

as pH is increased from 8.41 to 11.17. These variations in the peak intensities and linewidths with pH provide a measure of conformational change for Ser and Gly residues in BDPP. Two important transitions are noted: the first, occurring at low pH (pH 1.82), where no detectable signal is observed, and the second, occurring at pH 6.30 to 7.30, where both linebroadening and peak intensity effects are observed for the multiplet.

Peaks A, K, L, and M, representing the sideechain protons of Ser, Gly, and Pro, also experience perturbations during protonation and deprotonation. As shown in Figure 2, at low pH Peak A is initially obscured by chemical shift overlap with the β, β' protons of Asp residues. This situation changes above pH 5.5, where Peak A appears as a downfield shoulder region of Peak B, and is resolved as a narrow resonance above pH 8. Peak K (Pro δ-CH_2) undergoes variation in peak intensity, with maximal intensity at pH 1.82. As the pH increases, this signal intensity decreases, with the total disappearance of this signal from the ^1H NMR spectrum at pH 8.17 and above. Simultaneously, Peak M appears at pH 6.10 and increases in signal intensity as pH increases. Peak L exhibits maximal signal intensity at pH 1.82. This signal diminishes in intensity as the pH is increased from 1.82 to 5.28, but above pH 5.28, the signal intensity increases with pH, reaching a maximum at pH 7.42. From pH 7.42 to 11.17, the signal intensity begins to diminish again (Figure 2). In contrast, no significant changes in either linewidth, peak area, or lineshape are noted for Peak 10 (Leu/Ile β-CH_2 and/or Glu γ-CH_2 sidechain protons) as a function of pH. Overall, these pH induced effects on peak areas, intensities, and linewidths of Peaks G-M , Peak 17-20, and Peak 10 are reversible.

Conformational Effects Observed for Non- Asp/Pser Proton Resonances That Are Two or More Bonds Distant From the Peptide Backbone. Further evidence for conformational change is also provided by Peaks 1 and 2, which in part correspond to the δ, δ' methyl protons of Ile and Leu, the γ, γ' methyl protons of Val, and novel methyl groups in BDPP. These resonances exhibit alterations in peak intensity and peak area as a function of pH; however, the linewidths for these two resonances remain relatively constant (Figure 3). The ratios of the peak areas (Peak 2:Peak 1) is 1:1 at low and high pH, but in the pH range from 5 to 8, this ratio decreases, with a minima of 0.3 at pH 6.0. These changes in

peak areas and intensities may describe folding transitions in BDPP as a function of pH, in which the Val, Ile, and Leu sidechains that reside in certain clusters of BDPP, or, the silicon-modified amino acids, experience alterations in their tertiary structure, such that inhomogeneous broadening and spin-spin relaxation effects are observed for these sidechain protons.

Throughout the pH titration we note also the appearance and disappearance of well-resolved resonances ($\Delta\nu_{1/2}$ approximately 4-10 Hz) superimposed upon broader resonances (30-50 Hz) in the BDPP ^{1}H NMR spectrum centered at 0.876, 1.154, 1.301, 1.545, 1.610, 1.892, 2.045, and 3.202 ppm (at pH 7.22)(Figure 3). We do not observe any resonance frequency shifts for these signals over the pH range investigated in this study. However, these peaks are reproducible in their intensity and peak area at various pH values. A 2-D HOHAHA experiment reveals that several of these resonances exhibit scalar connectivities (Figure 4). From the *J*-connectivity patterns and ^{1}H NMR chemical shifts, we have tentatively assigned these resonances to the $A_2(T_2)MPX$ spin system of Pro γ, γ', β, β', and δ-CH_2 pyrrolidine ring protons. The relayed magnetization of these Pro γ→β→δ protons gives rise to distinct *J*-connectivities which can be categorized as follows: (1) One set of connectivities is comprised of the γ, γ' resonance at 0.876 ppm, the β, β' resonance at 1.545 ppm, and the δ resonance at 3.526 ppm; (2) A second set of connectivities is comprised of the γ, γ' resonance at 1.301 ppm, the β, β' resonance at 1.545 ppm [which demonstrates chemical shift overlap with the Pro β, β'-CH_2 resonances in (1)], and the δ resonance at 3.202 ppm (Figure 4). (3) A third set exhibits weak β, β' and γ, γ'-CH_2 resonances and connectivities at 1.154, 1.301, and 1.545 ppm respectively, but without δ-CH_2 or α-CH connectivities (Figure 4); (4) The fourth set exhibits β, β' and γ, γ'-CH_2 proton resonances at 1.892 and 2.045 ppm, respectively, with no observable *J*-connectivities to δ-CH_2 or α-CH protons. Obviously, since the γ and β imino ring protons of Pro exhibit considerable chemical shift overlap [Wishart *et al.*, 1991; Grob and Kalbitzer, 1988], our ability to distinguish between γ and β CH_2 protons in the HOHAHA spectrum is admittedly limited, and these assignments should be considered as interchangeable. Unlike the HOHAHA spectrum at pH 6.94 [Evans and Chan, 1992], no scalar couplings between pyrrolidine methylene ring protons and Pro α-CH protons are observed at this pH. This

may be explained by short T_2 values for the Pro α-CH protons which experience restricted motions on the peptide backbone. This would suppress spin-lock magnetization transfer for these protons [Torchia *et al.*, 1989]. Thus, in conjunction with the pH titration, the 2-D HOHAHA experiment have allowed us to identify motionally restricted proline-containing regions in the BDPP protein that are sensitive to pH conformational effects. Some of these proline ^1H signals are not initially observed in our earlier assignment studies of BDPP at pH 6.94 [Evans and Chan, 1992], and their appearance and disappearance from the ^1H NMR spectrum might have originated from changes in the the secondary and/or tertiary conformation of Pro-containing regions in BDPP . These conformational changes are induced by deprotonation of Asp/Pser sidechains which affect the folding of the BDPP molecule. Additonally, the effects of changing ionic strength and solvent interactions may also influence the protein conformation and result in similar linebroadening effects.

Discussion

In our study of pH titration effects on the Asp and PSer residues in BDPP, we established that these two anionic sidechain groups are involved in long-range electrostatic interactions along the polypeptide [Evans and Chan, 1992]. In the present investigation, we have observed apparent pH titration effects for "neutral" sidechain protons (Pro, Gly, Ser), as well as evidence for other pH effects occurring at Pro, Ala, Val, Glu, and Leu/Ile residues. These neutral amino acids are influenced to a great degree by the protonation states of Asp and/or Pser sidechains, as the protein undergoes conformational change in response to alterations in electrostatic interactions involving these residues.

One mechanism by which Asp and Pser sidechains can exert their effect on the resonance frequency of the neutral sidechain protons involves the through-space effects of the <u>electrostatic charge</u> on the detailed electron distribution in the vicinity of these nuclei. A charged chemical group (COO^-, $CH_2PO_4^{-2}$) which is in close proximity (< 5 Å) to nuclei on a neighboring amino acid would act to polarize the surrounding electrons of these nearest-neighbor atoms and thus alter the resonance frequency for these nuclei in response to protonation and deprotonation [Jardetzky and Roberts, 1981]. A second

mechanism that can affect this frequency shift involves torsion bond angles, in which the ionization of Asp and PSer residues lead to an alteration in the torsion bond angles of the peptide backbone (Ψ, Φ, Ω), as the charged sidechains rotate away from each other to minimize unfavorable electrostatic energies. Given that Peaks G-M represent in part the α-CH protons of Ser, Gly, and Pro amino acids in BDPP, the resonance frequency of these protons would be sensitive to perturbations in the secondary and tertiary structure of a peptide region. In fact, the proton chemical shifts for α-CH protons have been used as a means of deducing secondary structures in proteins [Szilagyi and Jardetzky, 1989]. Likewise, the Ser β, β' methylene and Ala, Val, Leu/Ile methyl protons, which are several-bonds removed from the peptide backbone, might also be expected to be sensitive to perturbations in peptide bond torsion angles, albeit with less sensitivity compared to the α-CH protons.

Apparent pK_a's Identify the Anionic Residues Which Influence The Non-Asp/Pser Sidechains. The ionization of a group in a protein results not only in the alteration of the chemical shift for that group, but also those of neighboring residues [Jardetzky and Roberts, 1981]. The presence of "apparent" pK_a's for methyl group protons, aromatic ring protons, and other sidechain protons of non-ionizing amino acid sidechains in proteins has been utilized to denote the close proximity of one or more charged amino acids in relation to the non-charged sidechain [Campbell *et al.*, 1975a,b; Jardetzky and Roberts, 1981]. In this work, the apparent pK_a's can be used to distinguish between the two types of ionizing acids -- carboxylate groups and monophosphate ester groups -- that reside in BDPP (Table I). On this basis, one can subdivide the Pro, Gly, and Ser residues in BDPP into four groups (Table II). Where the carboxylate and monophosphate ester pK_a's are both observed is taken as an indication of the presence of $(PSerAsp)_n$ sequence block. Group III possesses the same two inflection points noted for the Group II residues; in addition, a third, weak inflection point is observed near pH 7-8 (Peak K: Pro δ-CH_2)(Figure 1; Table 1). As for the Group II residues, these hinge region sidechains are experiencing the effects of protonation on Asp/Glu and Pser residues in the peptide chain; the third ionization may arise from the second ionization of PSer residues [Evans and Chan, 1992]. In the case of Group IV, the observed carboxylate pK_a for Peak 10 might represent the ionization of Glu δ-COOH; for Pro γ, β-CH_2

protons, this pK_a may reflect the ionization of neighboring Asp carboxyl groups, as is the case for Peak A. In general, the apparent pK_a's observed here are characteristic of weak polyacids: the pK_a values are higher than those of the free amino acid (Asp, Glu, or PSer) and the titration curves exhibit a broad profile covering a transition of 2 or more pH units [Evans and Chan, 1992] indicating a heterogeneous distribution of pK_a's arising from interacting charged groups. Our interpretation of the pH titration data is that the non-Asp/PSer amino acid groups in BDPP can be divided into two classes: those that are in close proximity to either homologous anionic clusters (Groups I, IV), comprised of either carboxyl or monophosphate ester sidechains, or, those that are in juxtaposition to the heterologous anionic clusters comprised of both types of charged sidechain groups (Groups II,III).

Evidence for conformational change involving sidechain and/or peptide torsional bond angle rotation to accommodate charge repulsion is also suggested by two related phenomena: (1) In the pH titration curves for Ser, Pro, and Gly residues (Figure 1), we note that only the pK_a^1 of PSer and the pK_a of Asp γ-COOH are observed in the titration of Peaks G-L; we do not observe the pK_a^2 of PSer monophosphate ester. If the frequency shifts for these neutral sidechains were induced by through-space electrostatic interactions from Asp and PSer, then the presence of a secondary inflection point (pK_a^2 = 6.9 - 7.0) from Pser would have been evident in the pH titration profiles for the neutral residues. Instead, our data indicate that above pH 5.17, the neutral residues represented by peaks G-L are not affected by the mono-to-dibasic deprotonation transition of PSer sidechains. These observations provide evidence that the neutral amino acid sidechains have undergone torsional bond and sidechain conformational change in response to anionic sidechain repulsion from deprotonating Asp and PSer anionic clusters in the pH range of 2 to 5. These conformational changes have led to the localization of these neutral residues away from the Asp/PSer anionic clusters, such that further deprotonation of PSer at pH 6.0-7.0 has no effect on the proton chemical shifts of these neutral residues. Thus, these Ser, Gly, and Pro sidechains represented by Peaks G-L must be localized in regions which flank these anionic clusters. It should be noted that an inflection point was observed for Peak K in the pH range expected for Pser pK_a^2 (Figure 1). It is clear that this

group of Pro imino ring protons arise from the residues that are influenced by Pser residues. (2) Direct evidence for peptide torsion bond angle perturbation is seen in the ^1H NMR BDPP spectrum (Figure 2), in which proton resonances of Ser (β,β',α), Gly (α,α'), and Pro (δ-CH_2) (Peaks G-M) experience changes in linewidth, peak area, and peak intensity in response to protonation of Asp/Pser sidechains. These resonances would be expected to be sensitive to peptide backbone conformational change. In addition, the appearance and disappearance of certain Pro β, γ, δ-CH_2 sidechain proton resonances (Groups A, B, C, and D, narrow signals seen superimposed on peaks 3-7, Figure 3, and the appearance of multiplet at 3.526 ppm, Figure 2) as pH is varied may also be the result of broadening effects induced by protein conformational change. Collectively, these effects are indicative of the non-Asp/PSer sidechains experiencing changes in torsional bond angles as the Asp and PSer sidechains become deprotonated and move away from one another. Thus, based on these data, there are certain populations of Ser, Gly, and Pro sidechains in BDPP that are conformationally sensitive to events occurring in polyelectrolyte clusters in BDPP, but do not directly experience through-space electrostatic charge effects from Pser and Asp sidechains in the protein.

It should be noted that the proton spins which resonate near 0.00 ppm in BDPP also exhibit pH-dependent intensity effects. Preliminary evidence suggests that these protons arise from novel chemical groups (e.g., methyl protons belonging to a silane or siloxane group) which is bound to BDPP in some manner [Evans and Chan, 1992]. The absence of a pH-induced chemical shift effect would imply that these protons are not in close proximity (i.e., > 5 Å) to charged groups, since perturbations in neighboring electrostatic charge would have induced a frequency shift in these proton resonances. These peaks, however, did show intensity effects as a function of pH (Figure 3), indicating that conformationally-induced changes in these proton environments was occurring in response to protonation. These inhomogeneous broadening effects occur in response to changes in spin-spin relaxation time (T_2) and resonance frequency for these protons as their local environment is perturbed.

BDPP Possesses Multiple Regions Which Interrupt Anionic Clusters. From the foregoing, it is evident that as the polyelectrolyte regions undergo

protonation or deprotonation, a certain population of nearby Ser, Pro, and Gly sidechain protons in BDPP experience electrostatically induced conformational changes with respect to either heterologous or homologous polyelectrolyte clusters. We argue that these amino acids, which may exist singly or as a cluster, are flanked by anionic clusters. In support of this contention are the following observations: (1) The stoichiometry of the amino acid composition of BDPP dictates that the number of non-Asp/PSer residues in BDPP is small relative to the total Asp/PSer content: 1 Leu: 28 PSer/Asp; 1 Ala:25 PSer/Asp; 1 Val:26 PSer/Asp; 1 Ile:150 PSer/Asp; 1 Ser:20 PSer/Asp; 1 Glu:13 PSer/Asp, 1 Pro:20PSer/Asp; and 1 Gly:10 PSer/Asp (per 1000 amino acids, reported as approximate values)[Stetler-Stevenson and Veis, 1983; Veis, 1984]. (2). Evidence for intervening Glu and neutral amino acids residing between anionic amino acids in the peptide sequence of phosphophoryn has been established recently in a number of PMMP proteins isolated from various extracellular matrices. For BDPP, a 23 residue N-terminus of this protein has been recently sequenced in our laboratory, and this fragment features Asp and Ser residues separated by Asn, Gly, Glu, and Pro amino acids [Evans and Chan, 1991]. Other evidence for intervening amino acids that are distributed between charged residues has been reported for rat α-phosphophoryn, in which a repeat motif of neutral amino acids, followed by a polyanionic sequence (e.g, $(PSer)_n$) was established in the peptide sequence [Sabsay *et al.*, 1991]. A recently isolated PMMP protein from tooth dentine also features polyanionic sequence domains separated by hydrophobic domains [George *et al.*, 1992]. In other PMMP proteins, the repeat motif Asp-Y-Asp was elucidated for soluble proteins isolated from the mollusca *Crassostrea*, *Mercenaria*, and *Nautilus* [Weiner and Hood, 1975], where *Y* is either Gly or Ser; and the sequence Asp-Pro-Thr-Asp was identified in two soluble proteins isolated from two shell layers of the bivalve *Mytilius californianus* [Weiner, 1983]. The dipeptide Gly-Ser and individual amino acids Gly and Ser was also isolated from formic acid hydrolysates obtained from *Crassostrea virginica* oyster shell [Wheeler *et al.*, 1988]. (3) From the titration data presented above, we have witnessed that certain amino acid residues experience conformational change in response to protonation/deprotonation events occurring at Asp and Pser residues. From the apparent pK_a's and alterations in ^1H NMR linewidth, peak area, and peak intensity, these non-Asp/Pser amino acids in BDPP must exist in close

approximation to either homologous [$(Asp)_n$, Group 4; $(PSer)_n$, Group 1], or to heterologous anionic clusters [$(Asp\text{-}PSer)_n$, Group 2,3]. Thus, substantial evidence from amino acid composition and peptide sequence data, as well as our NMR data, establishes the presence of neutral amino acids and Glu residues as intervening amino acids that flank sequences of charged amino acids such as Asp and Pser.

Evidence for Neutral Amino Acids Which Do Not Experience pH-Induced Conformational Change. In contrast to the evidence presented above regarding interactions between Asp/Pser and other sidechains in BDPP, there exists a group of non-charged amino acids in BDPP which do not exhibit this behavior. These involve the sidechain protons of Peaks 3-7, and to a certain extent, Peaks 1 and 2, which collectively represent a population of Pro (Groups A, B, C, D), Leu, Lys, Ile, Val, Ala, and Si-modified amino acids or other chemical groups in BDPP [Evans and Chan, 1992]. We observe peak intensity, peak area, and linewidth changes for certain populations of Pro sidechain pyrrolidine ring protons which comprise these signals (Figure 3, 4), which may indicate that these Pro residues are located in motionally restricted clusters in BDPP that are sensitive to polyelectrolyte domain-induced conformational changes. However, there remains a certain population of Pro ring protons whose resonances are broadened ($\Delta\nu_{1/2} > 30$ Hz)(Peaks 3-4, Figure 3) but demonstrate no chemical shift-pH dependency in the pH range of 3.0 to 7.0 (data not shown). Thus, it is apparent that a certain population of Pro, Leu, Ile, Val, Ala, and Thr sidechains in BDPP are insensitive to the ionization of PSer, Asp, or Glu sidechains in BDPP. We have observed a weak pH-chemical shift effect in the pH range of 2.0 to 3.0 for Peaks 3, 4, and 5 (data not shown), which may be in response to the first ionization of PSer, but this effect is weaker than that seen for the other sidechain protons (e.g., Peaks G-M, Peak 10, Figure 1), and from the data we are unable to establish an inflection point for these resonances .

One explanation for the absence of a pH-conformational effect here is that these residues are located in a region of BDPP which is not contiguous or exposed to a polyelectrolyte domain. This situation could arise if these neutral amino acids reside away from the polyelectrolyte clusters in a sequence motif.

Another possible explanation is that these non-ionic amino acids may be distributed as intervening regions throughout the BDPP molecule, but because of secondary and tertiary structural arrangement, these amino acid sidechains are oriented away from polyelectrolyte clusters by peptide folding. These folded, non-ionic regions of the peptide chain may be energetically stabilized by hydrogen bonding or hydrophobic interactions, and thus does not experience electrostatic or torsion angle perturbations resulting from charge-charge repulsion of anionic sidechains. In our previous NMR study, we established that these unaffected sidechain protons of this population of Pro, Leu, Ile, Val, and Ala (peaks 3-6) possess broad proton linewidths (> 30 Hz) that are not observed for other regions of the BDPP protein at pH 6.94. This is suggestive of a population of amino acid sidechains that experience slow, restricted motion in solution compared to the rest of the protein molecule [Evans and Chan, 1992]. In addition to NMR denaturation experiments, peptide and nucleic acid sequencing studies will aid in establishing where these amino acids are in the primary sequence of BDPP, and should help to explain their lack of response to the pH titration of Asp and Pser sidechains.

Hypothetical Model for the Solution Conformation of Ca^{2+} Depleted BDPP. From the foregoing, and our previous report [Evans and Chan, 1992], it is clear that BDPP is comprised of several types of sequence regions. Figure 5 presents one possible schematic interpretation of the structure of BDPP at neutral pH (Structure IV), in which the protein is comprised of four types of regions. (1) Polyelectrolyte Clusters. These regions, either homologous [$(Asp)_n$ or $(PSer)_n$] or heterologous [$(Asp\text{-}PSer)_n$] in sequence, are non-contiguous with one another and would be expected to assume an extended or open conformation upon deprotonation. The sequence length of these regions, their secondary and tertiary structure, and the number of such regions in BDPP are not known at present. Since these regions of BDPP may be considered the prime sites for Ca (II) binding, in subsequent studies we will also refer to these regions as *polyelectrolyte calcium binding domains*, or **PCBD**. (2) Intervening or Hinge Domains (**HD**). These flanking regions are comprised of Ser, Gly, and Pro, and may exist either as a cluster of amino acids or as individual amino acids that interrupt the polyanionic sequences. The sequence length, number of extant regions in BDPP, and secondary/tertiary structure of these flanking

regions is not known. (3) Lysine Containing Regions (not shown in Figure 5). From our previous report, we established that Lys residues in BDPP do not exhibit normal titration behavior and thus may participate in the formation of salt-bridges with anionic residues Asp, PSer, and Glu. These types of non-bonding interactions may either be intramolecular or intermolecular, and the actual sites in the BDPP sequence where these interactions occur, along with the distribution of Lys residues in the BDPP sequence, is uncertain at present. (4) Hydrophobic Cluster (**HC**) Regions. One or more regions of BDPP are comprised of Pro, Leu, Ile, Ala, Val, and Thr amino acids which form a stable cluster that is unaffected by protonation events occurring at anionic domain regions. The location(s) of cluster region(s) in the BDPP sequence is not known, but these regions must be located away from the polyelectrolyte clusters (in Figure 5, arbitrarily represented at the C-terminus). These regions may be hydrophobic or stabilized by other types of bonding and/or non-bonding interactions.

pH -Induced Folding of BDPP as a Function of Charge Neutralization. From the data presented in this report, we have observed a number of conformational transitions in BDPP (Table 1; Figures 1-4). These are presented as a protein folding model in Figure 5. The BDPP protein at low pH (Structure I) would assume a condensed conformation, involving folding of the polyelectrolyte cluster regions and the intervening regions into a more compact structure. The first transition is represented by Structures I and II, in which the Pser residues would undergo deprotonation between pH 2 and 3 and assume the monobasic form and Pser-containing anionic clusters begin to unfold due to charge repulsion (Structure II). Concomitantly, intervening clusters containing one or more Ser, Gly, and Pro residues (Group 1-4) experience conformational change as flanking Pser sidechains move away from one another. This unfolding process would be expected to continue in the second transition, in which the Asp carboxyl groups would undergo deprotonation between pH 3 and pH 7, and the Pser residues are either fully deprotonated or are in the monobasic form. This transition is represented by Structures II and III. The polyelectrolyte cluster regions would be anticipated to move farther away from one another due to the presence of increased electrostatic charge repulsion; in turn, the flanking intervening regions would experience a conformational

change brought about by the enhanced electrostatic charge repulsion between anionic clusters. To some degree, it is possible that a certain amount of folding in the polyelectrolyte and intervening regions may occur up to pH 7 (Structure III) due to the Pser residues being partially shielded in the monobasic form or by Lys salt-bridging interactions. Finally, in the third transition, with an increase in pH (> 7.0) the Pser residues convert to the dibasic form, at which point the BDPP protein undergoes unfolding to assume an extended structure (Structure IV). These transitions between condensed and extended forms are reversible. Thus, pH-induced transitions in the BDPP structure involve conformational changes not only in polyanionic regions, but at or near the flanking intervening regions. This results in the molecule progressing from a condensed conformation at low pH to an extended conformation towards physiological pHs.

Overall, the results reported in this study provide us with a different picture of BDPP and PMMP proteins in general. Rather than being thought of as "acidic", random-coil proteins that bind divalent cations, they are far more complex than previously realized, in that they are structurally organized into repeating charged clusters which are separated by non-ionic amino acids. Clearly, BDPP and other PMMP proteins have evolved as specialized polypeptides whose structure is responsible for mediating solution-to-solid inorganic mineral transformation.

Acknowledgements

We thank Dr. Arthur Veis for discussions regarding the preliminary amino acid sequence of rat a-phosphophoryn, and Dr. Heinz Lowenstam for helpful discussions and reading of the manuscript. The technical expertise of John Racs in CZE separation is also acknowledged.

References

Addadi, L., Moradian, J., Shay, E., Maroudas, N. G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* 84 2732-2736.

Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* 82 4110-4114.

Bax, A., and Davis, D. G. (1985a) *J. Mag. Res.* 65 355-360.

Bax, A., and Davis, D. G. (1985b) *J. Am. Chem. Soc.* 107 2820-2821.

Campbell, I. D., Dobson, C. M., and Williams, R. J. P. (1975a) *Proc. R. Soc. London* A 345 23-40.

Campbell, I. D., Dobson, C. M., and Williams, R. J. P. (1975b) *Proc. R. Soc. London* A 345 41-59.

Evans, J. S., and Chan, S. I. (1992) *Biochemistry* , previous manuscript.

Evans, J. S., and Chan, S. I. (1991) in *Proteins: Structure, Dynamics and Design.* (Renugopalakrishnan, V., Carey, P.R., Smith, I.C.P., Huang, S.G., and Storer, A.C., Eds.) ESCOM Press, Leiden, Netherlands pp. 251-260.

George, A., Simonian, P., Tylzanowski, P., and Veis, A. (1992) *Connect. Tiss. Res.* 27 (2-3) 107.

Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology* (eds. by Jardetzky and Roberts) Academic Press, New York, New York, pp. 143-186, 227-231, 234-270, 448-465.

Lee, S. L., Veis, A., and Glonek, T. (1977) *Biochemistry* 16 2971-2979.

Lee, S. L., Glonek, T., and Glimcher, M. J. (1983) *Calcif. Tiss. Int.* 35 815-818.

Mann, S. (1988) *Nature* 332 119-124.

Marsh, M. E. (1989) *Biochemistry* 28 339-345.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* 42 104-112.

Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) *Biochem. J.* 276 699-707.

Stetler-Stevenson, W. G., and Veis, A. (1986) *Calcif. Tiss. Int.* 38 135-141.

Szilagyi, L., and Jardetzky, O. (1989) *J. Mag. Res.* 83 441-449.

Termine, J. D., and Conn, K. M. (1976) *Calcif. Tiss. Int.* 22 149-157.

Termine, J. D., Eanes, E. D., and Conn, K. M. (1980) *Calcif. Tiss. Int.* 31 247-251.

Torchia, D. A., Sparks, S. W., Young, P. E., and Bax, A. (1989) *J. Am. Chem. Soc.* 111 8315-8317.

Veis, A. (1984) in *Extracellular Matrix Biochemistry.* (Eds Piez, K., and Reddi, A. H.) Elsevier Publishing, New York, New York, pg. 350.

Weiner, S. (1983) *Biochemistry* 22 (17) 4139-4145.

Weiner, S., and Hood, L. (1975) *Science* 190 987-989.

Weiner, S., and Traub, W. (1984) *Phil. Trans. R. Soc. London* B 304 425-434.

Wheeler, A. P., Rusenko, K. W., and Sikes, C. S. (1988) in *Chemical Aspects of Regulation of Mineralization* (Eds. C.S. Sikes and A.P. Wheeler) University of South Alabama Publication Services, Mobile, Alabama pp. 9-13.

Wishart, D.S., Sykes, B.D., and Richards, F.M. (1991) *J. Mol. Biol.* 222 311-333.

Zanetti, M., de Bernard, B., Jontell, M., and Linde, A. (1981) *Eur. J Biochem.* 113 541-545.

Figure 1: *^1H NMR Chemical Shifts vs. pH for Glu, Gly, and Ser Sidechain Protons in BDPP (12 mg/ml) at 298 K, in 99.96% D_2O at Low Ionic Strength.*

pH titration procedures are identical to those outlined in our accompanying paper [Evans and Chan, 1992]. pD readings from the pH meter have not been corrected for deuterium isotope effects. Symbol legends are given in the Figure. All ^1H chemical shifts (ppm) are reported relative to external TSS in D_2O.

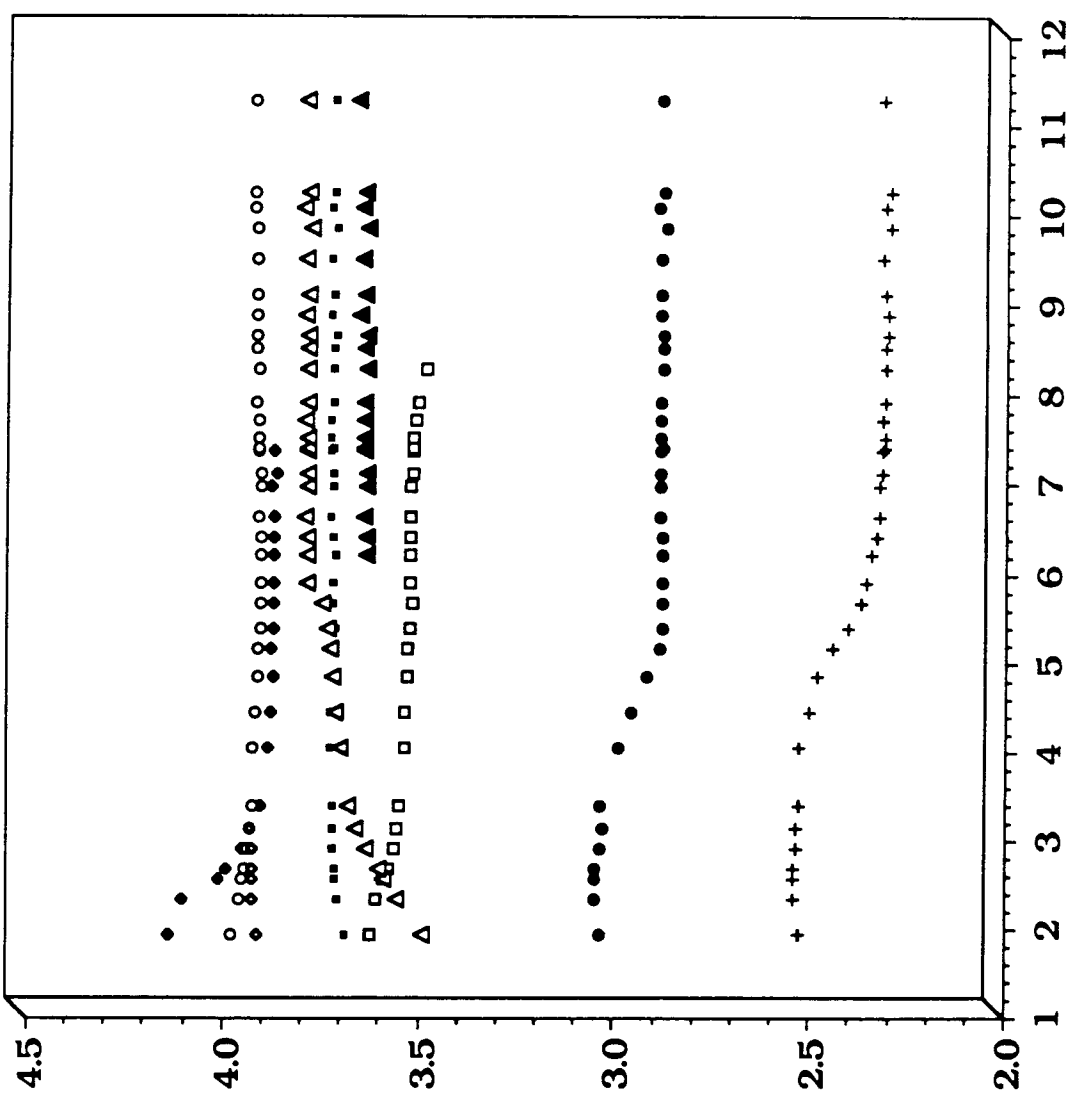

Peak A
Peak G
Peak H
Peak I
Peak J
Peak K
Peak L
Peak M
Peak 10
4.5
4.0
3.5
3.0
2.5
2.0
1
2
3
4
5
6
7
8
9
10
11
12
1H Chemical Shift (ppm, from TSS)

Table 1: †Observed inflection points are determined by non-linear least squares fit to chemical shift vs. pH data points obtained for BDPP sidechain resonances at low ionic strength. **Boldface** numbers indicate secondary inflection points on titration curves. *Inflection point for this peak non-resolvable prior to end of titration; pK_a is estimated.

Table I

Titration Behavior of BDPP Residues in D_2O at 298 K†

Peak	Sidechain	Inflection pK_a	Peak	Sidechain	Inflection pK_a
10	P β, E γ-CH_2	5.17	I	S β, α / G α, α'	2.54
A	S β, β' / G α, α'	4.14	J	S β, α / G α, α'	2.51
G	S β, β' / G α, α'	2.53	L	S β, α / G α, α'	2.55 **5.17**
H	S β, β', α-CH	2.50 **5.12**	K	P δ, δ'	2.50 **5.17** 7-8*

Figure 2: *^1H NMR 500 MHz Spectra of BDPP (12 mg/ml) in 99.96% D_2O as a Function of pH.*

Spectra are acquired using a homonuclear gated presaturation pulse (0.5 watts, decoupler offset centered on the HOD peak) applied during the recovery phase (utilizing 4 dummy scans) to achieve solvent suppression. Spectral acquisition parameters: 480 scans, 16 K data points, a pulse width of 6.5 µsec (approximating a 45^o pulse), a spectral window of 10 kHz, ^1H transmitter offset at 8500 Hz, and a Lorentzian apodization function of 1 Hz. All peak areas are scaled relative to the BDPP sample at pH 7.30. ^1H chemical shifts are referenced from TSS in D_2O.

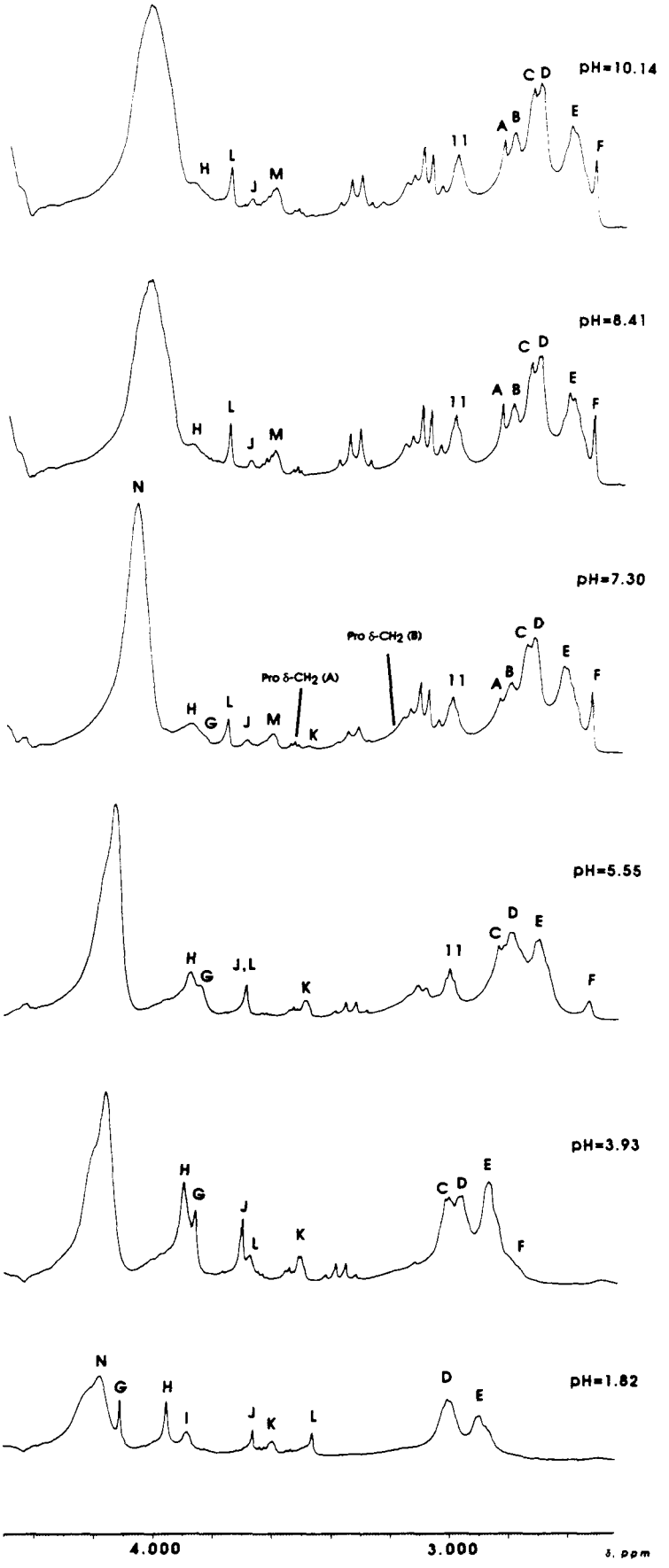

pH=10.14
H
L
J
M
11
A
B
C
D
E
F
pH=8.41
H
L
J
M
11
A
B
C
D
E
F
N
pH=7.30
Pro δ-CH2 (B)
Pro δ-CH2 (A)
H
G
L
J
M
K
11
A
B
C
D
E
F
pH=5.55
H
G
J,L
K
11
C
D
E
F
pH=3.93
H
G
J
L
K
C
D
E
F
N
G
H
I
J
K
L
D
E
pH=1.82
4.000
3.000
δ, ppm

Figure 3: *^1H NMR 500 MHz Spectra of BDPP as a Function of pH.*

Upfield regions of the same spectra presented in Figure 3 are presented here. ^1H chemical shifts are referenced from TSS in D_2O.

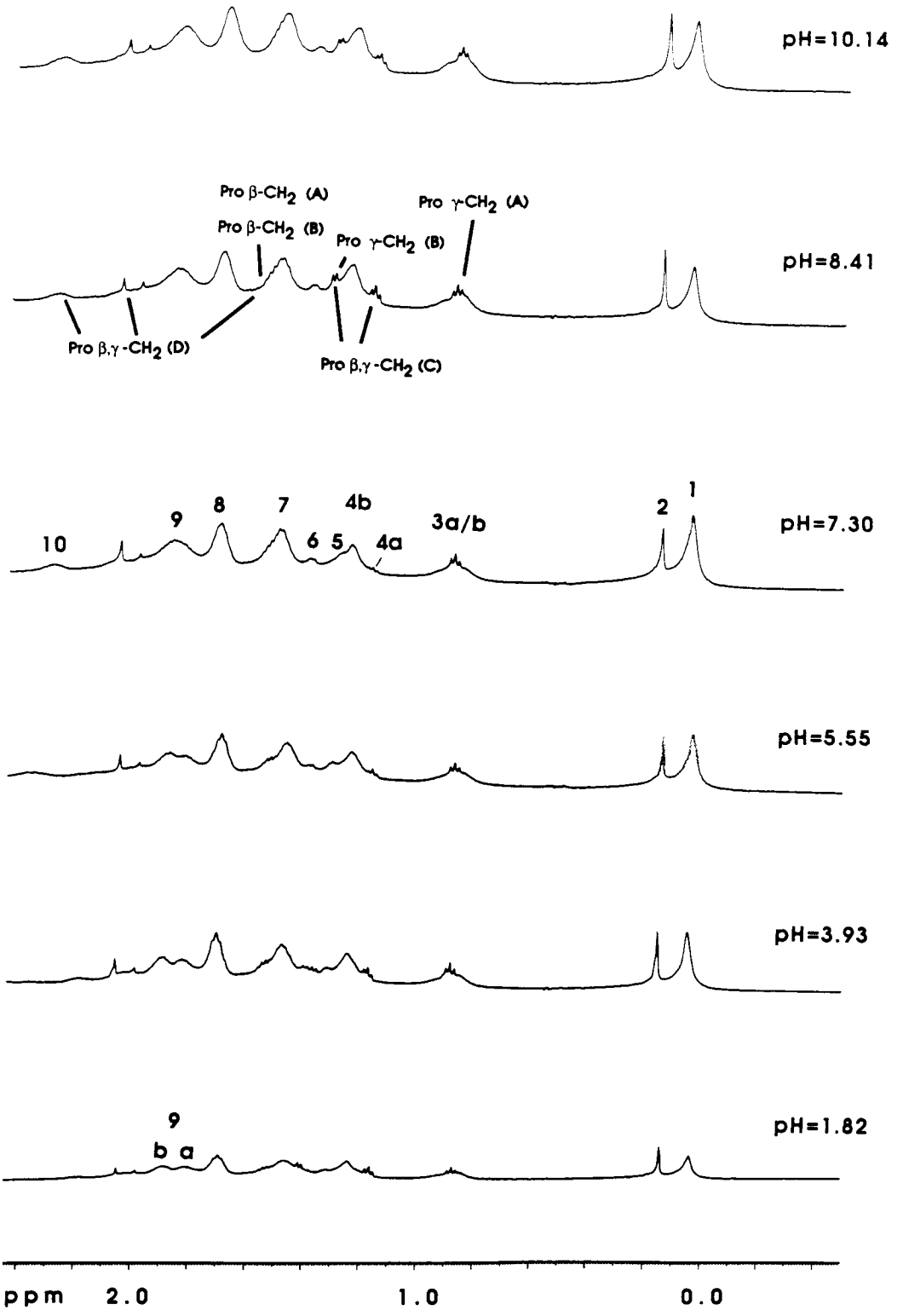

pH=10.14
Pro β-CH2 (A)
Pro β-CH2 (B)
Pro γ-CH2 (A)
Pro γ-CH2 (B)
pH=8.41
Pro β,γ-CH2 (D)
Pro β,γ-CH2 (C)
10
9
8
7
6
5
4b
4a
3a/b
2
1
pH=7.30
pH=5.55
pH=3.93
9
b a
pH=1.82
ppm 2.0
1.0
0.0

Figure 4: *2-D ^1H NMR 500 MHz HOHAHA Spin-Lock Correlation Experiment on a Sample of BDPP (12 mg/ml) in 99.96% D_2O at pH 7.22, 298 K.*

This TPPI phase sensitive contour plot (representing positive and negative levels) represents the upfield frequency region of a 2-D HOHAHA experiment. The experiment utilizes 2 dummy scans in the preparation phase and 512 scans per experiment, and the data represents the sum of 512 experiments. A 2 K x 2 K data matrix (4.883 Hz/pt in both dimensions) is created with zero filling in the F2 dimension, using a squared sine-bell window function in the phase sensitive mode. Transmitter ^1H 90^o pulse was 11 μsec, and decoupler 90^o pulse width is 60 μsec and 180^o pulse width is 120 μsec (at a decoupler power setting of 15 dB attenuation). Spectral window for F1 and F2 dimension is 5.0 kHz. A spin-lock time of 74 msec is used. ^1H chemical shifts are reported from 0.1 M TSS in D_2O.

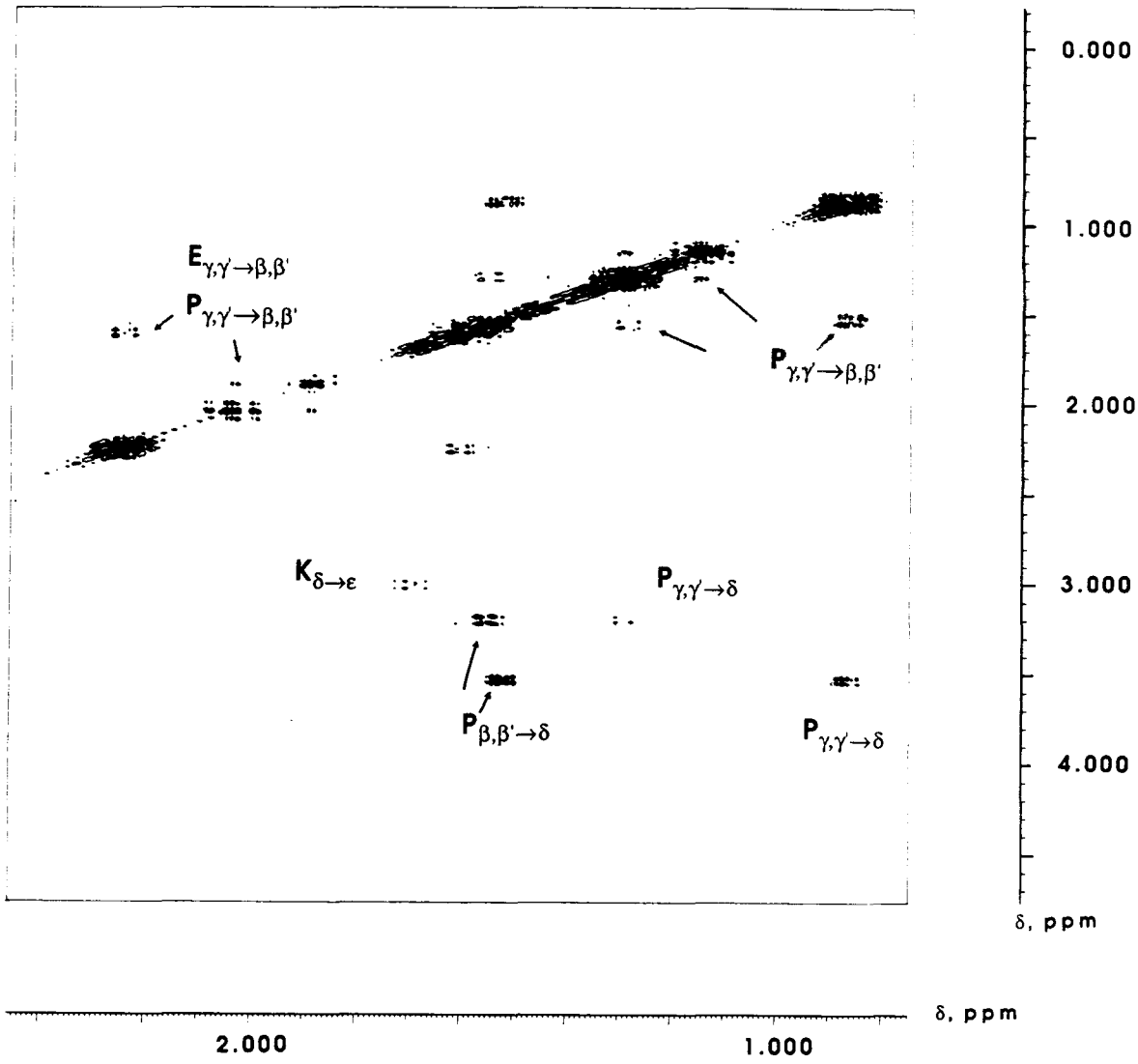

$E_{\gamma,\gamma' \rightarrow \beta,\beta'}$
$P_{\gamma,\gamma' \rightarrow \beta,\beta'}$
$P_{\gamma,\gamma' \rightarrow \beta,\beta'}$
$K_{\delta \rightarrow \varepsilon}$
$P_{\gamma,\gamma' \rightarrow \delta}$
$P_{\beta,\beta' \rightarrow \delta}$
$P_{\gamma,\gamma' \rightarrow \delta}$
0.000
1.000
2.000
3.000
4.000
δ, ppm
δ, ppm
2.000
1.000

Figure 5: *A Model for the Solution Conformation and pH-Dependent Folding of BDPP.*

Shown is a preliminary model of the BDPP protein at neutral pH (Structure IV). Structures I-IV describe BDPP folding transitions that occur at specific pH ranges.

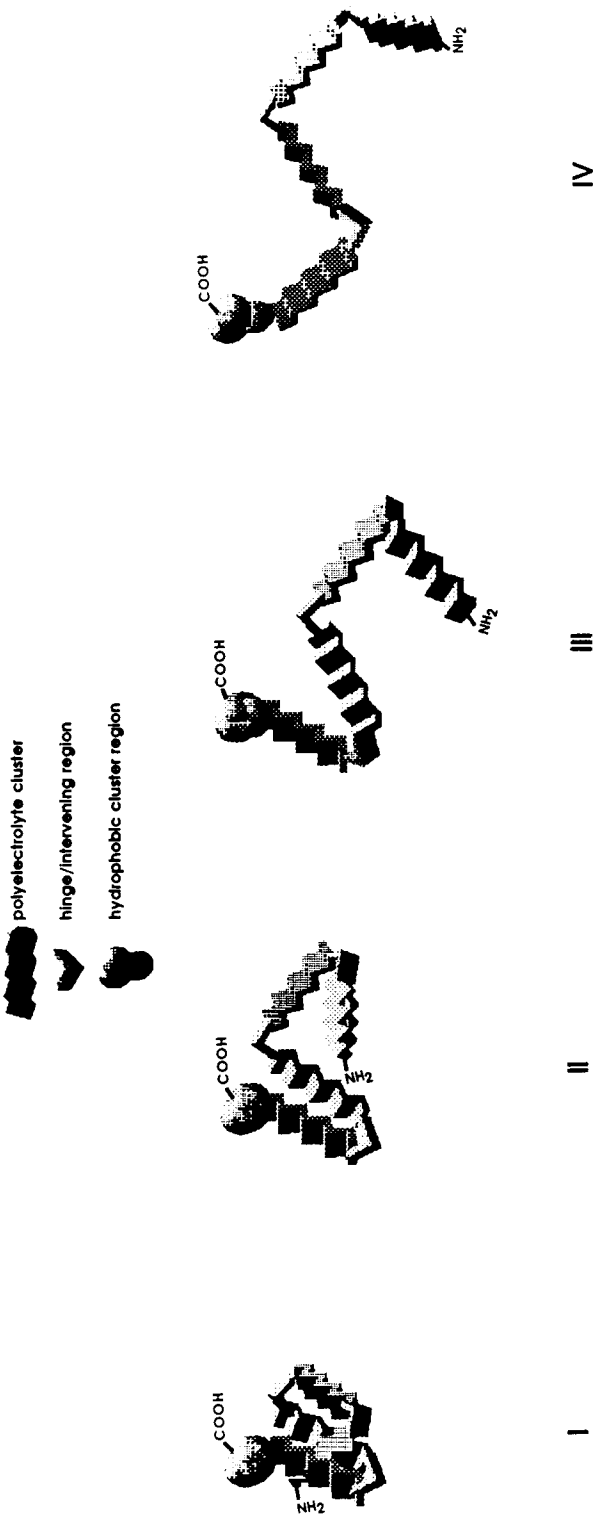

polyelectrolyte cluster
hinge/intervening region
hydrophobic cluster region
COOH
NH2
I
II
III
IV

Table II

Characterization of Hinge Region Resonances Affected by Neighboring Polyelectrolyte Sequences

Group	Peaks	Spin Systems	Neighboring PCBD
I	G, I, J	Ser, Gly	$(PSer)_n$
II	H, L	Ser, Gly	$(PSerAsp)_n$
III	K	Pro	$(PSerAsp)_n$
IV	A, 10	Ser, Gly, Glu	$(Asp)_n$

Six

NMR Conformational Studies of Phosphophoryn.

pH, Temperature, and Solvent Studies as a Probe for the Existence of Tertiary Structure in Polyelectrolyte Calcium Binding Domains.

Introduction

The polyelectrolyte mineral matrix proteins (PMMP) represent a class of proteins that contains one or more peptide clusters which possess a high negative charge density arising from the prevalence of Asp, Glu, *o*-phosphoserine (PSer), *o*-phosphothreonine (PThr), and *o*-phosphotyrosine (PTyr) residues in the cluster sequence. These amino acid sidechains are believed to act as oxygen ligand donors in the coordination of divalent cations (e.g., Ca^{2+}, Mg^{2+}). The role that these PMMP proteins play in Nature is that of templates in the biomineralization process. The interest in studying these proteins lies not only in elucidating the chemical mechanism for template-catalyzed inorganic complex formation, but also in investigating the unique feature of this fascinating class of proteins: namely, the peptide structure of anionic amino acid sequence clusters.

One example of this class of protein is phosphophoryn (bovine dentine isoform, 155 kD, abbreviated BDPP), a PMMP protein isolated from vertebrate dentine calcium phosphate mineral matrix. In earlier CD, NMR, and FT-IR investigations, the term "random coil conformation[1]" was used to describe the disorganized conformation of the anionic or Ca^{2+}-binding regions of phosphophoryn in the absence of Ca^{2+}, since the solution secondary structure of these regions did not fall into the categories of helical or sheet [Lee *et al.*, 1977, 1983; Cookson *et al.*, 1980; Zanetti *et al.*, 1981; Stetler-Stevenson and Veis, 1987]. Further refinement of phosphophoryn structure was achieved in recent studies which demonstrated that the phosphophoryn protein possesses three types of polypeptide cluster regions: polyelectrolyte calcium binding domain (abbreviated as PCBD)[Evans and Chan, 1991, 1992a,b; Sabsay *et al.*, 1991], hinge [Evans and Chan,

---

[1] In this context, random coil peptide structure is defined as "an ensemble of sidechain and backbone torsion angles which are independent of other bonds distant in the sequence, and where all conformations have comparable free energies" [Creighton,1990].

1992a,b], and hydrophobic [Rusenko *et al.*, 1991; Evans and Chan, 1992a,b]. Furthermore, the global conformation of the phosphophoryn molecule in the deprotonated state at pH 7.0 or above was found to exist in an extended or unfolded state [Evans and Chan, 1992a,c]. However, little is known about the local tertiary conformation of individual peptide regions in the phosphophoryn protein in the absence of Ca^{2+}, particularly in the PCDB regions. Are these anionic sequence regions "disorganized", or do they possess some degree of tertiary organization? It has been argued that even denatured, unfolded states for proteins involve very little random coil conformation, but in fact favor compactness of structure that promotes sidechain-sidechain interactions [Creighton, 1990]. Hypothetically, the cation binding or charged regions of a template should exist in some stabilized conformation that possesses specific spatial charge arrangement in three-space [Mann, 1988a,b]. This charge arrangement, i.e., tertiary structure[2], would provide three-dimensional control over nucleation and regulate the size of the nucleation sites.

In this present paper, we address the question of the local tertiary structure of the $(Asp)_n$, $(PSer\text{-}Asp)_n$, and $(PSer)_n$ polyelectrolyte cluster regions in the unfolded BDPP molecule. Our study utilizes two methods for probing conformational change in these regions: 1) Temperature and solvent denaturation of the polyelectrolyte cluster tertiary structure. This is accomplished by monitoring the changes in chemical shifts, linewidth and lineshape for PSer ^{31}P NMR resonances and non-exchangeable Asp and PSer sidechain ^{1}H NMR resonances which are several bonds removed from the peptide backbone in the protein [Evans and Chan, 1992a,c]. 2) pH-dependent perturbation of the polyelectrolyte cluster structure. By conducting pH titration experiments on BDPP in water, we are able to analyze pH-dependent effects for exchangeable Asp and PSer amide protons and thereby monitor peptide backbone conformational change for $(Asp)_n$, $(PSer\text{-}Asp)_n$, and $(PSer)_n$ clusters as a function of the protein folding state. As described herein, the three types of polyelectrolyte cluster regions exhibit

---

[2]Tertiary structure is defined here as the three-dimensional folding of secondary structures that occur within a defined polypeptide sequence.

some degree of local tertiary conformational folding or organization when phosphophoryn is in the extended global state.

Materials and Methods

Materials. The reagents used in the present study were obtained from the following sources: Deuterium oxide D_2O (99.9% *d*-atom) and *d*-acetonitrile (98% *d*-atom) from Cambridge Isotope Laboratories (Woburn, Massachusetts); deuterium chloride (DCl, 20% v/v solution, 98% *d*-atom) and sodium deuteroxide (NaOD, 30% v/v solution, 98% *d*-atom) from Aldrich Chemicals (Milwaukee, WI); 1N HCl and 50% w/v NaOH from Fisher Scientific. All glassware used in NMR sample preparation and analysis was first treated with EDTA/EGTA to remove unwanted divalent cations. Bovine mandibles (1-2 year-old steers) and their third molars were obtained fresh from a local slaughterhouse (Shamrock Packing Company, Vernon, California).

Isolation and Purification of BDPP for NMR Spectroscopy. The isolation and preparation of Ca^{2+} depleted BDPP from unerupted bovine third molars for the NMR experiments has been previously described [Evans and Chan, 1992a]. For NMR spectroscopy experiments involving temperature and solvent perturbation, the protein was solubilized in 99.9% *d*-atom D_2O (final concentration = 9.2 mg/ml for temperature experiments, approximately 7.8 mg/ml for solvent experiments). The pH was adjusted to 7.0 with microliter volumes of NaOD or DCl to minimize ionic strength effects, uncorrected for deuterium isotope effects.

One-Dimensional NMR Spectroscopy Experiments in D_2O. ^1H (500.138 MHz) and ^{31}P (202.458 MHz) one-dimensional NMR spectroscopy were performed on a Bruker AM-500 Spectrometer at a field strength of 11.7 T, with data processing performed on an Aspect 3000 computer. A Bruker low-gamma 5 mm broadband inverse probehead was utilized in all D_2O experiments. For experiments in D_2O, ^1H NMR acquisition parameters utilize a flip angle of 45° (6.5 μs), a spectral window of 10 kHz, 16 K data points, a digital resolution of 1.221Hz/pt, a relaxation delay of 1 second, and 640 scans (solvent experiment) or 1200 scans (temperature experiment). To achieve suppression

of the residual HOD solvent signal, a decoupler presaturation pulse (5 w) was applied during the recovery delay period. ^{31}P NMR acquisition parameters include a flip angle of 45° (8.3 μs), a spectral window of 5 kHz, 8 K data points, a digital resolution of 1.221 Hz/pt, a relaxation delay of 2 seconds, and 2400 scans. Proton chemical shifts are reported relative to external 2,2,2-trimethyl 2-silapentane sulfonic acid (TSS, 0.1 M in D_2O); phosphorous chemical shifts are reported from external EHDP (2-hydroxyethyl-diphosphonate), 0.5 M in D_2O, pH 7.0, with the ^{31}P centerpeak resonances at 10.55 and 10.48 ppm relative to 85% H_3PO_4 at 0.0 ppm).

Temperature Experiments. Temperature experiments were conducted using the Bruker variable temperature control unit, and liquid nitrogen boil-off vapor for below-ambient temperatures and dry nitrogen gas for above ambient temperatures. The temperature values given are those of the sample solutions in the NMR tubes as measured by the probehead thermocouple. Calibration of the temperature control unit was acheived by measuring the proton $\Delta\delta$ for hydroxyl and aliphatic C-H resonances for neat methanol (273 K to 310 K) and neat ethylene glycol (290 K to 383 K) for the temperature ranges specified. The temperature experiment was initiated by placing the BDPP sample (5 mm glass tube, teflon stopcock sealed, 400 μl final volume) in the probehead and equilibrating the sample for 15 min at the starting temperature. After equilibration at this temperature, room temperature shims and probehead Q tuning were optimized on the BDPP sample, and ^{1}H NMR FID acquisition commenced, followed by ^{31}P NMR data acquisition. In this manner, NMR experiments covering the temperature range of 4° to 90° C were accomplished with one BDPP sample. After the 90° C experiment, the BDPP sample was slowly cooled to room temperature and maintained at this temperature for 48 hrs, after which time room temperature $^{1}H/^{31}P$ NMR spectra were acquired. ^{1}H and ^{31}P NMR chemical shift references were obtained at each temperature and acquisition setting as per the BDPP/D_2O sample.

Solvent Exchange Experiments. d-acetonitrile (99.6% atom-d) was added in specific volumetric ratio to separate BDPP samples in D_2O. After acetonitrile addition, the pH was checked, and if necessary the pH was readjusted to 7.1 using microliter volumes of either DCl or NaOD in D_2O. The final sample

volume was 400 ± 15 μl. ^{1}H and ^{31}P NMR spectra were acquired at 298 K as described above in the NMR Spectroscopy section. After acquiring NMR spectra for the endpoint acteonitrile/D_2O addition (70% v/v), this particular BDPP sample was then subjected to <u>in situ</u> freeze-lyophilization in the NMR tube to remove the solvent. The lyophilized protein sample was then reconstituted in 99.8% D_2O, the pH was adjusted to 7.1, and $^{1}H/^{31}P$ NMR spectra were acquired for this reconstituted form of the BDPP protein

One and Two-Dimensional ^{1}H NMR Experiments in H_2O. NMR experiments were conducted utilizing a dedicated 500 MHz ^{1}H NMR 5 mm probehead. For 1-D and 2-D ^{1}H NMR experiments, solvent suppression was accomplished by using presaturation during the relaxation delay. BDPP protein was reconstituted in deionized distilled water (DD_{water}) (final concentration 23.2 mg/ml) in the absence of buffer salts with 10% v/v D_2O added to the sample to maintain the field frequency lock. For pH titration experiments, the pH was adjusted with microliter volumes of either NaOH or HCl in DD_{water} utilizing single pass titration methods to minimize ionic strength effects [Evans and Chan, 1992a,c]. pH measurements were made before and after each NMR experiment, using an Ingold 6030-02 glass microelectrode and Radiometer PHM83 Autocal pH meter. All pH units are reported without correction for deuterium isotope effects.

Two-Dimensional Homonuclear Spin-Lock Correlation NMR Spectroscopy Experiments in H_2O. Assignments for BDPP Asp and PSer amide NH proton resonances at pH 7.10 were achieved via homonuclear Hartmann-Hahn spin-lock correlation experiments (HOHAHA)[Bax and Davis, 1985a,b]. Acquisition parameters for the HOHAHA experiments included a relaxation delay of 1 sec.; a spectral window of 6 kHz in both dimensions; a 90° and 180° pulse of 26.5 μsec and 53 μsec , respectively (at a decoupler power setting of 12 <u>w</u>); SL_x trim pulses of 2.5 msec; 2 K data points, a spectral window of 6 kHz, 256 experiments, 320 scans per experiment, and a mixing time of 75 msec. Solvent suppression was achieved using a presaturation decoupler pulse (5 <u>w</u>) during the recovery phase of each experiment. Zero-filling in the F2 dimension was performed to create a 1 K x 1 K data matrix using a phase-sensitive (positive and negative levels) squared sine-bell window function in

both dimensions with no shift. Other parameters utilized in the 2-D NMR spectroscopy are described in the Figure legends.

Results

From the known amino acid composition of BDPP [Stetler-Stevenson and Veis, 1983; Rahima and Veis, 1988; Sabsay *et al,* 1991], the BDPP N-terminal amino acid sequence [Evans and Chan, 1991], and the partial amino acid sequence recently acquired for rat incisor phosphophoryn [Sabsay *et al.,* 1991], ≈ 80% of the phosphophoryn protein molecule is comprised of negatively charged amino acids. Our interest is in the tertiary organization of the Asp and PSer-containing PCBD regions in BDPP in the absence of divalent cations. For the sake of discussion, we will assume that the BDPP protein exists as a net negatively charged species at pH 7.0 that experiences counterion condensation (by Na^+ in our system), and possesses some degree of effective charge screening of these negative charges[3] [Manning, 1978; 1969]. One means of probing the organization of polyelectrolyte regions is via denaturation. Temperature denaturation (low and high temperature) induces conformational transitions in the structure of a protein [Freire *et al.,* 1992; Friere and Murphy, 1991; Murphy and Gill, 1991], which can be observed by ^1H and ^{31}P NMR. A second method of denaturation involves exchanging the bulk water solvent for a polar aprotic solvent that perturbs the structure of the solvent-accessible, hydrophillic polyelectrolyte clusters. Acetonitrile was chosen for this investigation for the following considerations: first, it is very miscible in water, and secondly, the residual singlet proton resonance arising from CD_3CN occurs in a chemical shift range (2.00 ppm) which does not overlap with any of the Asp, PSer, or non-Asp/PSer sidechains proton resonances in BDPP. From our previous NMR studies at pH 7.0 the divalent cation-depleted form of the BDPP protein exists in an unfolded or extended global conformational state [Evans and Chan, 1992a,c]. Thus, as we denature the protein, we wish to examine how stable (or unstable) the local tertiary

---

[3] The Cl^- ion in our system will also form ion pairs with a certain percentage of His residues and Lys residues that are positively charged at this pH [Manning, 1978, 1969].

structure of $(PSer)_n$, $(Asp)_n$, and $(PSer\text{-}Asp)_n$ polyelectrolyte regions are in the absence of divalent cations.

Temperature Experiments.

Temperature Effects on the Conformation of Polyelectrolyte Cluster Regions in BDPP at pH 7.0. From our previous work at 25° C, pH 6.94, Peaks B-F are assigned to the β, β'-CH_2 diastereotopic protons of Asp residues in BDPP [Evans and Chan, 1992a]. In addition, Peaks 12-16 are also assigned to Asp β, β'-CH_2 protons and the β, β' CH_2 and α, α' -CH protons of Ser and Gly residues, respectively. As shown in Table 1, each of the Asp β, β' CH_2 proton resonance frequencies demonstrates a linear response to temperature over the range of 4° C to 90° C in D_2O. For Peaks F, E, 12, and 15, there is a small downfield (+Δδ) resonance frequency shift with temperature (0.8 - 1.9 Hz/°C); for Peaks B, C, D, and 14 there is a small upfield (-Δδ) resonance frequency shift as a function of temperature (0.6 - 1.6 Hz/°C). We were unable to accurately determine the temperature-dependent frequency shifts for Peaks 14 and 16. Thus, the Asp β, β'-CH_2 resonances in BDPP exhibit a minor perturbation with temperature. Peak N, which is assigned to the PSer β,β' CH_2 diastereotopic proton resonances of BDPP, exhibits a temperature-dependent downfield chemical shift effect (Table 1).

In contrast to the ^1H NMR BDPP chemical shifts for PSer and Asp β,β'-CH_2 protons, the ^{31}P chemical shifts for BDPP PSer resonances (denoted as α, β, and χ, [Evans and Chan, 1992a]), exhibit a non-linear relationship with temperature (Figure 1). The temperature-dependent shift effect observed for each of the PSer ^{31}P NMR resonances encompasses a frequency range of nearly 1.5 ppm. The best fit for each of the three ^{31}P NMR resonances is a second-order polynomial (of the form $y = Ax^2 - Bx - C$) with positive slope (Figure 1).

The temperature dependence of the ^1H NMR linewidth and lineshape of BDPP Asp β, β'-CH_2 resonances were also examined (Figure 2). Spectral changes were observed by both increasing and decreasing the temperature from 25° C, demonstrating the change in the tertiary structure associated with

heat denaturation and cold denaturation, respectively. For Asp sidechains in BDPP, the transition temperature for cold denaturation (denoted as T_c) occurred between 15° C and 4° C, where noticeable differences in the lineshape of Peaks C, D, and E and their corresponding shoulder regions are observed (Figure 2). Since the hydrophillic Asp sidechains would be expected to interact significantly with water solvent, the small perturbation in the ^{1}H NMR lineshape at lower temperature may reflect temperature-dependent solvent reorganization (e.g., changes in hydrogen bonding distances and hydrogen bonding donor-acceptor geometry), and concommittantly, reordered sidechain-solvent and sidechain-sidechain interactions. Three transition temperature points for thermal denaturation (denoted as T_h1,2,3) are observed for the BDPP Asp residues. (T_h1): As the temperature progresses from 25° C to 42° C, there is a decrease in signal intensity for Peaks F and 13-16. At the 42° C, Peak 12 is no longer observable in the BDPP ^{1}H NMR spectrum. (T_h2): Peaks 13 and 14 experience linewidth broadening and assimilate as a single resonance at 3.114 ppm at 70° C (Figure 2). Peaks 15,16, and F are observed to be diminished in signal intensity at this temperature. We note that these changes in the ^{1}H NMR BDPP spectrum were irreversible at or above 65° C. (T_h3): The ^{1}H NMR BDPP spectrum exhibits several important temperature-dependent effects over the range of 70° - 90° C. The BDPP Asp β, β-CH_2 protons are observed to broaden and form three major resonances centered at 2.648 ppm, 2.702 ppm, and 2.817 ppm (all possessing minor shoulder peaks). Peaks A and B experience a signal intensity increase of nearly 45%. Peaks C, D, and E are reduced in signal intensity by a factor of 1.5, and Peak F has disappeared from the ^{1}H NMR spectrum. At 90° C, the resonance at 3.114 has broadened considerably and is reduced in signal intensity. Thus, the heat denaturation of BDPP at pH 7.0 results in three tertiary structure transitions in Asp-containing polyelectrolyte regions of the protein.

Temperature denaturation effects on the tertiary structure of the PSer-containing regions of BDPP were also examined. No significant spectral changes were noted for the β, β'-CH_2 protons of PSer residues from 4° C to 65° C (data not shown). Above 65° C, the PSer β, β-CH_2 resonance overlaps with the HOD solvent peak and thus no further ^{1}H NMR spectral information

regarding the PSer methylene protons could be obtained. However, temperature denaturation effects on the ^{31}P NMR BDPP PSer α, β, and χ resonance lineshapes and linewidths (Figure 3) proved to be more informative. The sensitivity of the phosphorus spins to alterations in the O-P-O bond distances, bond and torsion angles and solvent interactions are influenced to a great degree by temperature. No significant ^{31}P NMR spectral changes were noted for the PSer ^{31}P resonances at temperatures lower than 25° C. However, as for the Asp β, β-CH_2 protons, three transitions in the tertiary structure of the PSer residues are observed, *viz* (T_h1): In the temperature range of 25 - 42° C, an inverse relationship between ^{31}P linewidth and temperature is observed, i.e., a 33% decrease in linewidth. (T_h2): In the temperature range of 60-70° C, a previously unobserved ^{31}P NMR resonance (denoted as γ, with a $\Delta\nu_{1/2}$ of 20 Hz) appears at 4.6340 ppm. (T_h3): Between 70-90° C, Peak γ becomes more prominent (representing 42% of the total observable ^{31}P spins in the BDPP sample), Peak α experiences a decrease in linewidth, and Peaks β and χ are no longer discernable in the ^{31}P NMR spectrum. Upon slow cooling, the BDPP ^{31}P NMR spectrum (top, Figure 3) was found to be comprised of two overlapping ^{31}P spin contributions: the first, a narrow resonance ($\Delta\nu_{1/2}$=16 Hz) centered at 2.808 ppm, and the second, a broad resonance ($\Delta\nu_{1/2}$ = 76 Hz) superimposed upon the narrow resonance. Both resonances are in the reported chemical shift range for *o*-phosphoserine at this pH [Evans and Chan, 1992a; Lee *et al.*, 1977; 1983; Cookson *et al.*, 1980]. The total peak area of the ^{31}P spins in the renatured sample is equivalent (within 6% error) to the peak area of the ^{31}P spins in the BDPP sample at room temperature. From these observations, we conclude that BDPP PSer-containing polyelectrolyte regions undergo three tertiary conformational transitions as a result of heat denaturation to the protein. These three transitions (T_h1, T_h2, T_h3) occur in temperature ranges similar to those observed for the Asp-containing polyelectrolyte regions. Thus, both Asp and PSer-containing polyelectrolyte regions in divalent cation depleted BDPP may exist in a folded or stabilized tertiary conformation at room temperature.

pH Titration Experiments in Water

Assignment of BDPP Amide NH Resonances. Protein backbone amide proton assignments (Table 2) were determined from the 2-D ^1H-^1H HOHAHA experiment [Figure 4] of BDPP in H_2O (with 10% D_2O), at pH 7.10. The 1-D NMR spectrum indicates that BDPP possesses two major amide proton resonances (denoted as Peaks π and δ, respectively), and four other broad amide proton resonances of weaker signal intensity (Peaks ρ, ψ, ϕ,and ω)(Figure 4). Due to hydrogen-deuterium isotope exchange, these labile proton resonances were not originally observed in our earlier ^1H NMR BDPP experiments conducted in D_2O solvent [Evans and Chan, 1992a,c]. Although a concentrated BDPP protein sample was used (23.2 mg/ml), we were unable to detect any other BDPP amide NH resonances in the frequency range of 8.0 to 11.0 ppm.

The HOHAHA spectrum (Figure 4) reveals that the two major amide resonances, π, and δ, exhibit *J*-scalar couplings with α-CH protons at 4.800 to 4.382 ppm, and with β-CH_2 protons at 2.500 to 2.900 and at 4.000 to 4.100 ppm. Relayed *J*-scalar couplings are observed between Peak δ and resonances at 2.500 to 2.900 ppm, which correspond to BDPP Asp β-β'-CH_2 diastereotopic protons [Figure 4][Evans *et al.*, 1992a]. The corresponding Asp α-CH protons are observed at 4.500 to 4.800. Interestingly, the weak crosspeaks observed between Peak δ and the protons resonating at 4.000 ppm arise as a result of long-range couplings with the PSer β, β'-CH_2 protons, since it has been established that in BDPP there are no Asp α-CH protons resonating at 4.000 ppm [Figure 4][Evans *et al.*, 1992a]. Thus, Peak δ is assigned to BDPP Asp amide NH protons, with overlapping contributions from PSer amide NH proton resonances. Peak π exhibits relayed four-bond *J*-scalar couplings with β-CH_2 proton resonances at 3.900 to 4.200, and three-bond *J*-scalar couplings with α-CH proton resonances at 4.500 to 4.800 ppm. Thus, Peak π is exclusively assigned to BDPP PSer amide NH proton resonances.

Further analysis of the 2-D HOHAHA spectrum indicates that Peak ρ demonstrates $^3J_{H\text{-}H}$ and $^4J_{H\text{-}H}$ relayed connectivities exclusively with α-CH protons and PSer β, β'-CH_2 sidechain protons, respectively [Figure 4]; this peak is assigned to PSer NH amide protons from Peak N. Peak ψ exhibits weak

$^3J_{H-H}$ scalar couplings with α-CH protons at 4.520 to 4.477 ppm [Figure 4]. No amide NH $\rightarrow$ α-CH connectivities were observed for Peak ϕ. Peaks ϕ and ψ are therefore assigned to BDPP Asp NH amide protons, based in part on the HOHAHA spectrum and on pH titration behavior of these resonances (see below). Due to the weak intensity of the Peak ω resonance, we were unable to detect any scalar connectivities between this resonance and α-CH or sidechain proton resonances in BDPP. However, as discussed below, the pH-dependent behavior of this resonance indicates that it can be tentatively assigned to Asp NH as well.

pH-Dependent Peptide Backbone Conformational Change for Polyelectrolyte Cluster Regions. The amide proton chemical shift-pH titration profile for BDPP Asp and PSer residues [Figure 5] yields additional information regarding the chemical shift contributions of these exchangeable amide resonances, and, details the pH-induced folding transitions that occur in the PCBD regions of BDPP. Over the pH range investigated (pH 2.08 to 9.00), the titration profiles for Peaks π, δ, ρ, ψ, ϕ, and ω reveal that there are three distinct BDPP peptide backbone conformational change transitions. (1) The first conformational transition is observed for the *o*-phosphoserine monoester protonated-to-monobasic transition for PSer-containing regions in BDPP (pH 2.93, curve ψ). (2) The second conformational transition corresponds to the BDPP Asp carboxylate protonation/deprotonation event (Peaks δ, ψ, and ω). Peaks ψ and ω exhibit conformational transitions corresponding to an "observed" pK_a of 4.6 and 5.5, respectively. Peak δ exhibits a broad conformational transition centered at pH 4.4. (3) Perhaps the most significant transition in BDPP conformational folding occurs in the phosphate monoester monobasic-to-dibasic transition (pH range 7.4-8.0, curves π, ρ, ψ, ϕ, and ω). The presence of a conformational transition corresponding to an "observed" pK_a of 7.9, 7.6, and 7.4 for Asp amide NH proton resonances ω, ϕ, and ψ, respectively, clearly indicates that these Asp residues are predominantly located in the $(PSer\text{-}Asp)_n$ heteropolymer sequence regions in BDPP [Evans and Chan, 1992a]. The absence of this conformational transition for Peak δ implies that this amide NH resonance originates from Asp residues that reside in $(Asp)_n$ sequence repeat regions. Collectively, the polyelectolyte sequence regions of BDPP undergo

peptide backbone folding transitions as a function of the charge repulsion/charge shielding of the amino acid sidechain groups that reside in these regions.

Solvent Experiments

Solvent Denaturation Effects on the Conformation of Polyelectrolyte Cluster Regions in BDPP. Acetonitrile addition resulted in several changes in the ^1H NMR spectra of the BDPP protein (Figure 6). The Asp β, β'-CH_2 protons (Peaks B-F and 12-16), Ser β, β'-CH_2 and Gly α, α' protons (Peak A; Peaks 12-16), and Lys ε-CH_2 protons (Peak 11) manifest significant solvent effects at CD_3CN : D_2O volume ratios of 50% v/v or greater. At 50% v/v CD_3CN, we observe the following: (1) Peaks A, B, and C are reduced in signal intensity relative to the centerpeaks D and E. (2) Minor peaks are apparent on the upfield and downfield shoulder regions of centerpeaks C and E. (3) Peak F exhibits a 25% decrease in signal intensity and peak area compared to its original value in the absence of acetonitrile. (4) The multiplet signal for Asp β, β' CH_2 protons (Peaks 12-16) transforms into a doublet signal (Peaks 13 and 14) that is flanked upfield and downfield by two resonances of minor intensity (Peaks 12 and 15). These peaks experience a small solvent-induced upfield chemical shift effect (average $\Delta\delta$ = -0.022 ppm). (5) The Lys ε-CH_2 protons of BDPP has increased in signal intensity by 25%, and are resolvable as a centerpeak with an prominent upfield shoulder region. As the acetonitrile percentage increases from 50 to 70% v/v, Peaks A-E, 12, and 15 exhibit line broadening effects due to changes in spin-spin relaxation and/or chemical shift overlap. The Lys ε-CH_2 protons (Peak 11) exhibit a decrease in linewidth of approximately 100%, and appear as a narrow singlet resonance at this acetonitrile concentration. At this concentration of acetonitrile, the BDPP solution adopted an opaque appearance. Upon removal of the bulk solvent by lyophilization and reconstitution of the protein sample in 99.9% D_2O, pH 7.0, we observed that the ^1H NMR spectrum generated by this reconstituted sample was nearly identical to the ^1H NMR spectrum for the original, undenatured BDPP sample [Figure 6], with the only noticeable differences being (1) the increased signal intensity ($\approx$ 40% increase in proportion to the other resonances in the reconstituted sample) for Peak A in the ^1H NMR

spectrum of the reconstituted BDPP sample in D_2O, and (2) the reduction in total peak area ($\approx$ 35%) for the Asp β, $\beta'CH_2$ and Lys ε-CH_2 protons in the ^1H NMR spectrum of the reconstituted BDPP sample.

In the presence of acetonitrile, the ^{31}P NMR spectrum of BDPP demonstrates the effects of solvent exchange on the PSer residues in the protein [Figure 7]. As the acetonitrile concentration increases from 0% to 30% v/v, ^{31}P chemical shift overlap effects are observed : Peaks β and χ shift upfield by -0.1989 and -0.3617 ppm, respectively, and exhibit overlap with Peak α. At 50% v/v acetonitrile, the majority of the PSer ^{31}P spins (92%) experience a downfield shift of -0.5728 ppm relative to the α resonance of the native BDPP sample. However, 8% of the total ^{31}P NMR signal is observed as a minor resonance centered at 2.872 ppm (denoted as λ), which at pH 7.0 is characteristic for the monophosphate ester phosphorous species [Evans and Chan, 1992a; Lee *et al.*, 1977; 1983; Cookson *et al.*, 1980]. At 70% v/v acetonitrile addition, four ^{31}P PSer resonances are observed (4.591 ppm, a broad resonance centered at 4.235 ppm, a minor resonance, λ_1, at 3.005 ppm, and a corresponding shoulder peak, λ_2, at 2.977). Upon reconstitution in D_2O, the ^{31}P NMR spectrum for BDPP closely resembles that of the native BDPP sample [Figure 7], with one notable difference: compared to the native protein sample spectrum, the observable ^{31}P NMR signal for the reconstituted BDPP sample is reduced by a factor of 5.

The effects of acetonitrile on the BDPP protein demonstrate that the polyelectrolyte Ca^{2+} binding regions experience solvent-induced local tertiary conformational transitions. These transitions are marked by chemical shift overlap and spin-spin relaxation effects in the NMR spectra, and may arise from alterations in the solvent environment around Asp, Lys, and PSer residues in certain regions of the BDPP protein.

Discussion

Mann [1988a,b] has proposed that template macromolecules exist in a topological state where spatial charge distribution favors regions of high charge density. For a protein template, the primary amino acid sequence

would create a linear sequential clustering of sidechains. The question is, is this sequence further organized into three-dimensionally organized clusters of high negative charge density in the absence of divalent cations? In the case of a random-coil conformation, this would not be the case: there would be little or no local secondary or tertiary organization [Figure 8]. However, if the negatively charged Ca^{2+} binding regions existed in an organized or folded local state that minimizes sidechain charge-charge interactions, this would increase the negative charge density in each region, and provide an effective three-dimensional surface for nucleation [Figure 8]. This could exist even though the global conformation of the BDPP protein is in the extended state. From the data in our present report, it is clear that BDPP possesses tertiary folding in each of the three polyelectrolyte regions: $(Asp)_n$, $(PSer\text{-}Asp)_n$, and $(PSer)_n$. This is supported by the ability to denature the tertiary structure of the Asp and PSer-containing regions in the BDPP protein *in the absence of divalent cations* at a pH where most of the protein exists in an extended state [Evans and Chan, 1992a,c]. This suggests that these polyelectrolyte regions do not exist as disorganized random coil conformers in the absence of divalent cations. Conformational heterogeneity in PCBD regions was established in our earlier report [Evans and Chan, 1992a], where α-CH protons for PSer and Asp residues were observed over a broad frequency range. The observation of three denaturation transitions at elevated temperatures (42° C, 60-70° C, 90° C)[Figure 2,3], of three solvent-induced denaturation transitions (30%, 50%, 70%), and the inability of the BDPP protein to refold into its original room temperature conformational state, suggests that the folding or organization of polyelectrolyte Ca^{2+} binding regions are unique, and are somehow stabilized in the absence of divalent cations. Further evidence for stable tertiary folding of $(Asp)_n$, $(PSer)_n$ and $(PSer\text{-}Asp)_n$ containing polyelectrolyte regions was noted by the presence of pH-dependent backbone folding transitions [Figure 5]. The $(Asp)_n$ polyelectrolyte regions showed a broad peptide backbone conformational transition centered near the pK_a's for Asp sidechains in BDPP (pH 4-5.5). However, a more significant conformational transition was observed for PSer-containing polyelectrolyte regions in BDPP occurred in the pH range of 7.4 to 8.0, corresponding to the monobasic-to-dibasic protonation transition of PSer monophosphate esters. From our earlier work, we know the following:

(1) the divalent cation depleted BDPP protein exists in an unfolded state in this pH range, i.e., the polyelectrolyte regions are oriented away from one another due to charge repulsion [Evans and Chan, 1992a,c]. (2) Correcting for deuterium isotope effects, all the Asp sidechains exist in the fully deprotonated state in this pH range. Therefore, the observation of pH-dependent backbone transitions suggests that there is PSer-induced localized unfolding *within each PSer-containing polyelectrolyte region* in the BDPP protein. This could only occur if each PSer-containing region, at neutral pH, possessed some stabilized tertiary structure that became disrupted due to PSer-PSer charge repulsion.

The presence of stabilized local tertiary folding of anionic clusters presents some intriguing aspects. Goto and Fink [1989], in their studies of β-lactamase, have shown that salt stabilizes the native structure of the protein when the molecule is highly charged. Thus, the counterion condensation on the surface of BDPP may permit a clustering of the polyelectrolyte subregions if the cations shield the charge repulsions effectively, and, if the cations can penetrate into the structure of the protein [Stigter and Dill, 1990; Stigter *et al.*, 1990]. In addition to charge shielding, there may be other forces that counteract the destabilizing effects of sidechain charge repulsion as well. In several recent studies of intermediate molten globule form of cytochrome *c* [Goto and Fink, 1989, 1990; Goto *et al.*, 1990a,b; Goto and Aimoto, 1991; Goto and Nishikiori, 1991; Hughson *et al.*, 1991], it has been shown that the transitions between conformational states of cytochrome *c* are determined by a balance between electrostatic repulsions (i.e., between Lys residues) and other opposing forces (in this case, hydrophobicity). Thus, one way in which the PMMP template macromolecule can stabilize its conformational state to attain high negative charge density is to offset the destabilizing charge repulsion forces with other stabilizing forces that lower the overall free energy of the peptide conformation ($\Delta G_{conformation}$). Several forces could stabilize the peptide structure of the polyelectrolyte regions against ineffectively screened charge repulsion effects between anionic sidechains:

(a) *Hydrogen Bonding Network Effect:* At pH 7.0 in D_2O, the PSer residues in BDPP exist in an approximately 50-50 mixture of the monoanionic

form and dianionic form. Thus, one possible stabilizing force might involve hydrogen bonding between PSer HPO_4^- sidechains and other hydrogen bonding acceptors such as H_2O, COO^- and monophosphate PO_4^{-2}. In an earlier $^1H/^{31}P$ NMR study conducted on the 30 kD rat incisor phosphophoryn peptide fragment [Cookson *et al.*, 1980], it was hypothesized that a hydrogen bonding "network" existed between protonated PSer and the deprotonated Asp sidechains at neutral pH. This "network" phenomenon was used to explain both the cooperativity effect observed for the phosphophoryn ^{31}P PSer titration, and, the non-Henderson-Hasselbach titration behavior of 80% of the Asp and Glu sidechains which did not exhibit a simple pK_a [Cookson *et al.*, 1980]. A similar pK_a "effect" was also observed in our initial studies conducted on BDPP, where the sidechain pK_a's for Asp and PSer exhibited an alkaline shift from pK_a values reported for the free amino acids [Evans and Chan, 1992a]. Cooperative binding isotherms have been reported for proton binding in clusters of ionizable groups [Saroff, 1991]. Thus, at physiologic pH, the PSer-containing polyelectrolyte domains may exhibit monobasic-dibasic PSer hydrogen bonding. However, this stabilizing force would be unavailable to $(Asp)_n$ regions of BDPP, and thus the *Network Effect* would not help to explain the temperature stability of $(Asp)_n$ regions in BDPP

(B) *Interresidue Hydrogen Bonding.* A second stabilizing force, would be the presence of interresidue hydrogen bonding between amide NH and carbonyl oxygen atoms throughout the polyelectrolyte peptide sequence. Ben-Naim [1991] has calculated that intramolecular hydrogen bonding in the presence of solvent contributes favorably to the overall free energy of the folding process by retaining the solvation free energy. Thus, the formation of intramolecular hydrogen bonds in the solvent-accessible polyelectrolyte regions might stabilize the tertiary folding of these regions. Typically, intramolecular hydrogen bonding in polypeptides results in the formation of a defined secondary structure, e.g., α-helix, parallel or antiparallel β-sheet, or β-turn [Richardson, 1981; Wilmot and Thornton, 1988, 1990; Richardson *et al.*, 1975, 1977, 1981; Chou *et al.*, 1990; Sibanda *et al.*, 1989; Milner-White, 1990; Thomas, 1990; Tramontano *et al.*, 1989]. At this time we are unable to define the secondary structure(s), if any, assumed by BDPP $(Asp)_n$, $(PSer)_n$, and $(PSer\text{-}Asp)_n$ anionic cluster regions in solution. However, this hypothesis is an

attractive one, in that it would permit stabilization of any solvent accessible region of the polyelectrolyte sequence. One approach to determine the feasibility of polyelectrolyte sequence intrastrand hydrogen bonding is to investigate the solution structure of synthetic $(Asp)_n$, $(PSer\text{-}Asp)_n$ and $(PSer)_n$ peptides, and, to model the folding of these peptides using molecular dynamics and conformational search algorithms. These experiments are currently in progress.

(C) *Salt Bridging.* A third stabilizing force would be the formation of salt-bridges between Lys and Asp and/or PSer residues in the BDPP protein. Previously, we had noted the pH-dependent behavior of Lys residues in BDPP, in which two Lys conformers existed below pH 7.0, and became equivalent above this pH [Evans and Chan, 1992a,c]. Since pH 7.0 marks the midpoint of the monobasic-to-dibasic transition for PSer residues in BDPP in D_2O [Evans and Chan, 1992a], it is plausible that the PSer residues may electrostatically interact with the positively charged Lys amino groups, and that this interaction may change as the PSer residues transition between dianionic and monoanionic forms during the pH titration. Although ion pairing in proteins typically involves a small proportion of residues and thus does not contribute a large stabilizing energy (1-3 kcal/mol/pair [Dill, 1990; Fersht, 1972; Perutz and Raidt, 1975]), intraprotein salt-bridging may stabilize small PSer-containing regions into high charge density regions that are prerequisite for nucleation.

(D) *Hydrophobic Interactions.* Finally, there should be some mention of hydrophobicity as a stabilizing force in BDPP. Despite the paucity of hydrophobic amino acids in the BDPP sequence (< 5%), important regions of the protein on a limited scale may actually employ hydrophobic-hydrophobic sidechain interactions to stabilize a neighboring polyelectrolyte sequence into a local tertiary folding motif. Sabsay *et al.* (1991), in their preliminary sequence mapping of rat α-phosphophoryn, reported a trypsin digest generated 17 residue polyelectrolyte sequence:

-DDDDDDYSDSDSSDSDD.....

in which the 8th residue is Tyr. Phe, Tyr, and Trp sidechain pair formation, involving aromatic-aromatic interactions, have been observed in a number of proteins; it is estimated that 60% of the aromatic sidechains in proteins are involved in such pairs [Serrano *et al.*, 1991; Burley and Petsko, 1985]. Conceivably, a similar Phe, Tyr, or other hydrophobic amino acid sidechain in a separate polyelectrolyte sequence could form a hydrophobic-hydrophobic pairing that would bring the two polyelectrolyte regions closer together, thus fulfilling Mann's requirement for creating the highest possible degree of negative charge density on the surface of the template protein.

Our observation that the polyelectrolyte regions in divalent-cation depleted BDPP have some degree of organized tertiary conformation(s) at neutral pH suggests that the BDPP protein may have evolved a structural hierarchy for creating maximal negative charge density on its surface. These polyelectrolyte regions might be expected to undergo conformational change and reorganization in the presence of divalent cations to enhance the formation of nucleation centers on the protein surface. The creation of these high charge density regions would then result in the ability to sequentially bind divalent cations at specific regions on the protein, depending on the divalent cation:protein stoichiometry.

Acknowledgements

We thank Drs. Steve Weiner, Arthur Veis, and Heinz Lowenstam for their helpful discussions.

References

Addadi, L., Moradian, J., Shay, E., Maroudas, N. G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* 84 2732-2736.

Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* 82 4110-4114.

Ben-Naim, A. (1991) *J. Phys. Chem.* 95 (3) 1437-1444.

Burley, S.K., and Petsko, G.A. (1985) *Science* 229 23-28.

Campbell, I.D., Dobson, C.M., and Williams, R.J.P. (1975) *Proc. R. Soc. London B.* 189 485-502.

Chou, K-C., Carlacci, L., and Maggiora, G. G (1990) *J. Mol. Biol.* 213 315-326.

Chou, P. Y., and Fasman, G. D. (1974a) *Biochemistry* 13 (2) 211-221.

Chou, P. Y., and Fasman, G. D. (1974b) *Biochemistry* 13 (2) 222-245.

Cookson, D. J., Levine, B. A., Williams, R.J.P., Jontell, M., Linde, A., and de Bernard, B. (1980) *Eur. J. Biochem.* 110 273-278.

Creighton, T. E. (1990) *Biochem. J.* 270 1-16.

Dill, K.A. (1990) *Biochemistry* 29 (31) 7133-7155.

Evans, J. S., and Chan, S.I. (1991) in *Proteins: Structure, Dynamics, and Design* (Renugopalakrishnan, V., Carey, P.R., Smith, I.C.P., Huang, S.G., and Storer, A.C. Eds) ESCOM Press, Leiden, Netherlands, pp. 251-260.

Evans, J. S., and Chan, S.I. (1992a) *Biochemistry*, submitted

Evans, J. S., and Chan, S.I. (1992b) *Biochemistry*, submitted

Fersht, A.R. (1972) *J. Mol. Biol.* 64 497-512.

Friere, E., and Murphy, K.P. (1991) *J. Mol. Biol.* 222 687-699.

Freire, E., Murphy, K.P., Sanchez-Ruiz, J.M., Galisteo, M.L., and Privalov, P.L. (1992) *Biochemistry* 31 250-256.

Goto, Y., and Aimoto, S. (1991) *J. Mol. Biol.* 218 387-396.

Goto, Y., and Fink, A.L. (1989) *Biochemistry* 28 945-952.

Goto, Y., Calciano, L.J., and Fink, A.L. (1990a) *Proc. Natl. Acad. Sci USA* 87 573-577.

Goto, Y., Takahashi, N., and Fink, A.L. (1990b) *Biochemistry* 29 3480-3488.

Goto, Y., and Fink, A.L. (1990) *J. Mol. Biol.* 214 803-805.

Goto, Y., and Nishikiori, S. (1991) *J. Mol. Biol.* 222 679-686.

Hughson, F.M., Barrick, D., and Baldwin, R.L. (1991) *Biochemistry* 30 4113-4118.

Jaenicke, R. (1991) *Biochemistry* 30 (13) 3147-3160.

Lee, S. L., Glonek, T., and Glimcher, M. J. (1983) *Calcif. Tiss. Int.* 35 815-818.

Lee, S. L., Veis, A., and Glonek, T. (1977) *Biochemistry* 16 (13) 2971-2979.

Lowenstam, H. A., and Weiner, S. (1985) *Science* 227 51-53.

Lowenstam, H. A., and Weiner, S. (1989) *On Biomineralization* (Oxford University Press, New York) pp. 12-45.

Mann, S. (1988a) *Nature* 332 119-124.

Mann, S. (1988b) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Press, New York, New York, pp. 35-60.

Manning, G.S. (1969) *J. Chem. Phys.* 51 924-933.

Manning, G.S. (1978) *Quart. Rev. Biophys* 11 (2) 179-246.

Mayo, S.L., Olafson, B.D., and Goddard, W.A. (1990) *J. Phys. Chem.* 94

8897-8909.

Milner-White, E.J. (1990) *J. Mol. Biol.* 216 385-397.

Murphy, K.P., and Gill, S.J. (1991) *J. Mol. Biol.* 222 699-710.

Perutz, M.F., and Raidt, H. (1975) *Nature* 255 256-259.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* 42 104-112.

Richardson, J.S., Thomas, K.A., Rubin, B.H., and Richardson, D.C. (1975) *Proc. Natl. Acad. Sci. USA* 72 1349-1353.

Richardson, J. S. (1977) *Nature* 268 495-500.

Richardson, J. S. (1981) *Adv. Prot. Res.* 34 167-339.

Sabsay, B., Stetler-Stevenson, W. G., Lechner, J. H., and Veis, A. (1991) *Biochem. J.* 276 699-707

Saroff, H.A. (1991) *Biopolymers* 31 1037-1047.

Serrano, L., Bycroft, M., and Fersht, A.R. (1991) *J. Mol. Biol.* 218 465-475.

Sibanda, B. L., Blundell, T. L., and Thornton, J. M. (1989) *J. Mol. Biol.* 206 759-777.

Stetler-Stevenson, W. G., and Veis, A. (1983) *Biochemistry* 22 4326-4335.

Stigter, D., Alonso, D.O.V., and Dill, K.A. (1990) *submitted for publication.*

Stigter, D., and Dill, K.A. (1990) *Biochemistry* 29 1262-1266.

Thomas, D. J. (1990) *J. Mol. Biol.* 216 459-465.

Tramontano, A., Chothia, C., and Lesk, A. M. (1989) *Proteins* 6 382-394.

Weiner, S. (1983) *Biochemistry* 22 4139-4145.

Weiner, S., and Hood, L. (1975) *Science* 190 987-989.

Weiner, S., and Traub, W. (1984) *Phil. Trans. R. Soc. London* B 304 425-434.

Wilmot, C. M., and Thornton, J. M. (1988) *J. Mol. Biol.* 203 221-232.

Wilmot, C. M., and Thornton, J. M. (1990) *Prot. Engineering* 3 (6) 479-493.

Zanetti, M., de Bernard, B., Jontell, M., and Linde, A. (1981) *Eur. J Biochem.* 113 541-545.

Table I: *Temperature Dependence of ^1H NMR Chemical Shifts for BDPP Asp and PSer Residues.*

The °C/ppm values (mean) are determined as the slope of the linear least squares fit to ^1H NMR chemical shift vs. temperature data obtained for BDPP sidechain resonances at low ionic strength. Separate slope determinations (n=3) are utilized for each line to arrive at the mean value.

Table 1

Temperature Effects on BDPP $^{1}H/^{31}P$ NMR Chemical Shifts

Peak	ΔHz/°C x 10
F	+0.806
E	+1.207
D	-1.624
C	-1.637
B	-1.105
A	-1.315
12	+1.906
13	N.D.
14	-0.623
15	+0.981
16	N.D.

Figure 1: *Temperature Dependence of ^{31}P NMR Chemical Shifts for BDPP PSer Residues.*

Second-order polynomial ($y = Ax^2 - Bx - C$) curve fitting of BDPP ^{31}P chemical shift data points as a function of temperature.

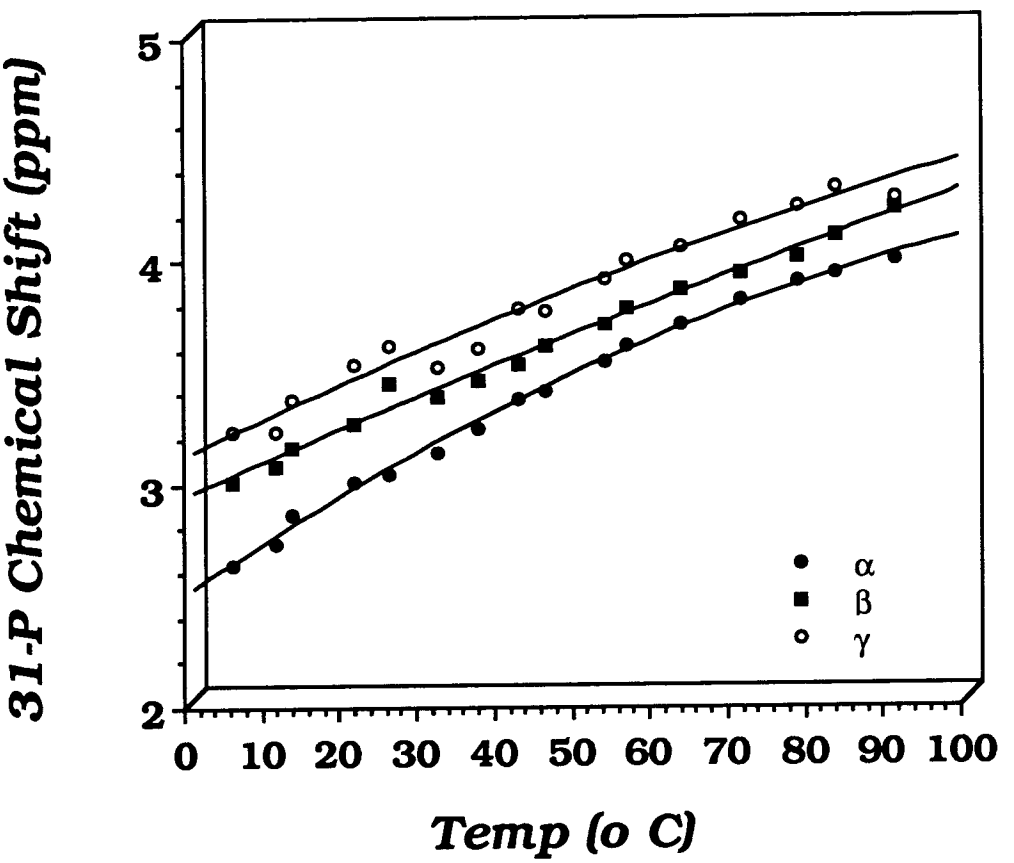

31-P Chemical Shift (ppm)
5
4
3
2
0
10
20
30
40
50
60
70
80
90
100
Temp (o C)
α
β
γ

Figure 2: *Temperature Dependence of ^1H NMR Spectra (Aliphatic Region) of BDPP at pH 7.0.*

NMR acquisition and processing parameters are given in the Materials and Methods section. Spectra correspond to BDPP sample (9.2 mg/ml) in 99.9% D_2O at the given pH. Temperature readings are given for each spectra. All peak areas are scaled relative to the BDPP sample at room temperature. The top spectrum represents the slow-cooling "renatured" BDPP sample after heating at 90° C. ^1H chemical shifts are referenced from 0.1 M TSS in D_2O at each temperature reading.

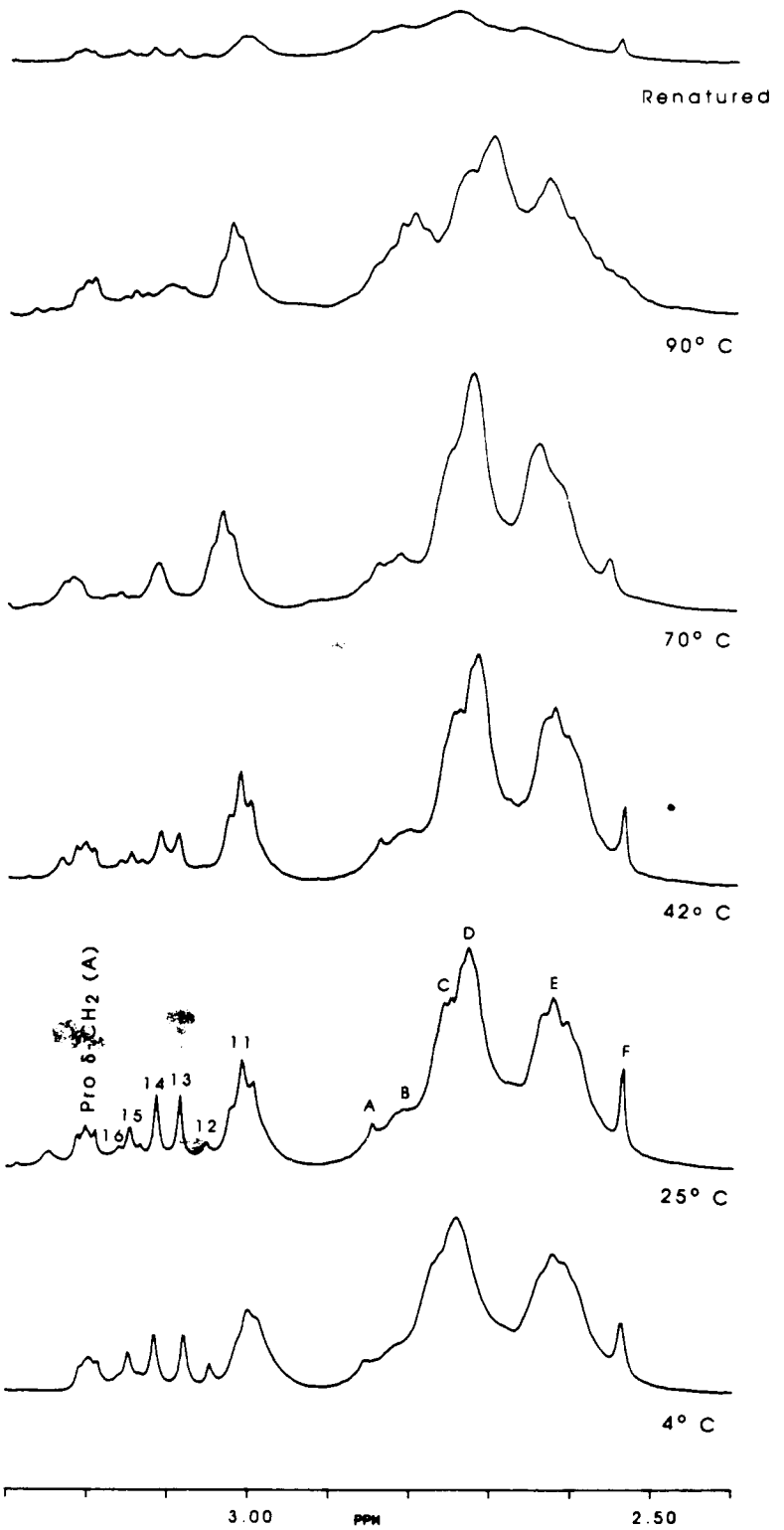

Renatured
90° C
70° C
42° C
Pro δ-CH2 (A)
11
14
13
15
16
12
A
B
C
D
E
F
25° C
4° C
3.00
PPM
2.50

Figure 3: *Temperature Dependence of ^{31}P NMR Spectra of BDPP at pH 7.0.*

NMR acquisition and processing parameters are given in the Materials and Methods section. Spectra correspond to BDPP sample (9.2 mg/ml) in 99.9% D_2O at the given pH. Temperature readings are given for each spectra. All peak areas are scaled relative to the BDPP sample at room temperature. The top spectrum represents the slow-cooling "renatured" BDPP sample after heating at 90° C. ^{31}P chemical shifts are referenced from 0.5 M EHDP in D_2O at each temperature reading.

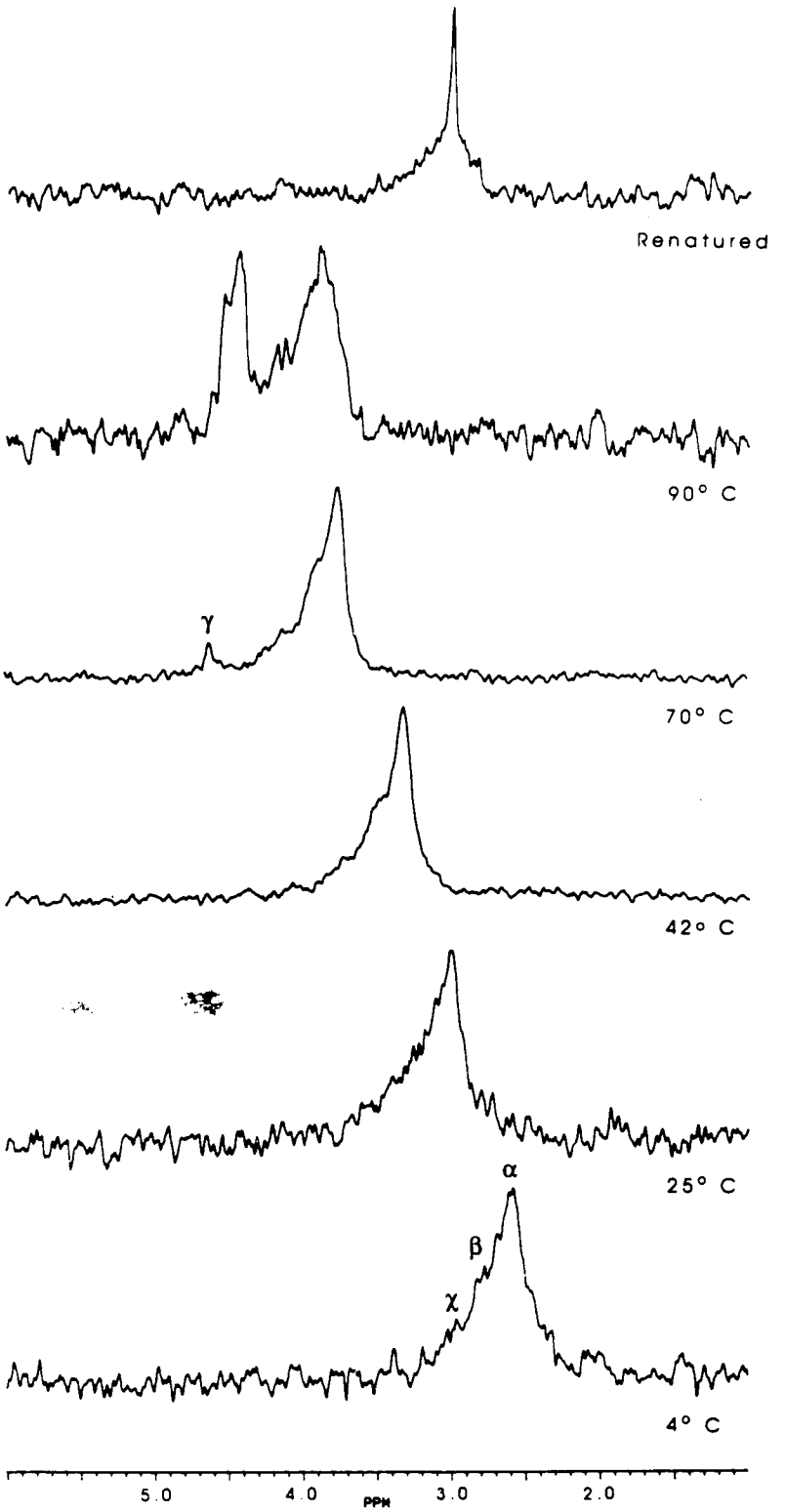
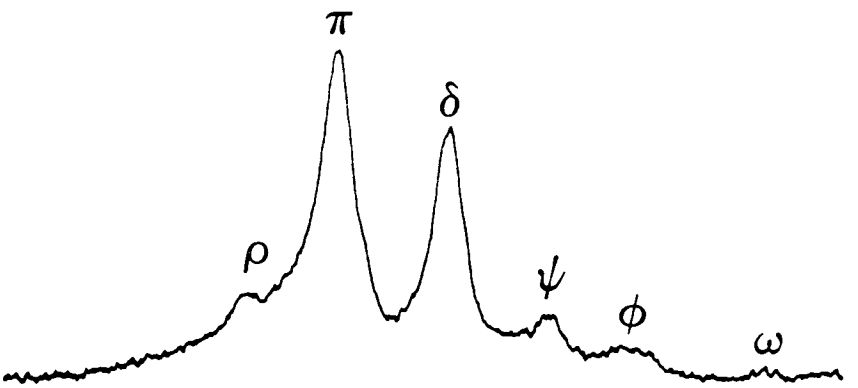
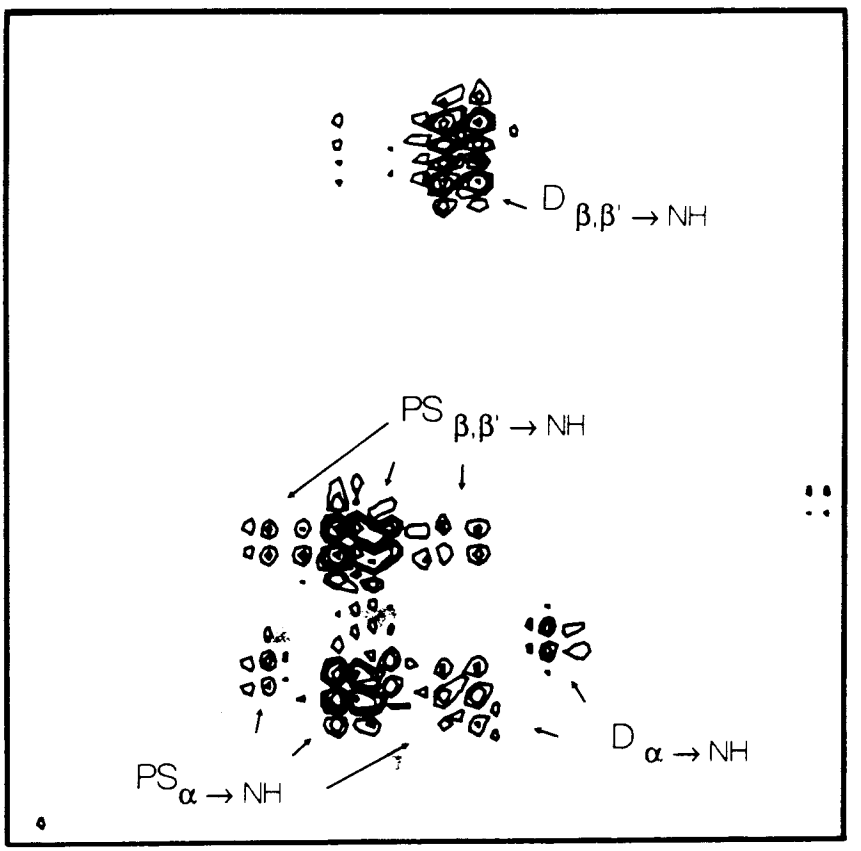
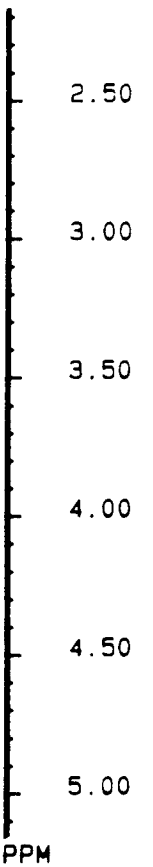

Renatured
90° C
γ
70° C
42° C
25° C
α
β
χ
4° C
5.0
4.0
PPM
3.0
2.0

Figure 4: *2-D HOHAHA Spin-Lock Correlation Spectrum (TPPI Phase Sensitive Contour Plot with F1 and F2 projection) of BDPP at pH 7.1, in 90% H_2O/10% D_2O.*

Details of the spin-lock correlation experiment are described in Materials and Methods. BDPP concentration was 23.2 mg/ml. Zero-filling in the F2 dimension was performed to create a 1 K x 1 K data matrix (4.883 Hz/pt in both dimensions) using a phase-sensitive (positive and negative levels) squared sine-bell window function in both dimensions with no shift. A total of 256 experiments were utilized. ^1H chemical shifts are reported from 0.1 M TSS in D_2O. Resonance assignments are given in Table 2.


π
δ
ρ
ψ
φ
ω


D β,β' → NH
PS β,β' → NH
D α → NH
PS α → NH


2.50
3.00
3.50
4.00
4.50
5.00
PPM

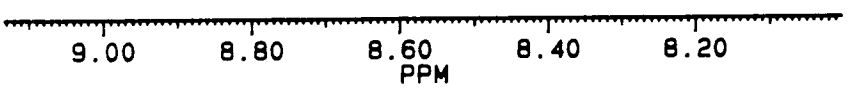

9.00
8.80
8.60
8.40
8.20
PPM

Table II: *Assignment of BDPP Amide NH Resonances.*

[0]All ^1H Chemical shifts are reported from 0.1 M TSS in D_2O. [&] This peak also possesses overlapping contributions from PSer NH resonances in the protein. Percentage of overlap not determined.

Table 2

BDPP Amide NH Assignments (pH 7.1)

Resonance	δ (ppm)[0]	Assignment
P	8.8038	PSer-NH
Π	8.6758	PSer-NH
Δ	8.5273	Asp-NH&
Ψ	8.3906	Asp-NH
Φ	8.2759	Asp-NH
Ω	8.1024	Asp-NH

[0] All ^1H Chemical shifts are reported from 0.1 M TSS in D_2O.

& This peak also possesses overlapping contributions from PSer NH resonances in the protein. Percentage of overlap not determined.

Figure 5: *pH Titration of BDPP PSer and Asp Amide Protons in H_2O.*

The pH titrations are conducted in H_2O/D_2O in the absence of buffer, using a BDPP concentration of 23.2 mg/ml. No correction has been made for deuterium isotope effects. ^1H Chemical shifts (ppm) are referenced from TSS (0.0 ppm). Symbol legend is given in the Figure.

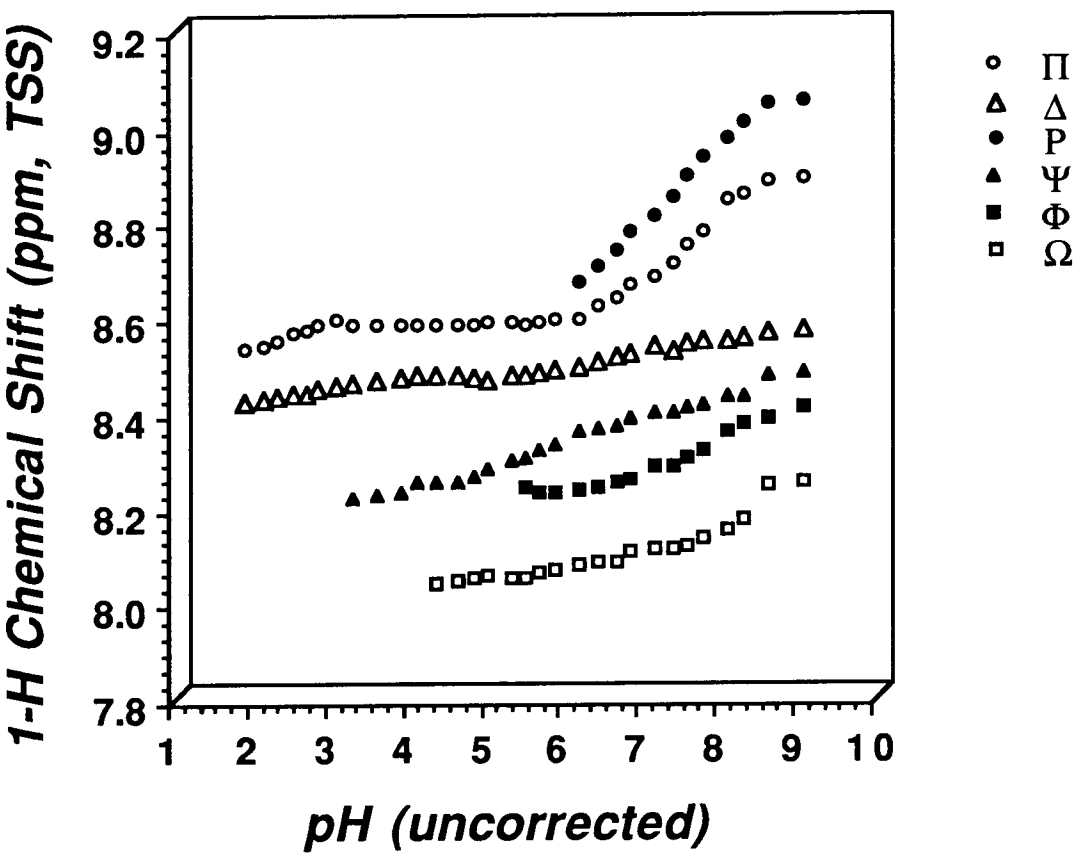

1-H Chemical Shift (ppm, TSS)
9.2
9.0
8.8
8.6
8.4
8.2
8.0
7.8
1
2
3
4
5
6
7
8
9
10
pH (uncorrected)
Π
Δ
P
Ψ
Φ
Ω

Figure 6: *AcetonitrileSolvent Effects on the ^1H NMR Spectra (Aliphatic Region) of BDPP, pH 7.0.*

NMR acquisition and processing parameters are given in the Materials and Methods section. Spectra correspond to BDPP sample (7.8 mg/ml) in 99.9% D_2O at the given pH, with the specified volume percentage of acetonitrile-*d*. All peak areas are scaled relative to the BDPP sample at 0% v/v acetonitrile. The top spectrum represents the acetonitrile solvent treated BDPP sample, recovered from the solvent via freezing and lyophilization, and subsequently reconstituted in 99.9 % D_2O, and adjusted to pH 7.0 ^1H chemical shifts are referenced from 0.1 M TSS in D_2O

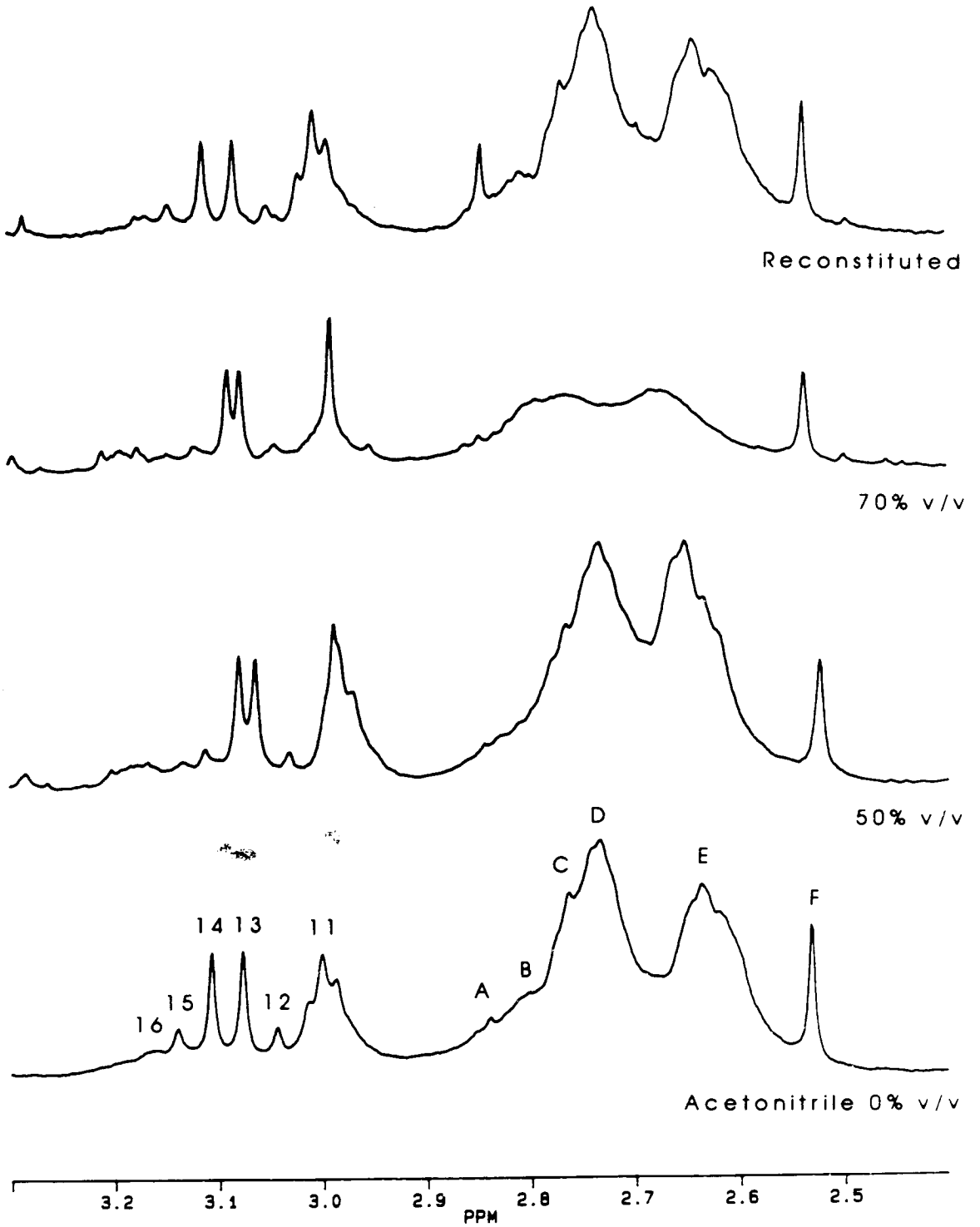

Reconstituted
70% v/v
50% v/v
Acetonitrile 0% v/v
D
C
E
F
14 13
11
B
A
15
12
16
3.2
3.1
3.0
2.9
PPM
2.8
2.7
2.6
2.5

Figure 7: *Solvent Effects on the ^{31}P NMR Spectra of BDPP, pH 7.0.*

NMR acquisition and processing parameters are given in the Materials and Methods section. Spectra correspond to BDPP sample (7.8 mg/ml) in 99.9% D_2O at the given pH, with the specified volume percentage of acetonitrile-*d*. All peak areas are scaled relative to the BDPP sample at 0% v/v acetonitrile. The top spectrum represents the acetonitrile solvent treated BDPP sample, recovered from the solvent via freezing and lyophilization, and subsequently reconstituted in 99.9 % D_2O, and adjusted to pH 7.0. ^{31}P chemical shifts are referenced from 0.5 M EHDP in D_2O

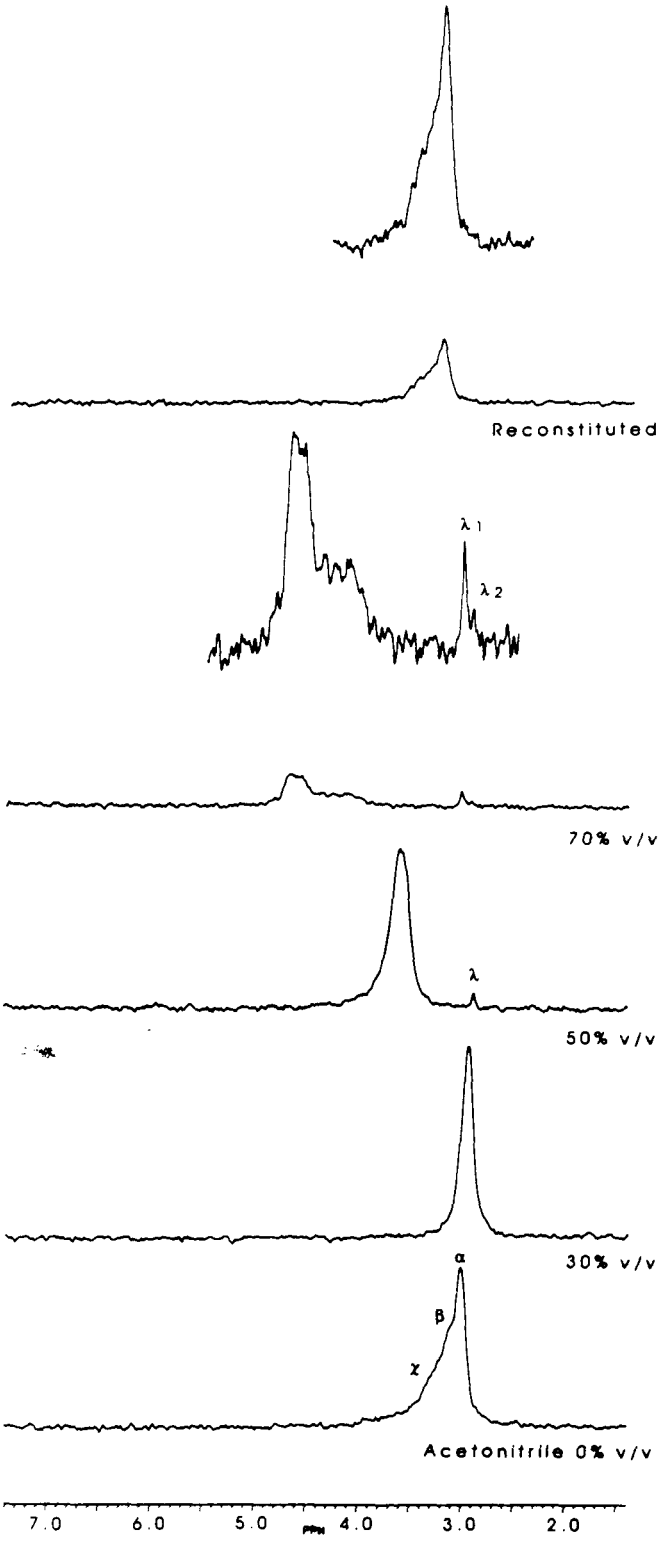
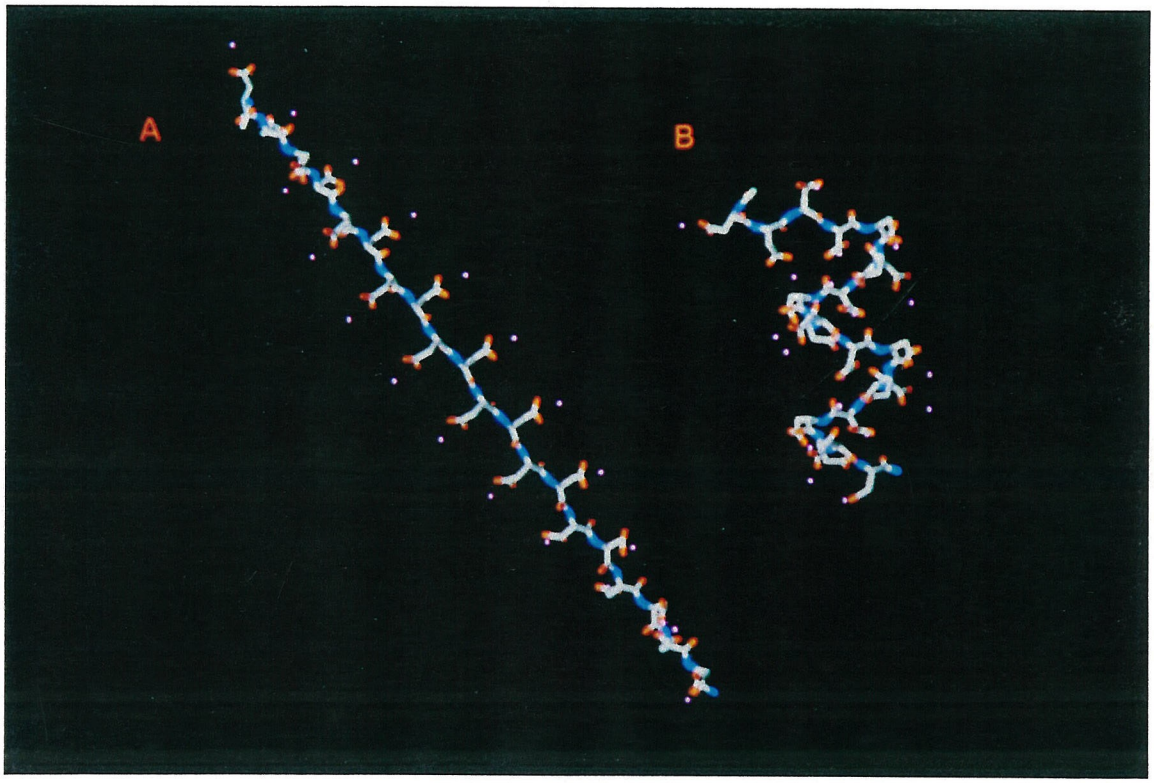

Reconstituted
λ1
λ2
70% v/v
λ
50% v/v
30% v/v
α
β
χ
Acetonitrile 0% v/v
7.0
6.0
5.0
4.0
3.0
2.0

Figure 8: *Examples of Spatial Charge Distribution in PMMP Proteins that Result in High Negative Charge Density.*

A ball and stick diagram of a hypothetical polyelectrolyte sequence region (N-acetyl-L-$(Asp)_{20}$-C-amide, Na^+ salt) is presented in two types of folding arrangments. In (A), the extended or random-coil arrangment is portrayed. In (B), a hypothetical fold in the peptide structure, in which spatial surface charge distribution is maximized, and charge-charge repulsion between sidechains is minimized, is presented. All figures were produced with the molecular graphics program BIOGRAF (BIODESIGN, Pasadena, CA) and the force field, DREIDING II (Mayo *et al.*, 1990). Atom legend: Gray: carbon; Red: oxygen; Blue: nitrogen Pink: sodium.


A
B

Seven

Monte-Carlo Conformational Search Algorithm. A Computational Method for Polyelectrolyte Peptide and Protein Folding Prediction.

ABSTRACT

Using the Dihedral Probability Grid-Monte Carlo conformational search algorithm/Metropolis sampling method devised by Mathiowetz and Goddard (preceeding manuscript) and the DREIDING forcefield, we investigated the lowest energy conformers for a series of N-acetyl, C-terminal amide capped polyelectrolyte peptides $(Asp)_{20}$, $(PSer)_{20}$, and $(PSerAsp)_{10}$. These sequences represent cation binding regions in phosphophoryn, a protein template macromolecule responsible for the nucleation of calcium phosphates, and a potential model for inorganic material synthesis. The starting structures possessed the following features: 1) extended secondary structure, fully deprotonated state; 2) Na counterion (1Na : 1 COO, 2 Na : 1 PO_4) placed in a non-bonding bridging geometry between oxoanions of each Asp and PSer sidechain; 3) charge equilibration for all atoms of the Na complexed peptides, with the net molecular charge set = 0.0000 e.u.; 4) minimization to convergence of all structures prior to Monte Carlo ψ, ϕ backbone dihedral simulation. Using the branching algorithm, each Na atom is considered to be the last atom off of the sidechain branch for the purposes of dihedral motion. The accepted structures generated from these simulations were evaluated according to internal potential energies [for dielectric constant (ε_o) values of 1, 4, 10, 40, and 80] and to solvation potentials calculated via the protein dipole-langevin dipole method (PDLD)(Russell and Warshel, 1985). Our results indicate that by utilizing explicit counterions for ion pair charge balance, no major differences are found between structures generated with different ε_o values. In addition, our results also predict that certain sequences exhibit a preference for certain global conformational minima: $(Asp)_n$ form supercoil conformers; $(PSer)_n$ form hairpin structures; and $(PSerAsp)_n$ exhibit a preference for spiral conformers.

INTRODUCTION

The optimal spatial structure of proteins and peptides are achieved by a balance of intermolecular forces within the protein structure and with its environment. This optimum, which can be defined as the native conformation, represents a global potential energy mimina, or lowest free energy state of the system as defined by the "thermodynamic hypothesis" of Anfinsen[1,2]. In addition to the global potential energy minima, multiple conformational states may also exist for a particular protein[4-7]. For example, these states may represent intermediates in protein folding[8,9], or, conformational change resulting from molecular interaction with a ligand or substrate[10]. In principle, identifying the global conformational minima state for a protein can be achieved by a rigorous evaluation of the free energies for each possible conformer[11,12]. Various statistical mechanical analysis algorithms have been developed for global conformational minima search; some examples include (1) *The nearest-neighbor Ising model*[19], and its semi-empirical variations[20,21], (2) *Monte-Carlo methods*[14], and (3) *Two-dimensional and three-dimensional lattice fitting routines*[11,12]. With varying degrees of success, each method has made relatively accurate predictions regarding secondary and tertiary folding of peptides and small molecular weight proteins. In principle, if one possessed the primary amino acid sequence, theoretical methods could determine the lowest energy conformational state of a protein or peptide.

As it turns out, nothing is as simple as it seems. First, the conformational energies of a protein or peptide are not necessarily limited to a single minima, but in fact possess multiple minimas[13,14]. The problem then becomes one of avoidance of local minima trapping, and instead determining the "absolute" minima, which corresponds to the global minima. Secondly, a computational search that extends over all possible conformational space is time-consuming, particularly as the number of residues in the sequence grows large. Thirdly, the thermodynamic hypothesis of protein folding includes the interaction of the solvent, or environment, into its determination of the native state[1,2]. However, the inclusion of environmental factors, such as counterions[15-17] and

hydration[18] complicates the development of a statistical algorithm that can accurately account for all stabilizing and destabilizing forces that contribute to the native state conformation. It becomes apparent that if one wishes to rigorously determine the lowest potential energy minima structure for a protein or peptide, a considerable amount of computational effort will be required.

However, there are several cases where the need to determine the absolute lowest energy potential is not crucial. One such instance is in the study of polyelectrolyte polymers, where internal motion and electrostatic interactions leads to an ensemble of structures in solution that possess conformational isomerizations, extended shapes and intermediate bending and twisting rigidities[18]. Thus, although there may be only one true low energy minima structure for a polyelectrolyte polymer, there may also be a great number of other structures energetically accessible to the polymer, such that the polymer is overwhelmingly likely not to be in the lowest energy state at any given instant in solution[19-21]. In this case, the identification of multiple minimas would be of great interest, since the collection of generated structures would yield a series of conformational "snapshots" detailing the dynamics of polyelectrolytes in solution.

In our studies of polyelectrolyte mineral matrix proteins (abbreviated PMMP) which act as templates for inorganic solid phase synthesis in a process known as biomineralization, several observations have been made regarding repeating anionic sequences. These polyelectrolyte sequences, estimated at 10 to 100 or more contiguous residues[23,24] , are the sites at which divalent cations coordinate in the process of nucleation. Under conditions of low ionic strength (i.e., Na^+) in the absence of divalent cations, protein thermal and solvent denaturation NMR experiments demonstrated that bovine dentine phosphophoryn (BDPP), a PMMP protein, possesses secondary and tertiary folding in these polyelectrolyte sequence regions[25]. This protein acts as a nucleation and crystal growth regulator of calcium phosphate mineral in tooth dentine[24]. This folding may be of importance in creating localized regions of high negative charge density which would attract and bind metal ions. Because of the large number of amino acids present and sequence redundancy, the use of conventional NMR spectroscopy to solve

the structures of these sequences is severly restricted at the present time. In the meantime, computational methods for the prediction of polyelectrolyte conformational minimas might provide useful information regarding secondary structure probabilities and the minimization of sidechain electrostatic interactions, provided that (1) an effective method for conformational search is employed, and (2) a realistic manner of depicting counterion condensation and solvent electrostatic potential is developed.

Recently, we have devised a Monte-Carlo method that employs importance sampling[14, 26] of dihedral angles to model peptide and protein conformations. This new method, termed Dihedral Probability Grid Monte-Carlo (DPG-MC), employs either backbone (ψ, ϕ) or sidechain (χ) dihedral angle motion to generate new amino acid conformations and evaluates the generated structures on the basis of the Metropolis criterion[27]. These simulations proceed by choosing amino acids at random and, depending on the simulation, modifying their backbone or sidechain dihedrals. The new dihedral angles for the amino acid are chosen from an M-dimensional grid of size N^M, where $N = 360/S$ and M is the number of dihedrals involved. Backbone dihedrals, for example, are selected from N x N grids representing the possible ϕ, ψ pairs, where the probability of choosing one of the N^M grid points is determined by partitioning the conformations in a selected subset of protein crystal structures from the Brookhaven Protein Database.

In this first series of reports, we discuss the usage of DPG-MC method in the predictive modeling of folding states as a function of charge screening (Na^+ counterion) in defined sequence length homopolymers and heteropolymers of Asp and o-phospho-L-serine (PSer)[i.e., $(Asp)_n$, $(PSer\text{-}Asp)_m$, and $(PSer)_n$, where n = 20, m = 10], which represent Ca (II)-binding motifs derived from bovine dentine phosphophoryn[28]. We also investigate the effectiveness of the distance-dependent dielectric constant in representing force-field electrostatic potentials and influencing the outcome of the conformational search, and, the application of the protein dipole-langevin dipole method (PDLP) of Warshel and colleagues[29-31] in evaluating lowest energy potential structures on the basis of solvent electrostatic potentials. As described in this report, the application of DPG-MC permits effective

searching for polyelectrolyte peptide conformational minima that provides an insight into the nature of tertiary folding and backbone conformation in PMMP proteins.

METHODS

Construction of the protein database subset. The dihedral probabilities which are integral to our method require a judicious choice of structural data. Therefore, we sought to use a subset of protein structures from the Brookhaven protein database (PDB) which was both diverse and accurate. The PDB dataset contains approximately 515 protein crystal structures, excluding structures with only C_α coordinates. However, there are many proteins which are represented numerous times or are highly homologous to other proteins in the PDB dataset. Such identical, or nearly identical, structures would tend to distort our probabilities in favor of geometries found in those particular proteins. In order to eliminate highly redundant structures, we carried out pairwise sequence comparisons among 515 proteins in our initial PDB dataset. Any protein with greater than 75% sequence identity with another protein of higher resolution was eliminated. Although 75% sequence identity represents a high degree of homology, there can still be a substantial amount of conformational information unique to each structure. This homology-elimination process reduced our dataset from 515 proteins to 201. This dataset of 201 proteins, which we call UNIQUE-201, is useful for a wide variety of statistical analyses. However, geometric analyses such as those required here require high resolution data, so we further reduced the dataset to 81 proteins which had crystals of 2.0 Å resolution or better and R-factors of better than 20%. We call this dataset HIGHRES-81.

Phi/Psi Monte Carlo Method. For conformational sampling of the backbone dihedral angles ϕ and ψ, we create a markov chain of conformations, where each structure differs from the previous by changing the ϕ and ψ of one amino acid. The amino acid to be modified is chosen completely at random, with each residue having the same likelihood of being chosen. The new ϕ,ψ pair chosen for the residue is chosen from a grid of probabilities where the spacing between the grid points = S°. The grid, therefore, contains $N \times N$

points, where N = 360°/S. We allow for the possibility of *cis-* peptide bonds (ω = 0°) for proline residues, but the ω dihedral is treated independently and not as an additional dimension in the probability grid. The probabilities were determined by partitioning every ϕ,ψ pair in the proteins comprising HIGHRES-81 dataset into bins of size S° x S° and normalizing. We have determined separate probability grids for each amino acid, as well as different grids for *cis-* and *trans-* prolines , but it is usually sufficient to use only three different probability grids: one for prolines, one for glycines, and one for the other 18 residues. We have also used the secondary structure designators in the protein database (HELIX, SHEET, and TURN) to obtain separate probability grids for alpha helix, beta sheet, and coil regions. We considered coil regions to be all regions not specified by one of the three structural designators. We have used grid spacings of various sizes, from 5° to 60°. The narrower spacing allows for much greater conformational variation, but the resulting increase in the total conformational space necessitates a great increase in the number of steps needed for adequate sampling. This is somewhat mitigated by the fact that for the finer grids, many grid points have zero populations in the HIGHRES-81 dataset. For instance, even the most general probability matrix, for all types of residues regardless of secondary structure, 126 of the 144 possible 60° grid points (87.5%) are populated, while only 1597 out of 5184 grid points (30.8%) are populated on a 5° grid.

Sampling Criteria. The process of generating and assessing conformations is the same for both backbone and sidechain Monte Carlo. The protein is assigned an initial conformation, and the potential energy of the structure (E_i) is calculated. The simulation proceeds by modifying this conformation at one amino acid, either by changing its backbone or sidechain conformation as explained above. This potential energy of the new conformation (E_{i+1}) is then evaluated and this structure is either accepted or rejected according to the Metropolis criterion[27]. If the new structure is lower in energy ($E_{i+1} < E_i$), it is accepted. If it is higher in energy, the probability of acceptance, P, is defined by:

$$P = e^{(E_{i+1}-E_i)/kT} \qquad (1)$$

This probability is enforced by comparing the exponential quantity to a random number between 0 and 1. There is, therefore, a 50% chance of a structure being accepted if $E_{i+1} - E_i = -kT \ln(0.5) = 0.413$ kcal/mol at T = 300 K. If the structure is accepted, it becomes the new structure i. Otherwise, structure **i** remains the same. The process of generating new structures and accepting or rejecting them continues for a pre-defined number of steps or until endpoint criteria are met.

All calculations presented here were performed using a modified version of BIOGRAF running on Silicon Graphics workstations (models 4D/380 and 4D/25). Most of our simulations have used the DREIDING forcefield[32], but we have also used the AMBER forcefields[33] for comparison purposes. Our implementation includes the use of the numbering scheme of Abe and co-workers[34] which allows us to calculate torques and update atomic coordinates rapidly. Although torques are not used by the DPG method itself, they are necessary for a conjugate-gradients internal-coordinate energy minimizer which we use to optimize the structures produced by the Monte Carlo simulations.

Conformational Search of PMMP Ca (II) Binding Domain Peptides $(Asp)_n$, $(PSer)_n$, and $(Pser\text{-}Asp)_m$. The polyelectrolyte sequence peptides L-$(Asp)_n$ and L-$(PSer)_n$, where n=20 residues, and L-$(PSerAsp)_m$, where m=10 residues, were constructed in the peptide build function in the extended backbone conformation (ψ, ϕ=180°), such that consecutive residues of the peptide were *anti* conformers relative to one another. The peptides were modified by the base fragment addition of H_3PO_4 to Ser residues and subsequent removal of superfluous O and H atoms to create the deprotonated *o*-monophosphate ester, and, the deprotonated carboxylate form of Asp. Likewise, the peptides were "capped" by the addition of an acetyl group to the N-terminal α-NH_2 group and the conversion of the C-terminal α-carboxylate group to the amide form. Explicit hydrogen atoms were added to the sidechain β-carbon atoms and to C_31 and N_R atoms of the peptide backbone, acetyl, and C-terminal amide. No hydrogen atom types were added to the carboxylate or monophosphate ester groups; this assumes that our system reflects a solution pH of 8.0,

where both chemical groups would exist in the deprotonated form[34], and we are only interested in Na counterion screening.

In our model of Na^+ counterion condensation of polyelectrolyte peptide sequences, we have adopted the following assumptions. The Na^+ in our system is considered as a condensed counterion species associated with the peptide and does not exchange out into the bulk solvent. However, *the Na atom is free to reposition itself relative to the other atom types on the peptide structure during each phase of mechanistic minimization of non-bonding forces that occurs <u>after</u> each Monte-Carlo dihedral motion, and, during the quenching routine for each "best" accepted structure.* Next, each Na atom is used to balance the net anionic charge of each sidechain group, with no Cl^- counteranion or solvent present. Thus, in the fully deprotonated form (e.g., pH 8.0 or above), the situation Na : COO and 2 Na : PO_4 exists, i.e., the system is minimally saturated with Na, and the net molecular charge of the Na:peptide complex should be neutral. Secondly, in DREIDING II the atom type for lone pair oxygen (O_2) is the same for all sidechains and for all peptide carbonyl oxygens[32]. With this in mind, we have elected to place the Na atom in an equidistant fixed bridging geometry relative to each pair of "anionic" O_2 atoms of each amino acid sidechain, with a single Na atom lying in the O-C-O bond plane of the carboxylate group, and, in each of the O-P-O bond planes as shown in Figure 1. The justification for using the Na - O bridging geometry is supported by structural evidence from crystallographic databases compiled for organic and inorganic carbonates, carboxylates, and phosphates[35], and, has been determined as an optimal geometry from SCF *ab initio* calculations on carboxylate and phosphate-metal complexes[36,37]. The net charge on the Na:peptide complexes were assigned by using the charge equilibration approach of Rappe and Goddard[38]. In the case of Na:peptide complexes, the total charge for each peptide sequence in the extended starting conformation is equilibrated to a value of 0.000 electron units (e.u.), with a convergence criteria of <0.0005 e.u. [38]. Mean equilibration Q values for the C_R, P, O_2 and Na atoms that comprise the various Na:peptide complexes are shown in Table I. Partial charges for the peptide backbone atoms C_R, C_31, O_2, and N_R are given in Table I. The Na:peptide complexes were then re-minimized in Cartesian coordinate space, using ten

steps of steepest descents minimization[39] to relieve improper contacts, followed by conjugate-gradient minimization[39] to convergence [rms atom force criteria < 0.100 kcal/mol/Å]. A non-bonding cutoff of 9.0 Å and a hydrogen bonding cutoff of 5 Å were utilized in all minimization protocols.

One particular feature of the DPG-MC program is the ability to define each dihedral ϕ, ψ pair sidechain as a "branching" structure off of the peptide backbone for the puposes of determing torsional motion[34]. This "branching" feature permits the assignment of a non-bonded species, such as a cation or water molecule, to a specific sidechain residue. In this instance, each sidechain Na+ atom was designated the last atom(s) of that particular branch structure off of the backbone. This sidechain-metal "pair" will then remain associated via a minimzed non-bonding distance throughout the dihedral motion and over the course of the conformational search. Although this approach is simplistic and qualitative in nature, it does provide a unique approach to maintaining non-bonding and electrostatic interctions with metal ions and sidechain ligand atoms, and allows these interactions to influence the outcome of the global conformational minima search.

DPG-MC Simulations of Polyelectrolyte Peptides. Using the minimized Na peptide extended structure as the initial conformation, dihedral phi/psi Monte-Carlo simulations were conducted at Boltzmann temperatures of 500, 1000, 1500, and 2000 to 10,000 K in 1000 K increments, yielding a total of nine runs for each peptide case. Backbone dihedrals were chosen exclusively, since it has been shown that to a large degree phi/psi dihedral angles determine tertiary folding in peptides. For 20-mer structures 2×10^5 steps were utilized in each temperature run. In some runs, other starting structure conformations, such as α-helix (ϕ, ψ = -57, -47), β-sheet (parallel: ϕ, ψ = -119, 113) and β-turn (Type IIA: ϕ, ψ = -50, 105) were also substituted for the extended state, to assess if the starting structure conformation had any affect on the outcome of the DPG-MC calculations. We utilized probability grids with increments between phi/psi pairs of 60, 30, and 5 degrees; 5 degree grids were found to be optimum in terms of generating lowest energy minimas. Backbone dihedrals were incremented one pair at a time for each step, with the pairs randomly selected at each step of the Monte-Carlo. One step of

conjugate-gradient minimization was performed at each Monte-Carlo step prior to Metropolis evaluation of total energy. "Accepted" structures represent the lowest energy minima generated during a particular run; from this set, the DPG-MC program determines the overall "best" conformer, which was then subjected to Cartesian minimization to convergence as described above.

Representing Electrostatic Forces and Simulations in the Presence of Solvent. In the DREIDING II force field, the non-bonding potential energy function (E_{NB}) is given by the expression[32]:

$$E_{NB} = E_Q + E_{vdW} + E_{HB}, \qquad (2)$$

where E_Q, E_{vdW}, and E_{HB} denote the electrostatic, van der Waals, and hydrogen bonding potential energies of the system of atoms. The energy term E_Q represents a through-space electrostatic potential for nonbonded atoms[32]:

$$E_Q = \frac{(322.0637)Q_iQ_j}{\varepsilon R_{ij}} \qquad (3)$$

In this equation, Q_i and Q_j are the charges in electron units for atoms **i** and **j**, R_{ij} the distance in angstroms, ε is the dielectric constant, and the constant 332.0637 converts the expression into kcal/mol units. The ε term is expressed as a distance-dependent dielectric constant[32]:

$$\varepsilon = \varepsilon_o R_{ij} \qquad (4)$$

This non-linear scaled distance-dependent Coulomb's law relationship should yield a good first approximation for short-range metal-sidechain and sidechain-sidechain electrostatic interactions (i.e., 9.0 Å or less)[40], while permitting long-range electrostatic interactions to occur as well. It should be noted that the use of the "straight" or linear distance-dependent dielectric constant (D = r_{ij}) can lead to an overestimation of electrostatic forces[17,18,20], especially for the case ε_o = 1 (*in vacuo*). As a means of analyzing the effect of ε_o on the internal and electrostatic potential energies of the DPG-MC generated conformers, and modeling solvent electrostatic screening and its effects on

defining the lowest potential energy conformers[15], we employed dielectric values of 1, 4, and 10 in the minimization and execution of DPG-MC runs for all Na^+ peptide complexes for each temperature run conducted. Historically, considerable use has been made of the dielectric continuum model ($\varepsilon_o > 1$, with the upper limit is taken to be 80 for water) for implicit modeling of bulk solvent effects on mechanistic and dynamic simulations[29-31,41-44]. We have excluded induced polarization effects [29,31,43,45] and the variation in the dielectric constant over the region of the molecule.

Since the variation in dielectric constant will lead to scaling of the electrostatic forces between charges and thus the electrostatic energy of the system, another evaluation criteria needs to be introduced in order to select from among the dielectric-variable accepted structures a "best" conformer. One approach is to evaluate the free energy of solvation ($\Delta\Delta G_p^{sol}$) for each accepted DPG-MC structure, according to the Protein Dipole Langevin Dipole model (PDLD) of Warshel and colleagues[29-31]. These calculations were carried out using the POLARIS force-field (version 3.0, BIODESIGN, Inc.). This model simulates solvation free energy in polar solvents via the Langevin function for average orientations of the solvent dipoles around charged atoms, and includes two additional solvation terms, viz:

$$\Delta\Delta G_p^{sol} = \Delta G_{Born} + \Delta G_{Langevin} + (-T\Delta S), \qquad (5)$$

where ΔG_{Born} is the residual bulk energy of the solvent, $\Delta G_{Langevin}$ is the enthalpy of the interaction of the water dipoles with the solute, and ($-T\Delta S$) is the entropy of solvation at 300 K. For each Na^+ - peptide complex the two-region model for energy calculations were conducted in single run mode, defining the Region by an outer radius of 30 Å, an inner radius of 12 Å, and an inner and outer grid spacing of 3 Å and 2 Å, respectively. The value of 30 Å for the outer radius was found to accomodate all of the DPG-MC generated structures in our study. Other solvation energy terms utilized in the POLARIS calculations are given in the appropriate Figure and Table legends.

Results

Charge Distribution on Na-Peptide Complexes. The point charges obtained by the application of the Rappe-Goddard equilibration routine to Na-counterion peptides in the extended secondary structure are shown in Table I. This approach to data analysis has been utilized elsewhere[40]. One goal of the equilibration routine is to obtain a consistent treatment of charges for counterions, and, for the sidechain atoms that ion pair with the metal atoms. These sidechain atom values will directly influence the folding of the peptide into a energy stablized conformation. Of additional interest are the peptide backbone charges, which can have important consequences for intrapeptide chain interactions (i.e., dipole moments)[40]. As shown in Table I, for each atom type throughout a particular peptide sequence, the equilibration program generates consistent peptide backbone charge values that are within statistical agreement. To a certain degree this holds true for different sequence arrangements of Asp and PSer, and, for different residue lengths of the same sequence. It has been shown in *ab initio* charge distribution calculations that the partial charge on C_α depend strongly on the residue type and to a certain degree, sidechain conformation[40]. Overall, the Rappe-Goddard charge equilibration was successful in assigning backbone partial atomic charges in a consistent manner for each peptide. A comparison of the partial atomic charges obtained for the sidechain and Na atoms [Table I] reveals that the equilibration routine assigns consistent values (within statistical error) for carboxylate and monophosphate ester atoms in each peptide, regardless of the sequence or the residue length. However, the assignment of charges for the Na atoms showed interesting variations. For Na atoms which were assigned to ion pair with carboxylate groups, the Na partial atomic charge values were found to be very similar regardless of the residue length or sequence. In comparison, Na atom pairs which were assigned to the monophosphate ester, while statistically similar, exhibited a larger degree of variation in charge assignment values within a particular peptide-Na cluster [e.g., in $PSerAsp_{10}$, the S.D. for phosphate-associated Na atomic charge values was ±0.08 and ±0.39]. Since the geometries of the sidechains and Na atoms are fixed identically in the starting structures, and the distance between adjacent sidechains is greater than 9 Å, the variation in Na charge

assignments is not the result of nearest-neighbor sidechain interactions. Rather, these variations may have occurred as a result of the charge delocalization within each PSer residue, i.e., influence from nearby negatively charged O_2 atom types of the monophosphate ester on the bridged Na atom.

Analysis of DPG-MC Generated Polyelectrolyte Peptide Potential Energies. For each peptide case, the percentage of accepted conformations is very small, indicating that the Metropolis sampling criteria was successful in eliminating unwanted structures from consideration [Table II]. An inspection of the DREIDING II-determined low energy minima generated by the DPG-MC method for each peptide sequence [Table II] reveals that the ensemble of "best", or accepted, conformations are significantly lower in overall energy as compared to the initial structures (as given by ΔE, or the difference in total energies between accepted and starting conformer). This is true for every dielectric constant utilized. Although the data in Table II is reported for accepted conformers generated from starting structures set in the extended state, we found no significant differences in results when other secondary structures were employed as a starting conformation (data not shown).

The effect of increasing ε_o on the electrostatic potential energies and total energies of the best conformers [Table II] leads to a reduction in the potential energies in both cases for all peptide systems. For each 20 mer sequence, increasing the dielectric from 1 to 4 resulted in an average 5-fold reduction in the electrostatic potential; from 1 to 10, this average reduction was nearly 10 to 20-fold. An evaluation of the energy decomposition for each dielectric case also reveals that reductions in other energy terms have occurred, most notably the van der Waals potential [Table II] and total bonding energies. The van der Waals energies for ε_o = 4 and 10 represent a 100-300% decrease over values obtained in vacuo. Correspondingly, the bonding energies are also observed to decrease with increasing ε_o: 10% for $PSerAsp_{10}$ and $PSer_{20}$, and 30% for Asp_{20} [Table II]. Thus, the overestimation of electrostatic forces using ε_o =1 leads to an overestimation of other internal energy parameters as well. By scaling the magnitude of electrostatic forces via the dielectric constant, accepted structures are selected for based on a

balance of electrostatic and van der Waals interactions, and the bonding geometries, in each peptide complex.

PDLD Solvation Potential as a Criteria for Selecting Best Conformers. As one can see, "accepted" structures generated by the DPG-MC program at various dielectric values all feature significant lowering of potential energies as compared to the starting structures. Since the variation in ε_o prevents a direct comparison amongst conformers of a particular sequence, an important issue thus arises: if one chooses to represent the electrostatics of polyelectrolyte sequences by counterion pairing, under what force-field conditions are "best" accepted structures generated that are representative of polyelectrolyte structures in solution? To investigate which e_o value leads to the generation of the lowest energy conformation, (1) we calculated single energies for the case ε_o = 1 for all accepted structures, and (2) peformed PDLD solvation potential calculations[29-31]. The results are shown in Table III. It is evident that the lowest mean energies in vacuo were obtained for the case ε_o =1. The corresponding mean $\Delta\Delta G_p^{sol}$ values, however, were found to be least favorable by at least 80-100 kcal. The PDLD model has been reported to exhibit sensitivity to small changes in protein coordinates (e.g., 0.2 Å rms deviation in x-ray structure coordinates results in a 2kcal/mol change in $\Delta\Delta G_p^{sol}$ values)[30]. In our study, the major variant in the $\Delta\Delta G_p^{sol}$ term was found to be the $\Delta G_{Langevin}$ energies; differences in the ΔG_{Born} and $(-T\Delta S)$ energies were found to be negligible for each dielectric case and for each sequence peptide studied. This is hardly surprising, since all residues are hydrophilic, and the only discernable differences between the peptide conformers are the interactions between permanent and induced dipoles of the peptide and the Langevin dipoles. In general, conformers generated with a dielectric of 10 possessed the lowest solvation potentials. If one adds the two potentials together for each accepted structure, an interesting result emerges: the lowest overall energies are obtained for the case ε_o = 1, and there is approximately a 10-15% difference in overall energies for each dielectric case studied (Table III). Therefore, using two criteria to evaluate the potential energies of the structures, the "best" conformers are apparently those generated for the case ε_o = 1. Thus, positioning Na or other counterions to balance partial atomic charges is not only a useful way of modeling

counterion condensation, but may also fulfill the role of representing solvent electrostatics as well in polyelectrolytes.

Analysis of DPG-MC Accepted Structures. Several criteria can be employed to evaluate the structures generated by the DPG-MC program in the DREIDING II force field. Since the initial structures started at the extended conformation, where Φ, Ψ = 180°, the initial peptide geometry is linear in three-space. As a consequence of dihedral motion, lower potential energies are achieved by some degree of deviation from the linear conformation. This is evident from the dihedral phi/psi distribution plots obtained for accepted structures (Figure 2). For $PSer_{20}$ and Asp_{20}, dihedral pairs are cleanly distributed between β-sheet, helical and β-turn values (Type I, IIA), regardless of the dielectric utilized in the DPG-MC simulation. An examination of the "best" conformers (Figure 3) indicates that these dihedral values translate either as (1) "supercoil" structures (Asp_{20}), where the conformer possesses numerous turns, or (2) "hairpin" or "distorted hairpin" structures ($PSer_{20}$), in which a single major turn is noted. For $(PSerAsp)_{10}$, a slightly different situation exists. For the conformers generated under ε_o = 1 conditions, the dihedral library for "accepted" conformers exists exclusively in the extended and β-sheet regions (Figure 2). Such a structure resembles a "spiral" conformer (Figure 3). For ε_o = 4 and 10, the dihedral pairs are largely distributed in the β-sheet and helix/turn regions. However, an examination of the "best" conformers (Figure 3) reveal that these bear a striking similarity to the "spiral" structures, with larger turn radii. These "spirals" also resemble "hairpins". Therefore, a variation in the dielectric continuum does not dramatically alter the outcome of the DPG-MC search.

How can these structural differences be explained? The key "reason" behind the DPG-MC algorithm's choice of certain folded structures can be explained by polyelectrolyte theory and optimizing electrostatic interactions. If we compare the three sequences, the $PSerAsp_{10}$ polymer consists of polyion groups (PSer) separated from one another by monovalent ion groups (Asp). Therefore, unlike $PSer_{20}$ and Asp_{20}, the axial charge density of the polymer (ξ) varies over its length, viz[15,16]:

$$\xi = Q^2/\varepsilon kTb, \qquad (6)$$

where **Q** is the charge, ε the bulk dielectric constant, **k** is the Boltzmann constant, and **b** is is the average axial charge spacing along the contour axis for flexible polymers. It has been experimentally observed that delocalization of cations can occur along the length of the polymer if the nearest neighbor charge spacing is sufficiently close such that two or more groups can cooperate in holding the cation (i.e., effective multivalency)[16]. In $PSer_{20}$, the larger Coulombic sphere of the PSer group permits Na atoms to diffuse and cluster between nearest-neighbor PSer sidechains. This is evidenced in Table IV, where the trend in mean Na-O distances in the phosphate complexes tends towards slightly larger values over those in carboxylate groups. Thus, electrostatic potentials for $PSer_{20}$ peptide structures are more favorable if the peptide collapses, allowing additional Na-bridging interactions to form between adjacent sidechains. Hence, the formation of "hairpin" structures, in which the majority of PSer groups are oriented towards the inside of the hairpin, stabilize electrostatic energies in $PSer_{20}$. Conversely, in Asp_{20}, the Asp carboxylate groups, which are smaller in radii and charge density, cannot participate as easily in the "sharing" of counterions between sidechains since both oxoanions of each carboxylate group are effectively tied up by one Na atom. Thus, in Asp_{20}, there is no electrostatic stabilizing force in condensing the peptide structure; to do so would drive up the van der Waals repulsion energies. Thus, the "supercoil" structure predominates, where only a slight amount of sidechain-Na-sidechain bridging interactions exists. It becomes clear why the "spiral" structure is the most favorable conformer for $PSerAsp_{10}$ at $\varepsilon_o = 1$: this structure features chain segregation of PSer polyion and Asp monovalent groups on either side of the peptide bond [Figure 4]. Thus, at $\varepsilon_o = 1$, the DPG-MC program cannot fold $PSerAsp_{10}$ into a more condensed conformation, since no effective electrostatic attractions exist between Asp sidechains. The only alternative is to promote as much PSer clustering as possible by linearly arranging PSer groups via backbone twisting. As ε_o increases, favorable electrostatic interactions between PSer-Na sidechain complexes are scaled. Thus, linear alignment of PSer groups hold no distinct energetic advantage, and $PSerAsp_{10}$ peptides adopt a modified "spiral" or

compact "hairpin" structure that still features segregation of Asp sidechains from PSer sidechains.

Thus, the Coulombic interactions between charged groups and counterions, and between neighboring charged groups along a polymer chain, directly influence the outcome of the random-walk conformational search. As described below, both of these effects may be desirable where the modeling of polyelectrolytes is concerned.

Discussion

We have approached the problem of defining the conformation of a conformationally elusive group of protein-polymer structures – namely, polyelectrolytes – by the use of a Monte-Carlo/Metropolis sampling algorithm and a generic force-field microscopic description of electrostatic interactions. This initial report describes the simple strategy of employing monovalent counterions in non-bonding interactions with mono- and polyion groups along the length of the peptide chain, and then permitting these "counterions" to remain with each sidechain during dihedral motion in the random walk process. Using this approach, we have found that the value of ε_o does not significantly change the outcome of the Monte-Carlo search, in that all of the accepted lowest energy conformers generated for each ε_o case still exhibit tertiary folding. In addition, the use of ion pairs, i.e., counterions placed in non-bonding geometries with charged sidechains, also simulates the effect of solvent electrostatics (i.e., charge shielding) on peptide conformation.

Agreement of DPG-MC Predicted Polyelectrolyte Structures with Polyelectrolyte Theory and Experimental Data. For each of the sequences tested in the simulation, the accepted structures exhibited significant deviations from the linear or rod structure. In the development of polyelectrolyte theory, the term persistence length (L_p)[46] has been coined to describe the bending flexibility of a polyelectrolyte, and can include terms relating the elastic persistence length of the neutral polyion as well as the electrostatic contribution to the persistence length[46,47]. In theory, as the ionic strength is increased, the electrostatic contributions to the persistence length should decrease as a

direct result of charge shielding. It is apparent that our representation of counterion condensation leads to the generation of what appear to be short persistence length structures that exhibit a great degree of flexibility, i.e., folding [Figures 2, 3]. Therefore, our microscopic treatment of polyelectrolyte counterion condensation appears to be a feasible approach for conformational search studies involving random walk methods. Although very little experimental data is available regarding the solution structures of the sequences utilized in this simulation, it is known from earlier CD studies that poly-L-(Asp) polypeptides exhibit only a small degree of ordering (partial helicity) either at low pH or in the presence of counterions, apparently due to steric repulsion effects[48,49]. This was also observed in our study in the presence of Na^+ (Figures 2), where Asp_{20} demonstrated "supercoil" structures featuring dihedral pairs in the ϕ, ψ region of α-helix. In addition, $(Asp)_n$, $(PSerAsp)_n$, and $(PSer)_n$ sequence regions in the template protein, BDPP, exhibit thermal and solvent denaturability, which would occur if tertiary folding of these sequences was induced by counterion condensation at low ionic strength [Evans and Chan, 1992d]. Thus, the DPG-MC simulation has made a qualitatively correct guess: polyelectrolytes in the counterion condensation state should exhibit tertiary folding.

Conformational Snapshots of Polyelectrolyte Peptides Reveal that the Interconversion Between Lowest Energy Conformers May Be Either "Conservative"or "Liberal." One of the stated goals of this investigation was to examine the low energy conformers and derive a picture of conformational transitions that polyelectrolyte sequences might adopt in solution under conditions of counterion condensation. As shown by the distribution of phi/psi dihedrals and in the accompanying ball-and-stick figure representations [Figure 2,3], Asp_{20} and $PSer_{20}$ most likely convert between supercoil and hairpin structural variants, respectively. How rapidly these interconversions occur, and if any particular secondary structures remain stable during the interconversion (e.g., the turn regions, Figures 2,3), is not known. However, the major point to be gleaned from these results is that a particular sequence may prefer a certain conformer. This raises the possibility that the structural interconversion of $(PSer)_n$ and $(Asp)_n$ in solution may be conservative, i.e., Asp_n will convert between coil structures in solution, but not necessarily into a

hairpin conformation. This preference may exert a great influence on the ligand atom density in that sequence region, and consequently, the attraction of cations to these binding sites. In comparison, the situation regarding the prediction of the $PSerAsp_{10}$ solution structure and conformer interconversion is somewhat more complex, since our simulations predict that the PSer and Asp sidechains preferentially avoid one another in the lowest energy conformers. If this is true, then the $PSerAsp_{10}$ may be more "liberal," in that it could conceivably interconvert between a linear "spiral" structure and a modified "spiral" or "hairpin" [Figure 2,3]. The $PSerAsp_{10}$ peptide is also intriguing from the vantage point of counterion condensation and charge-charge interactions, since the structure is heterogeneous in terms of charged group composition. This may have some bearing on the solution behavior and cation binding properties of BDPP and other PMMP template proteins, perhaps by modulating the affinity of certain charged species for specific regions of the protein. One example can be cited in support of this hypothesis: The alkaline pK_a shifts observed for certain Asp clusters in phosphophoryn, arising as a direct result of neighboring PSer residues exerting short-range electrostatic stabilization of protonation-deprotonation events at adjacent Asp carboxylate groups[25]. Thus, by subtle variations in the Asp and PSer sequence composition of a polyelectrolyte cation binding domain, the tertiary folding of polyelectrolyte sequence regions may be changed in such as way as to bring about variation in oxoanion density, and consequently, metal ion binding and template-induced nucleation[50]. Thus, the use of molecular mechanics and dynamics can assist in the predictions of which particular amino acid sequences will lead to more effective nucleation of certain metals.

From the results obtained in this study, it is also possible that the "condensed" counterion model can be modified to mimic conditions of low ionic strength where sidechain charge repulsion occurs. This may be accomplished either by removing selective counterions, or, by modifying partial atomic charge assignments on either the metal atom or the sidechain atom types. In this way, polyelectrolyte conformation and the resulting effects on osmolarity and mobile-ion activity coefficients[15,16,51,52], viscosity[53], and pK_a's[54] may be examined in greater detail. Future studies in these particular areas are currently in progress.

Acknowledgements

We wish to thank Dr. Siddharth Dasgupta for his comments and helpful advice during the course of this study. JSE acknowledges a Postdoctoral National Research Service Award from the NIH (NIDR DE-05445), and a fellowship award from AMGEN Pharmaceuticals. AMM acknowledges a NIH Predoctoral Biotechnology Traineeship.

References

1. Anfinsen, C.B. (1973) *Science* 181 223-230.

2. Anfinsen, C.B., and Scheraga, H.A. (1975) Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* 29 205-300.

3. Smith, J.L., Hendrickson, W.A., Honzatko, R.B., and Sheriff, S. (1986) *Biochemistry* 25 5018-5027.

4. Creighton, T.E. (1990) Protein folding. *Biochem. J.* 270 1-16.

5. Jaenicke, R. (1991) Protein folding: Local structures, domains, subunits, and assemblies. *Biochemistry* 30 (13) 3147-3160.

6. Alonso, D.O.V., and Dill, K.A. (1991) Solvent denaturation and stabilization of globular proteins. *Biochemistry* 30 5974-5985.

7. Goto, Y., and Hagihara, Y. (1992) Mechanism of the conformational transition of melittin. *Biochemistry* 31 732-738.

8. Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* 6 87-103.

9. Goto, Y., and Nishikiori, S. (1991) Role of electrostatic repulsion in the acidic molten globule of cytochrome c. *J. Mol. Biol.* 222 679-685.

10. Gerstein, M., and Chothia, C. (1991) Analysis of protein loop closure. Two types of hinges produce one motion in lactate dehydrogenase. *J. Mol. Biol.* 220 133-149.

11. Lau, K.F., and Dill, K.A. (1989) A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules* 22 3986-3997.

12. Chan, H.S., and Dill, K.A. (1990) Origins of structure in globular proteins. *Proc. Natl. Acad. Sci USA* 87 6388-6392.

13. Gibson, K.D., and Scheraga, H.A. (1988) The multiple minima problem in protein folding. *in Structure and Expression: Volume I: From Proteins to Ribosomes* (Sarma, M.H., and Sarma, R.H., eds.) Adenine Press, Guilderland, New York pp. 67-94.

14. Paine, G.H., and Scheraga, H.A. (1985) Prediction of the native conformation of a polypeptide by a statistical-mechanical procedure. I. Backbone structure of enkephalin. *Biopolymers* 24 1391-1436.

15. Record, M.T., Anderson, C.F., and Lohman, T.M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the role of ion association or release, screening, and ion effects on water activity. *Quart. Rev. Biophys.* 11 (2) 103-178.

16. Manning, G.S. (1978) The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of nucelotides. *Quart. Rev. Biophys.* 11 (2) 179-246.

17. Smith, P.E., and Pettitt, B.M. (1991) Peptides in ionic solutions: A comparison of the Ewald and switching function techniques. *J. Chem. Phys.* 95 (11) 8430-8441.

18. Schurr, J.M., and Schmitz, K.S. (1986) Dynamic light scattering studies of biopolymers: Effects of charge, shape, and flexibility. *Ann. Rev. Phys. Chem.* 37 271-305.

19. Katchalsky, A. (1971) Polyelectrolytes. *Pure and Applied Chem.* 26 (3-4) 327-372.

20. Schmitz, K.S., and Yu, J-W. (1988) On the electrostatic contribution to the persistence length of flexible polyelectrolytes. *Macromolecules* 21 484-493.

21. Noda, I., Kato, N., Kitano, T., and Nagasawa, M. (1981) Thermodynamic properties of moderately concentrated solutions of linear polymers. *Macromolecules* 14 668-676.

22. Rahima, M., and Veis, A. (1988) Two classes of dentin phosphophoryns, from a wide range of species, contain immunologically cross-reactive epitope regions. *Calcif. Tiss. Int.* 42 104-112.

23. Veis, A. (1988) Biochemical studies of vertebrate tooth mineralization. in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Publishers, New York, New York, pp. 189-216.

25. Evans, J.S., and Chan, S.I. (1992) manuscript in preparation.

26. Hammersley, J.M., and Handscomb, D.C. (1964) *Monte Carlo Methods.* Methuen, London, U.K. pp. 10-25.

27. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953) *J. Chem. Phys.* 21 1087-1092.

28. Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) Domain structure and sequence distribution in dentin phosphophoryn. *Biochem. J.* 276 699-707.

29. Warshel, A., and Aqvist, J. (1991) Electrostatic energy and macromolecular function. *Ann. Rev. Biophys. Biophys. Chem* 20 267-298.

30. Warshel, A., and Russell, S.T. (1985) Calculations of electrostatic energies in proteins. The energetics of ionized groups in bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* 185 389-404.

31. Warshel, A., and Russell, S.T. (1984) Calculations of electrostatic interactions in biological systems and in solutions. *Quart. Rev. Biophys.* 17 (3) 283-422.

32. Mayo, S.L., Olafson, B.D., and Goddard, W.A. (1990) DREIDING: A generic force field for molecular simulations. *J. Phys. Chem.* 94 8897-8909.

33. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alonga, G., Profeta, S., and Weiner, P.J. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106 765-782.

34. Abe, H., Braun, W., Noguti, T., and Go, N. (1984) Rapid calculation of first and second derivatives of conformational energy with respect to dihedral angles for proteins. General recurrent equations. *Computers and Chem.* 8 (4) 239-247.

35. Carrell, C.J., Carrell, H.L., Erlebacher, J., and Glusker, J.P. (1988) Structural aspects of metal ion-carboxylate interactions. *J. Am. Chem. Soc.* 110 8651-8656.

36. Einspahr, H., and Bugg, C.E. (1980) *Acta. Cryst.* B36 264-271.

37. Einspahr, H., and Bugg, C.E. (1981) *Acta. Cryst.* B37 1044-1052.

38. Rappe, A.K., and Goddard, W.A. (1991) Charge equilibration for molecule dynamics simulation. *J. Phys. Chem.* 95 3358-3363.

39. Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling W.T. (1986) *Numerical Recipies: The Art of Scientific Computing.* Cambridge University Press, Cambridge, England pp.274-328.

40. Bellido, M.N., and Rullmann, J.A.C. (1989) Atomic charge models for polypeptides derived from Ab Initio calculations. *J. Comput. Chem.* 10 (4) 479-487.

41. Solmajer, T., and Mehler, E.L. (1991) Electrostatic screening in molecular dynamics simulations. *Protein Engineering* 4 (8) 911-917.

42. Mehler, E.L., and Solmajer, T. (1991) Electrostatic effects in proteins: comparison of dielectric and charge models. *Protein Engineering* 4 (8) 903-910.

43. Harvey, S.C. (1989) Treatment of electrostatic effects in macromolecular modeling. *Proteins: Struct. Funct. and Genet.* 5 78-92.

44. Mehler, E.L. (1990) Comparison of dielectric response models for simulating electrostatic effects in proteins. *Protein Engineering* 3 (5) 415-417.

45. Matthew, J.B. (1985) Electrostatic effects in proteins. *Ann. Rev. Biophys. Biophys. Chem* 14 387-417.

46. Schmitz, K.S., and Yu, J-W. (1988) On the electrostatic contribution to the persistence length of flexible polyelectrolytes. *Macromolecules* 21 484-493.

47. Le Bret, M. (1982) Electrostatic contribution to the persistence length of a polyelectrolyte. *J. Chem. Phys.* 76 (12) 6243-6255.

48. Jacobson, A.L. (1965) Magnesium binding to poly(L-aspartic acid). *Biopolymers* 3 249-259.

49. Berger, A., and Katchalski, E. (1951) Poly-L-aspartic acid. *J. Am. Chem. Soc.* 73 4084-4088.

50. Mann, S. (1988) Crystallochemical strategies in biomineralization. in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P., eds.) VCH Publishers, New York, New York, pp. 35-61.

51. Manning, G.S. (1969) Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties. *J. Chem. Phys.* 31 (1) 924-933.

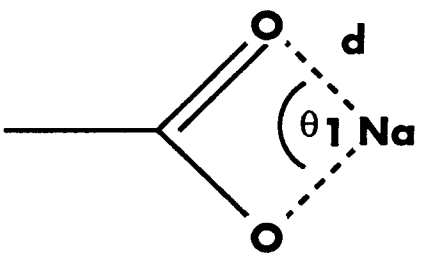
52. Manning, G.S. (1969) Limiting laws and counterion condensation in polyelectrolyte solutions. II. Self-diffusion of the small ions. *J. Chem. Phys.* 31 (1) 934-938.

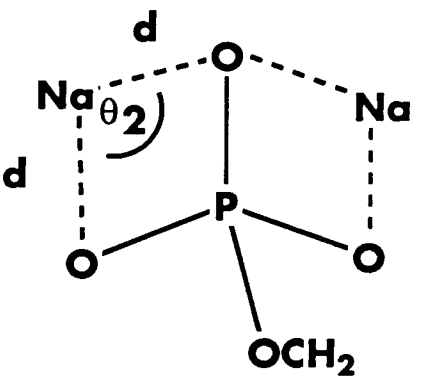
53. Davis, R.M., and Russel, W.B. (1987) Intrinsic viscosity and Huggins coefficients for potassium poly(styrenesulfonate) solutions. *Macromolecules* 20 518-525.

54. Manning, G.S., and Holtzer, A. (1973) Application of polyelectrolyte limiting laws to potentiometric titration. *J. Phys. Chem.* 77 (18) 2206-2212.

Figure 1: *Starting Na-O Geometries in Polyelectrolyte Peptides.*

The bridging geometry for 1 Na:1 Asp carboxylate and 2 Na: 1 PSer monophosphate ester are shown.





d = 2.5 Å

Table I: *Mean Partial Atomic Charges Obtained for Backbone and Sidechain Atoms in Na+ Complexed Peptides.*

Partial atomic charges were determined by the charge equilibration routine of Rappe and Goddard[38] for each Na:peptide sequence class in the extended conformation. Values (in electron units) represent the mean charge value (± S.D., where N = 20) determined for the specified atom types of the peptide backbone and in each of the Asp COO and PSer PO_4 groups in a specific peptide sequence. **C_α C_31**: C_α atom. **Carbonyl C_R**: C_1 carbonyl carbon atom. **Carbonyl O_2**: carbonyl oxygen. **Amide N_R**: amide nitrogen. **Amide H___A**: amide hydrogen. **O_2A**: "anionic" oxygen atoms of the PO_4 group. **O_2E**: "ester" oxygen atom of the PO_4 group. **Na_1**, **Na_2**: sodium atoms of the PO_4 complex , individually averaged.

Table I

Mean Partial Atomic Charges for Backbone Atoms in Na Complexed Peptides

Peptide	Atomic Charge (± S.D.)				
	C_α C_31	**Carbonyl C_R**	**Carbonyl O_2**	**Amide N_R**	**Amide H___A**
Asp_{20}	0.04 ± 0.003	0.45 ± 0.005	-0.53 ± 0.007	-0.51 ± 0.003	0.24 ± 0.006
$PSerAsp_{10}$	0.05 ± 0.01	0.47 ± 0.03	-0.59 ± 0.02	-0.49 ± 0.02	0.22 ± 0.01
$PSer_{20}$	0.10 ± 0.01	0.44 ± 0.02	-0.58 ± 0.02	-0.49 ± 0.02	0.22 ± 0.02

Table I

Mean Partial Atomic Charges for Sidechain Atoms in Na Complexed Peptides

Peptide	**Atomic Charge (±S.D.)**										
	C_R	**O_2**	**O_2**	**Na**	**P**	**O_2 A**	**O_2 A**	**O_2 A**	**O_2 E**	**Na_1**	**Na_2**
Asp_{20}	+0.59 ±0.003	-0.62 ±0.005	-0.54 ±0.02	+0.81 ±0.03	———	———	———	———	———	———	———
$PSerAsp_{10}$	+0.60 ±0.005	-0.59 ±0.008	-0.59 ±0.003	+0.86 ±0.02	+0.48 ±0.01	-0.48 ±0.01	-0.49 ±0.02	-0.49 ±0.01	-0.42 ±0.01	+0.68 ±0.08	+0.78 ±0.39
$PSer_{20}$	———	———	———	———	+0.52 ±0.03	-0.48 ±0.02	-0.49 ±0.08	-0.48 ±0.03	-0.45 ±0.03	+0.72 ±0.15	+0.79 ±0.13

Table II: *Statistical Properties of DPG-MC Accepted Conformers.*

All DREIDING II energy values are given in kcal/mole. The five best conformers are examined. **Ratio %** is the percentage of the total number of conformers which fit the criteria of accepted structures. **S.D.** = standard deviation, for N = 5. **Total energy:** the sum of the internal bonding and non-bonding energy terms. ΔE_{total} : the difference in total energies between the extended starting conformer and the accepted conformers. **vdW**: van der Waals potential. **H-Bond**: energy of hydrogen bond interactions between donor and acceptor atoms. **Bond**: summation of bond potentials[32]:

$$E_{Bond} = E_{harmonic} + E_{angle} + E_{torsion} + E_{inversion}$$

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o	Number of accepted conformations	Ratio (%)		Total Energy	ΔE_{Total}
Asp20	1	145	0.008	Mean	-810.5	-361.9
				SD	±24.9	±24.9
				Upper value	-770.2	-321.5
				Lower value	-836.8	-388.1
	4	325	0.033	Mean	-43.1	-119.6
				SD	±11.2	±11.2
				Upper value	-28.7	-105.3
				Lower value	-64.7	-141.1
	10	344	0.034	Mean	-66.1	-102.7
				SD	±3.0	±7.2
				Upper value	-61.8	-94.3
				Lower value	-72.0	-116.3

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o		Electrostatic Energy	vdW Energy	H-Bond Energy	Bond Energy
Asp20	1	Mean	-1078.0	165.0	-45.3	147.8
		SD	±15.2	±6.4	±11.0	±2.1
		Upper value	-1050.2	172.8	-30.9	151.7
		Lower value	-1108.2	154.2	-66.0	143.7
	4	Mean	-191.8	84.9	-45.7	109.4
		SD	±5.4	±3.2	±5.5	±2.8
		Upper value	-181.7	89.0	-34.3	114.8
		Lower value	-204.6	78.8	-57.3	108.1
	10	Mean	56.9	72.4	-49.4	100.1
		SD	±7.2	±6.0	±4.1	±2.2
		Upper value	65.2	80.3	-41.7	106.1
		Lower value	43.3	57.5	-55.2	98.0

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o	Number of accepted conformations	Ratio (%)		Total Energy	ΔE_{Total}
PSerAsp10	1	57	0.003	Mean	-928.9	-233.1
				SD	±17.9	±12.0
				Upper value	-915.0	-166.6
				Lower value	-958.9	-288.7
	4	238	0.030	Mean	-94.2	-87.8
				SD	±26.2	±25.6
				Upper value	128.7	-53.2
				Lower value	54.7	-127.1
	10	314	0.031	Mean	259.4	-65.7
				SD	±13.3	±13.3
				Upper value	278.2	-46.8
				Lower value	239.7	-85.3

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o		Electrostatic Energy	vdW Energy	H-Bond Energy	Bonding Energy
PSerAsp10	1	Mean	-1486.6	173.5	-20.5	404.2
		SD	±154.4	±32.3	±5.1	±3.7
		Upper value	-1468.5	177.7	-7.6	411.9
		Lower value	-1509.7	165.6	-25.3	397.2
	4	Mean	-309.2	63.9	-19.5	366.1
		SD	±16.0	±10.4	±8.6	±8.3
		Upper value	-286.1	78.2	-20.6	372.2
		Lower value	-331.4	49.6	-35.4	361.4
	10	Mean	-100.2	44.5	-43.3	358.4
		SD	±2.7	±7.5	± 3.6	±1.6
		Upper value	-97.1	51.6	-34.3	361.6
		Lower value	-107.1	26.2	-47.6	356.3

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o	Number of accepted conformations	Ratio (%)		Total Energy	ΔE_{Total}
PSer20	1	25	0.0014	Mean	-1247.5	-199.4
				SD	±21.9	±21.9
				Upper value	-1214.7	-166.6
				Lower value	-1276.8	-228.7
	4	449	0.045	Mean	108.6	-76.7
				SD	±22.3	±22.6
				Upper value	143.9	-62.9
				Lower value	66.5	-118.8
	10	314	0.031	Mean	336.4	-23.6
				SD	±23.6	±12.5
				Upper value	361.1	-2.0
				Lower value	312.2	-46.9

Table II

Statistical Properties of DPG-MC Generated Peptide Conformations

Structure	Dielectric, ε_o		Electrostatic Energy	vdW Energy	H-Bond Energy	Bonding Energy
PSer20	1	Mean	-2006.7	119.1	-39.4	679.5
		SD	±32.9	±8.9	±3.5	±3.3
		Upper value	-2056.1	131.1	-34.6	682.8
		Lower value	-1963.0	106.1	-44.7	674.6
	4	Mean	-407.6	-37.9	-62.9	616.9
		SD	±14.0	±9.5	±8.0	±4.9
		Upper value	-394.3	-29.0	-54.5	622.1
		Lower value	-426.8	-50.1	-80.6	606.1
	10	Mean	-140.2	-62.5	-49.4	608.7
		SD	±2.3	±8.2	±20.2	±2.9
		Upper value	-138.0	-55.0	-55.0	604.7
		Lower value	-145.9	-77.6	-77.6	612.2

Table III: *DREIDING II Internal Energy and POLARIS Free Energy of Solvation ($\Delta\Delta G_p^{sol}$) Calculated for $\varepsilon_o = 1$.*

All energies given in kcal/mole.

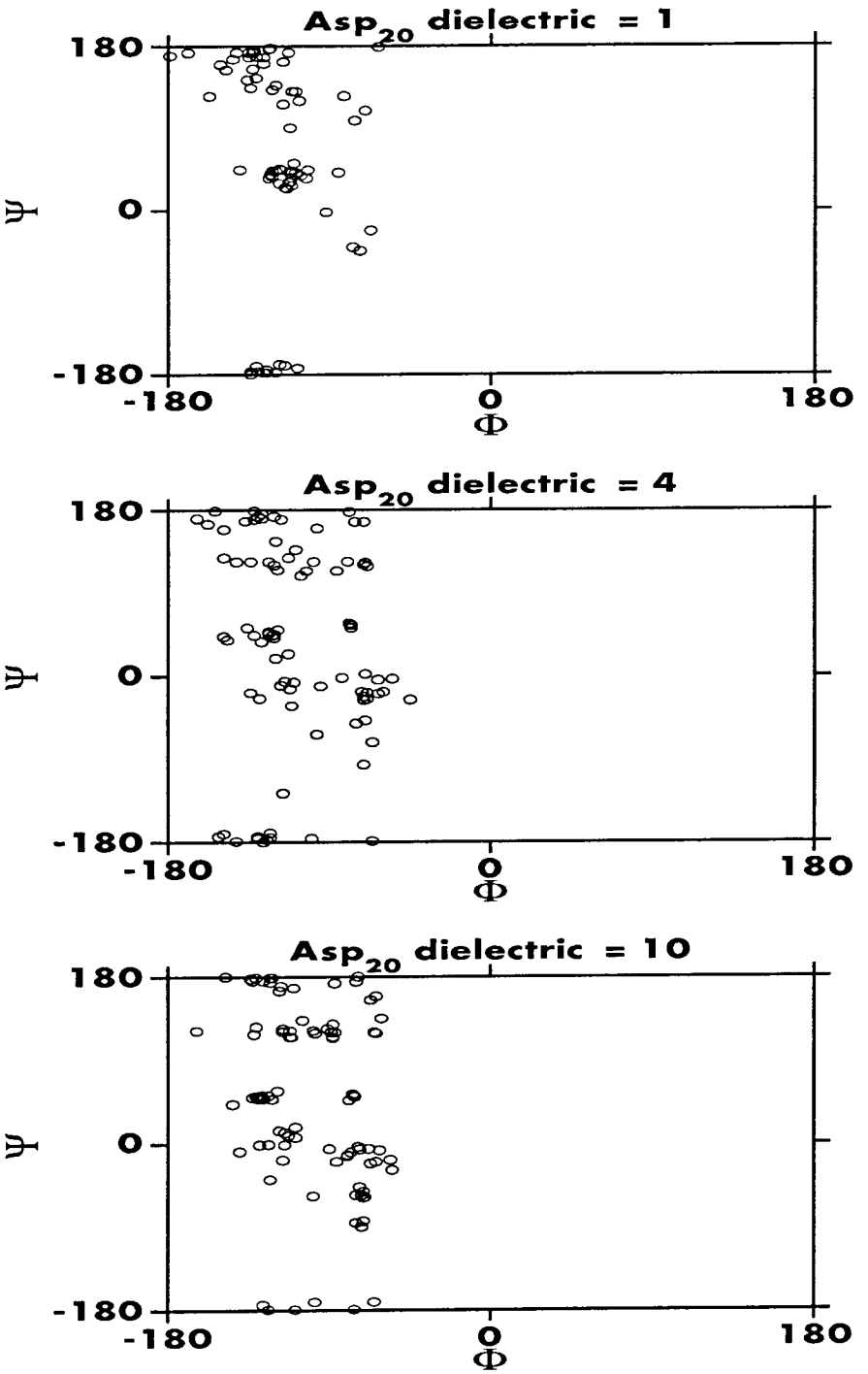
Table III

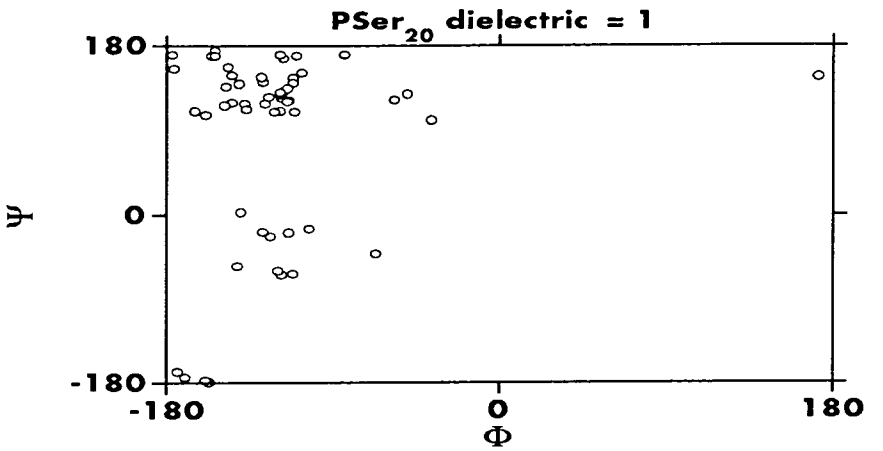
Internal Energy and Free Energy of Solvation ($\Delta\Delta G_p^{sol}$) Calculated for $\varepsilon_o = 1$

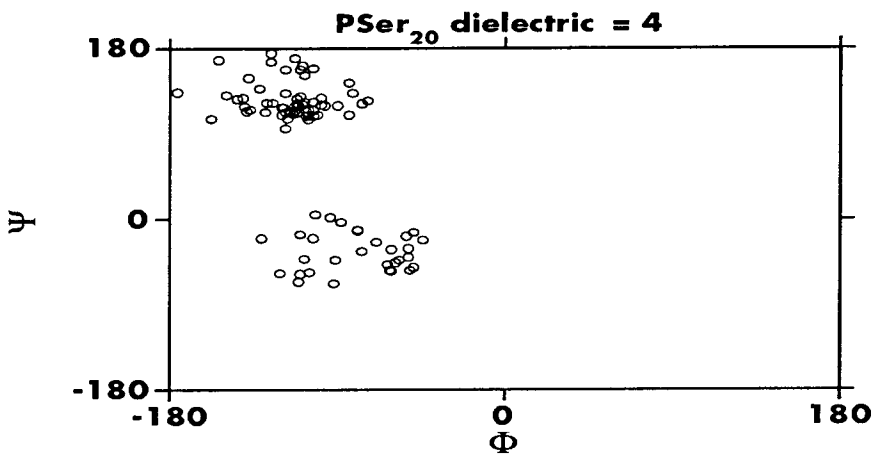
Structure	Former ε_o	E_{vacuo}	$\Delta\Delta G_p^{sol}$	$E_{vacuo} + \Delta\Delta G_p^{sol}$
Asp_{20}	1	-817.0 ± 25.0	-433.2 ± 12.7	-1250.3 ± 16.1
	4	-615.3 ± 29.1	-562.3 ± 29.0	-1185.6 ± 23.5
	10	-541.7 ± 31.2	-551.9 ± 21.5	-1093.6 ± 45.1
$PSerAsp_{10}$	1	-941.6 ± 15.2	-496.9 ± 25.8	-1438.5 ± 19.6
	4	-781.5 ± 60.9	-570.7 ± 11.6	-1352.2 ± 69.7
	10	-640.8 ± 34.3	-656.1 ± 34.3	-1275.0 ± 26.2
$PSer_{20}$	1	-1259.3 ± 27.2	-557.8 ± 28.2	-1817.0 ± 11.5
	4	-1119.0 ± 65.5	-630.6 ± 33.7	-1749.6 ± 31.8
	10	-930.4 ± 30.4	-700.7 ± 27.4	-1631.1 ± 18.1

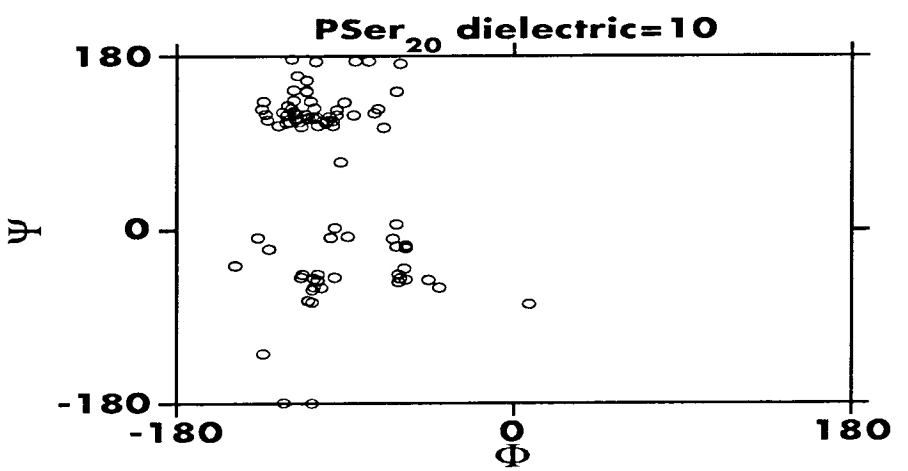
Figure 2: *Phi/Psi Dihedral Scatterplots for DPG-MC Accepted Conformers.*

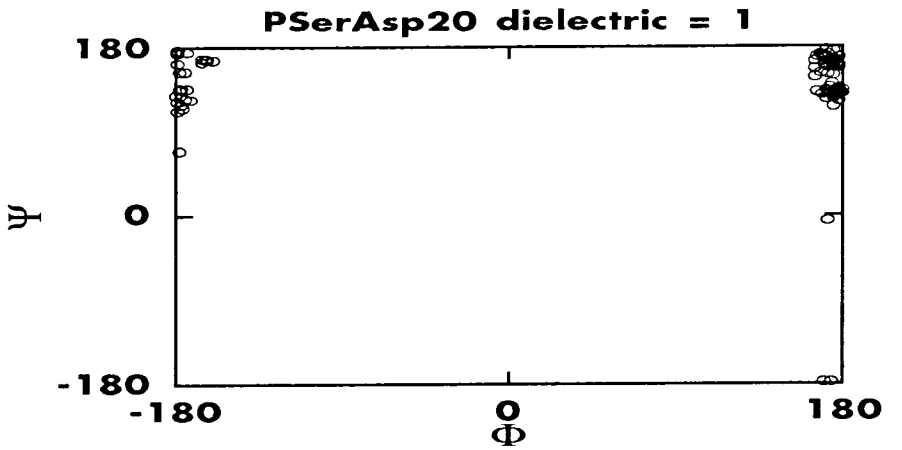
Dihedral pairs were obtained for all accepted structures for each polypeptide sequence and dielectric constant studied. For cases ε_o = 4 or 10, N = 95; for ε_o = 1, N = 76, where *N* is the number of dihedral pairs.

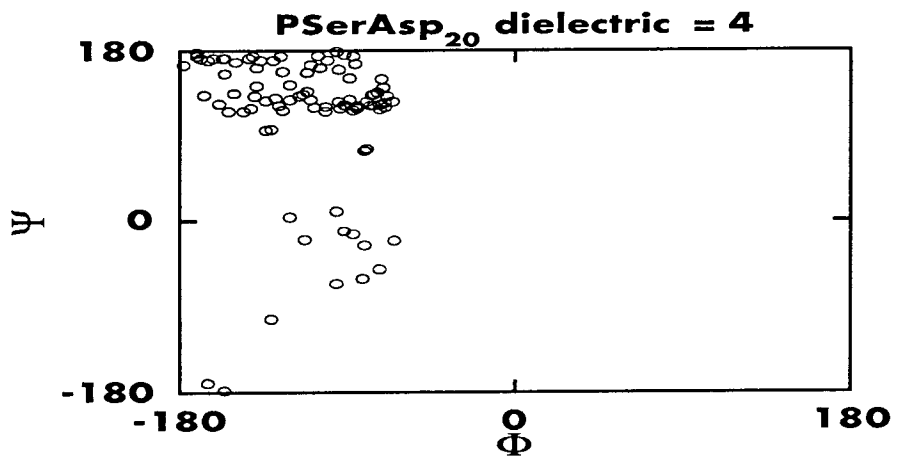

Asp20 dielectric = 1
180
0
-180
Ψ
-180
0
180
Φ
Asp20 dielectric = 4
180
0
-180
Ψ
-180
0
180
Φ
Asp20 dielectric = 10
180
0
-180
Ψ
-180
0
180
Φ


PSer20 dielectric = 1
180
0
-180
Ψ
-180
0
180
Φ


PSer20 dielectric = 4
180
0
-180
Ψ
-180
0
180
Φ


PSer20 dielectric=10
180
0
-180
Ψ
-180
0
180
Φ


PSerAsp20 dielectric = 1
180
0
-180
Ψ
-180
0
180
Φ


PSerAsp20 dielectric = 4
180
0
-180
Ψ
-180
0
180
Φ

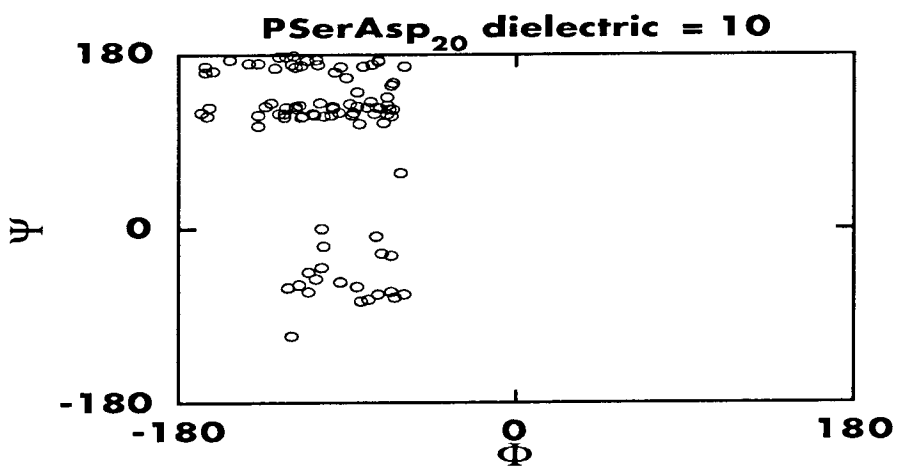
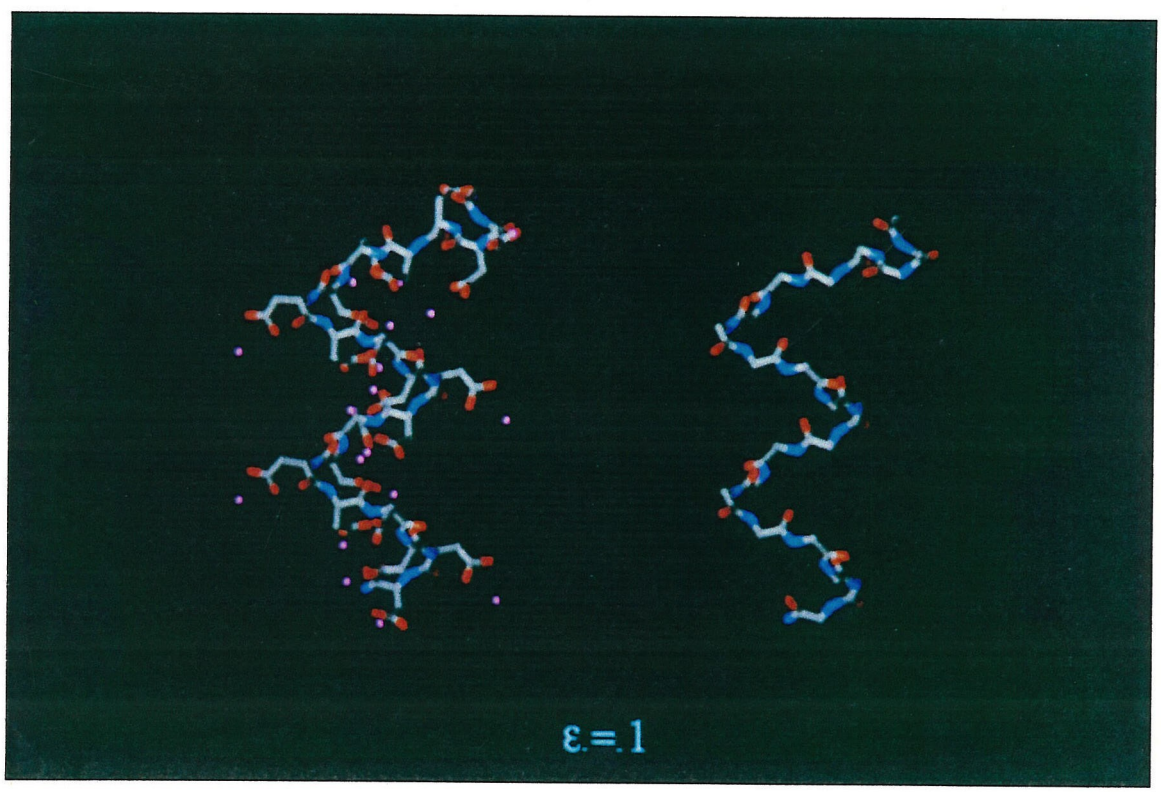
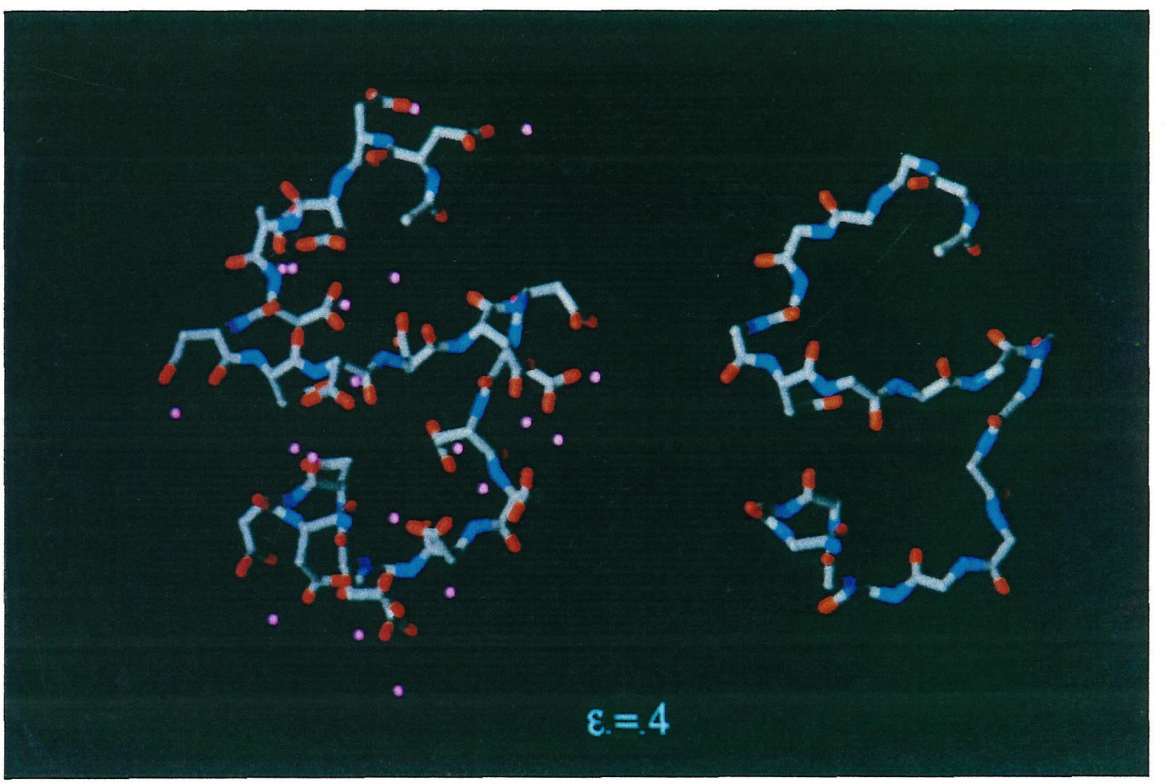
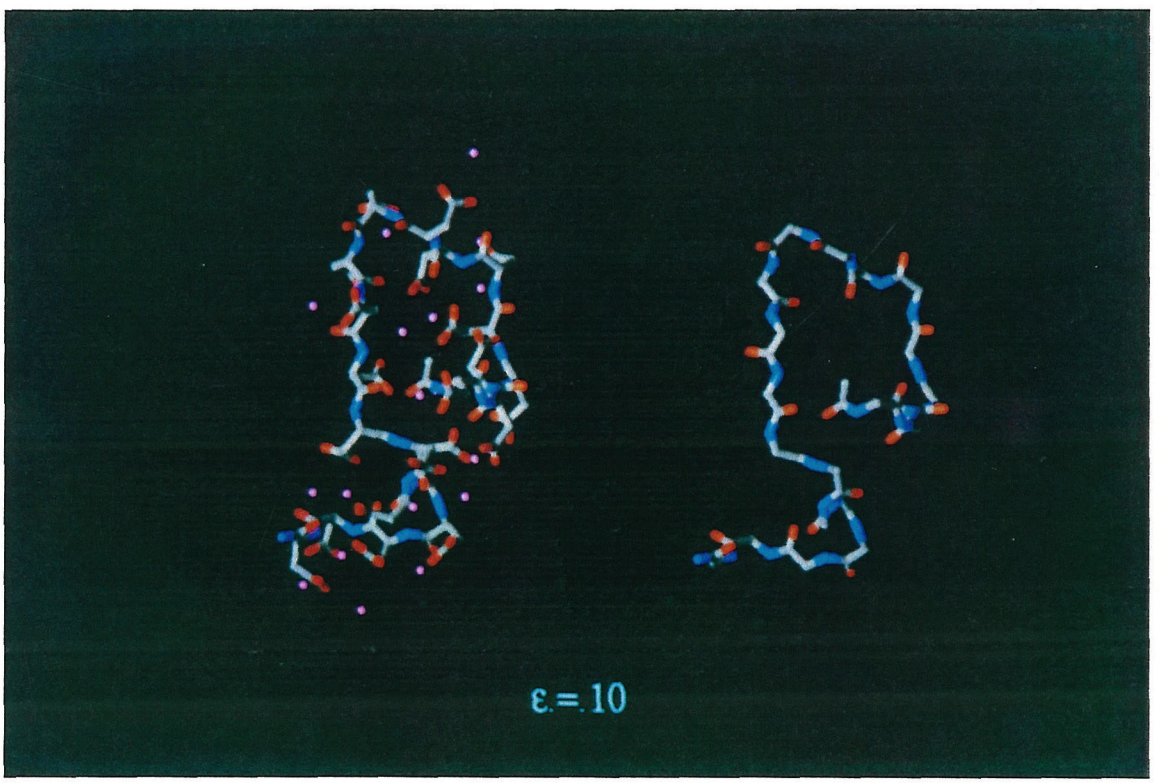

PSerAsp20 dielectric = 10
180
0
-180
Ψ
-180
0
180
Φ

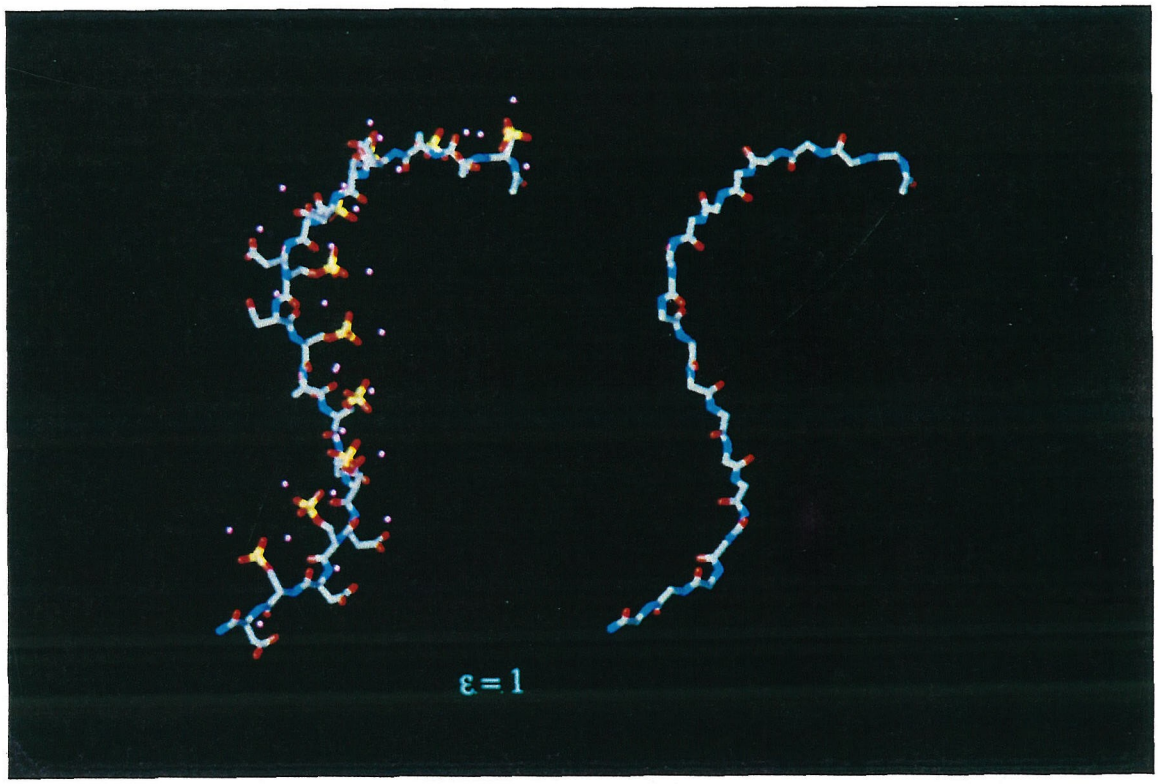
Figure 3: *Ball-Stick Representations of "Best" Accepted Polyelectrolyte Polypeptide Conformers.*

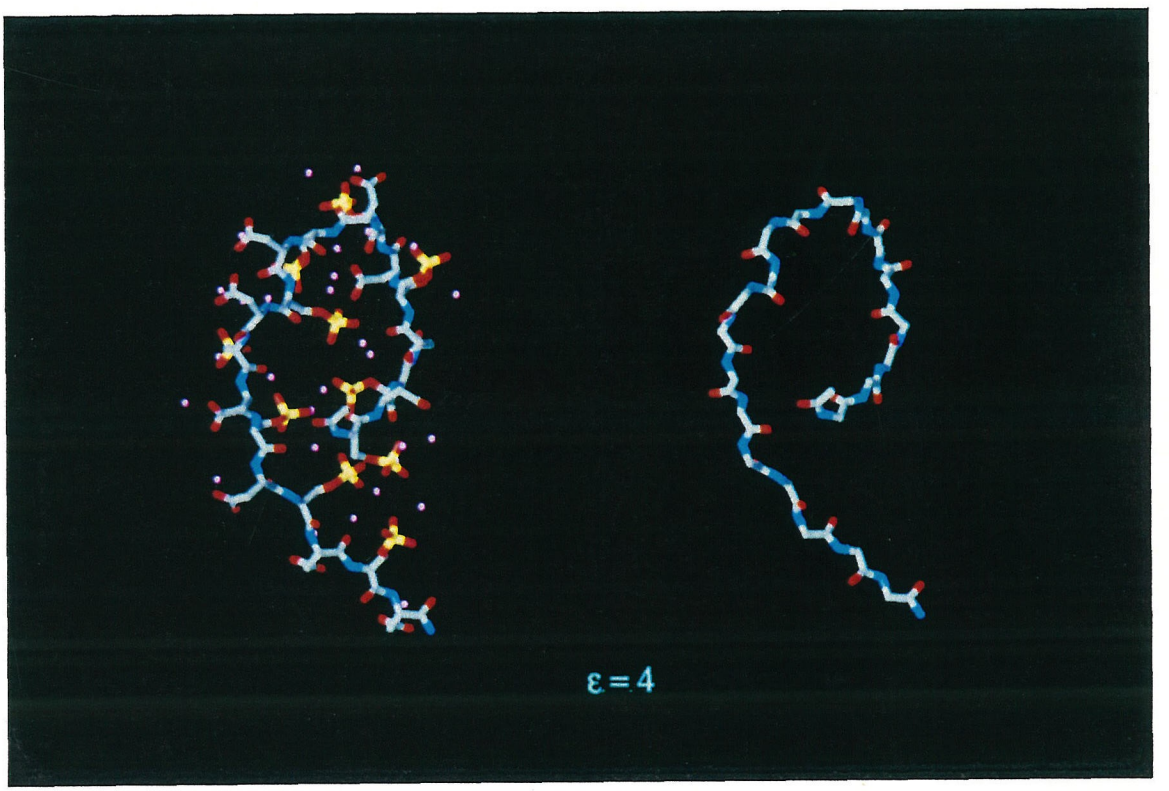
Each plate shows the structure with sidechain groups and Na atoms, *sans* H atoms (Left), and the backbone projection alone (Right). (**A**). N-acetyl-L-$(Asp)_{20}$-C-amide; (**B**) N-acetyl-L-$(PSerAsp)_{10}$-C-amide; (**C**) N-acetyl-L-$(PSer)_{20}$-C-amide. Atom type color legend: **Yellow**: P; **Blue**: N_R; **Red**: O_2; **Lavender**: Na; **Grey or Lite Blue**: C_31 or C_R. Corresponding ε_o values are given at the bottom of each Figures. An example of a "supercoil" structure is given in (**A**), for ε_o = 1,4, and 10. An example of a "hairpin" structure is shown in (**B**), for ε_o = 1; a "distorted hairpin" is shown in the same Figure, for ε_o = 4, 10.

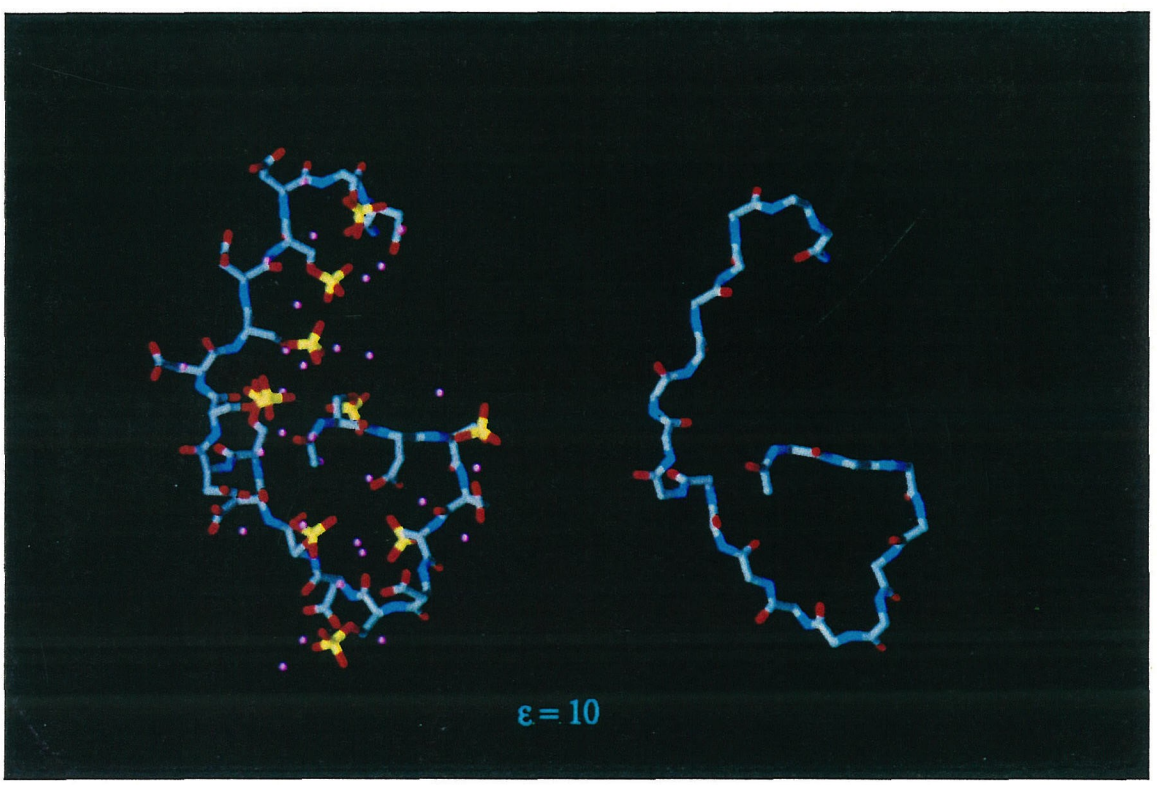

ε=.1

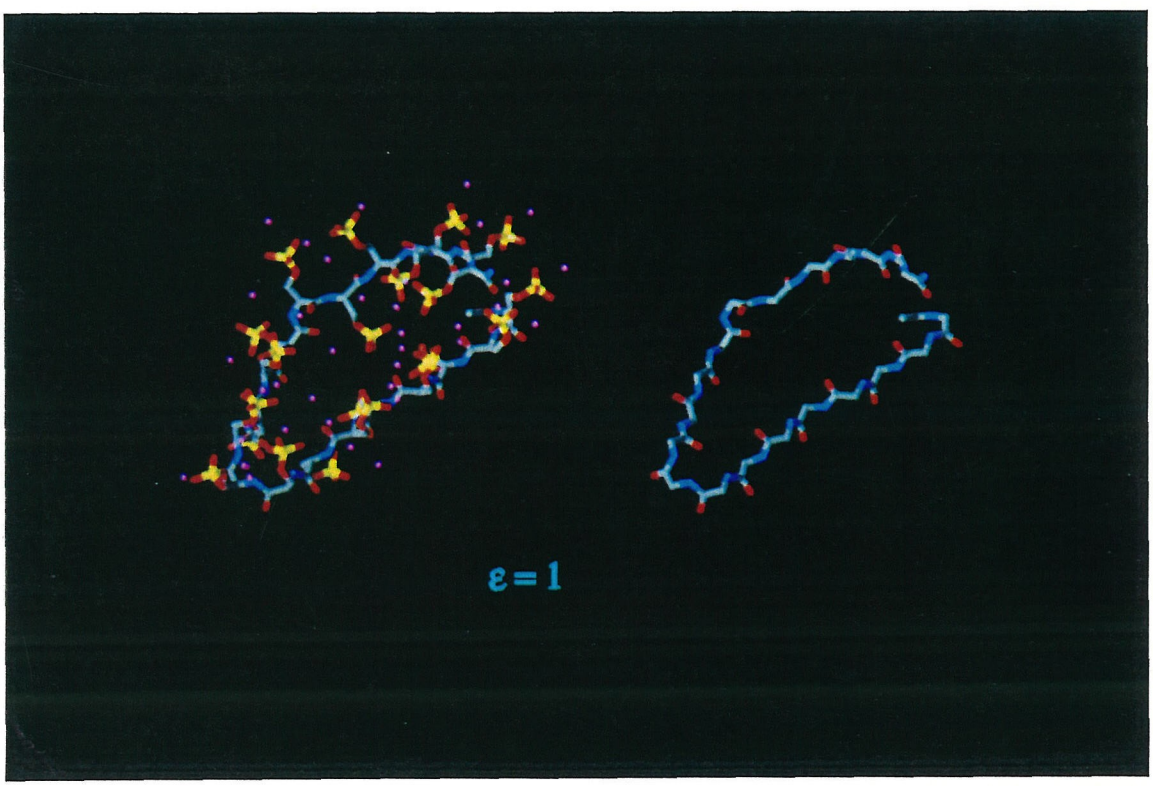

ε.=.4

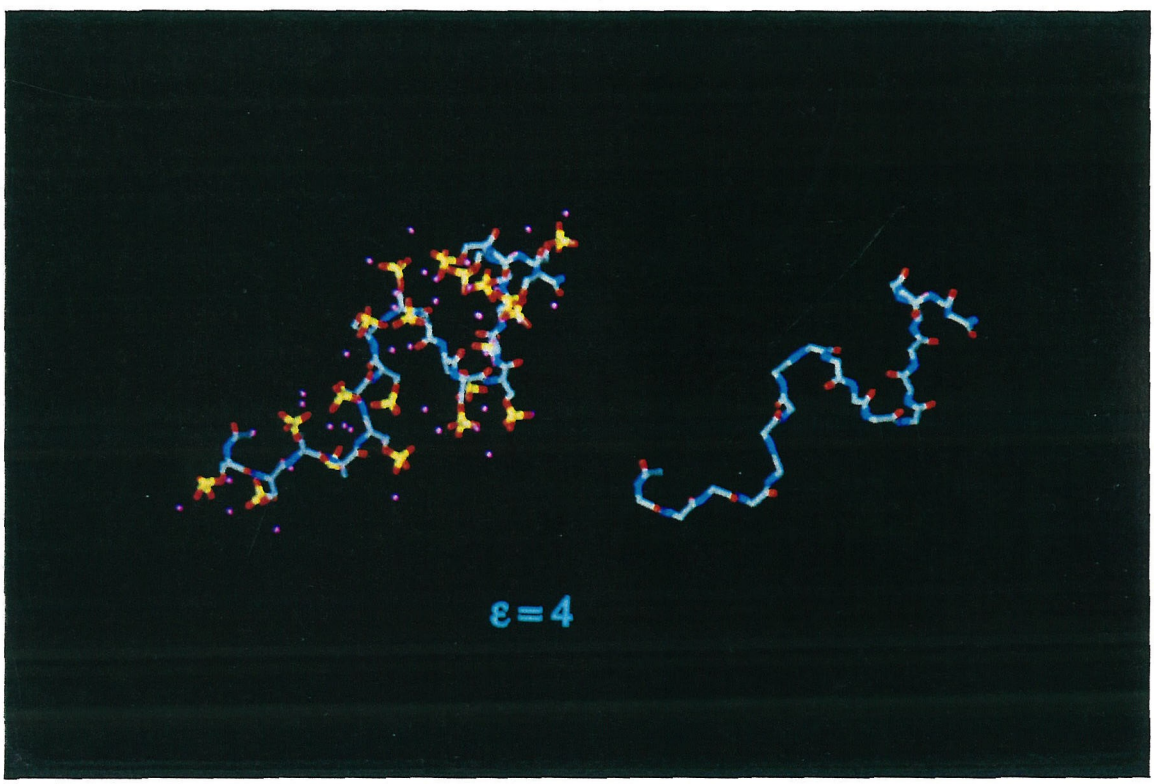

ε.=.10


ε=1


ε = 4


ε = 10


ε = 1


ε=4

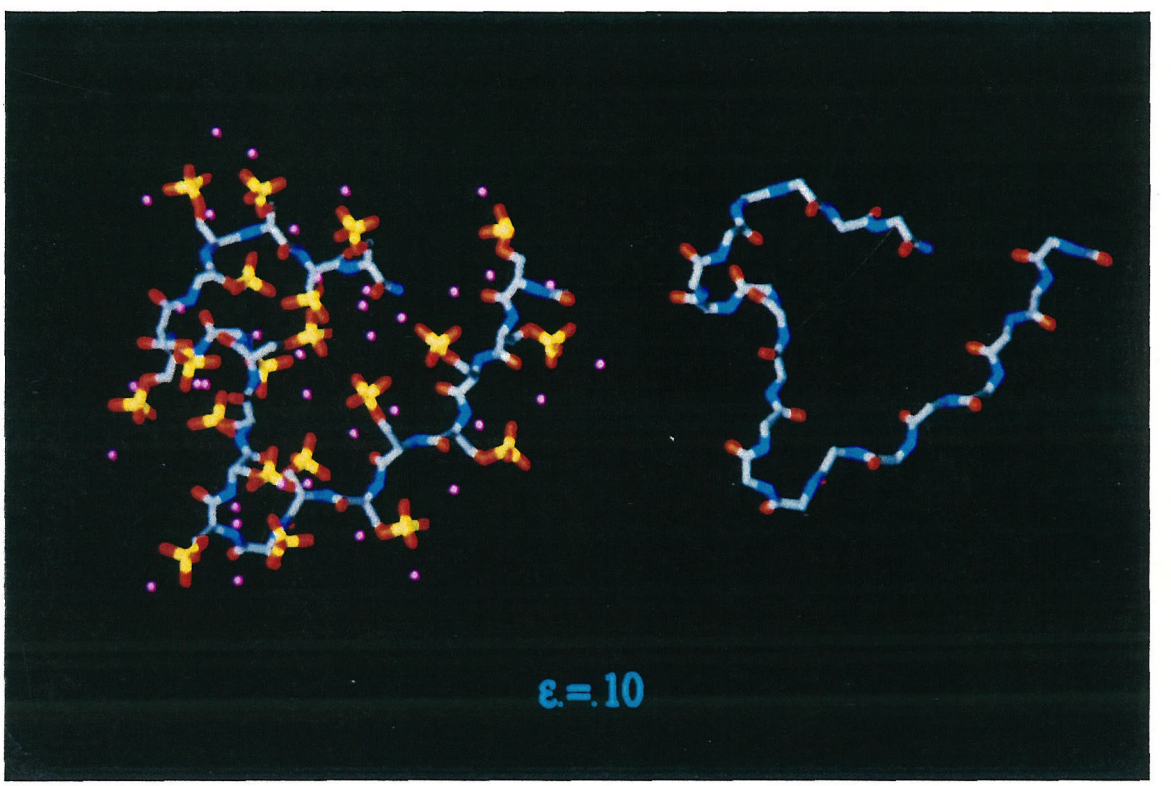

ε.=.10

Table IV: *Mean Sodium-Oxygen Distances.*

Distances (in Å) between O_2 and Na atom types were measured for every sidechain-Na complex in each of the accepted structures for the *in vacuo* case. **COO** and **PO_4** represent the O_2 - Na distances for the carboxylate and monophosphate ester complexes in $PSerAsp_{10}$.

Table IV

Mean Metal-Oxygen Distances in DPG - MC Generated Peptide Conformers

Structure	Mean Sidechain d(O-Na), Å
Asp_{20}	2.77
$PSer_{20}$	3.44
$PSerAsp_{10}$	COO: 2.75 PO_4: 3.31

Eight

Bovine Dentin Phosphophoryn, a Polyelectrolyte Mineral Matrix Protein.

Conformational Folding in the Presence of ^{113}Cd (II).

Introduction

The concept of a template, or molecular surface upon which a new chemical structure is synthesized, has been avidly explored in an area of biology known as biomineralization. Here, the template concept has been at work for at least 80 million years (Weiner and Hood, 1975). The most widely studied of these biomineral templates are a class of proteins, the *polyelectrolyte mineral matrix proteins* (PMMP), that are localized in the extracellular matrix of a diverse array of biogenic minerals (Mann, 1988). These interesting proteins possess significant proportions of Asp, Glu, and/or monophosphate esters of Ser, Thr, and Tyr (Shiraga *et al.*, 1992; George *et al.*, 1992; Borbas *et al.*, 1991; Gotoh *et al.*, 1990; Zhang *et al.*, 1990; Rahima and Veis, 1988; Constantz and Weiner, 1986; Swift *et al.*, 1986; Veis *et al.*, 1986; Worms and Weiner, 1986; Marsh and Sass, 1984; Stetler-Stevenson and Veis, 1983; Weiner, 1983; Weiner and Hood, 1975), that generate a high negative charge density over the surface of the protein and a large number of metal ion binding sites ($\approx 10^2$-10^3/mol protein)(Mann, 1988.). The PMMP protein has been observed both in vivo and in vitro to act as a surface for the coordination of metal ions in a precise stereochemical array , and, to bond to inorganic solid surfaces along a preferred crystallographic axis, thus inhibiting further ion apposition (Addadi and Weiner, 1985; Addadi *et al,*, 1987; Addadi and Weiner, 1988; Addadi *et al.*, 1992).

The presence of divalent cations has been shown to induce a conformational transition from the "random coil" state to the "predominantly" β-sheet structure in a number of PMMP proteins (e.g., phosphophoryn (Lee *et al.*, 1977; Stetler-Stevenson and Veis, 1987); invertebrate shell calcite-derived proteins (Worms and Weiner, 1986)) and in polyanionic sequence peptides that mimic Ca (II) binding domains in the PMMP proteins (e.g., Poly-L-$(Asp)_n$, Poly-$(Ser\text{-}Glu)_n$ (Addadi *et al.*, 1987, 1991, 1992; Lee and Veis, 1980; Addadi and Weiner, 1985; Nancollas, 1988)). A significant percentage of β-sheet

conformation has also been detected in calcite and aragonite-associated PMMP proteins and in polyanionic sequence peptides that are bound to inorganic calcium mineral phases (Addadi*et al.*, 1987; Addadi and Weiner, 1988; Weiner and Traub, 1984). Although the β-sheet structure has been shown to comprise a significant proportion (60% or more) of the Ca(II) loaded PMMP conformation (Worms and Weiner, 1986), the distribution of β-sheet structure over the length of the PMMP molecule is not known.

The major impetus for studying the conformation of PMMPs in the presence of metal ions is to ascertain how the PMMP protein folds into a three-dimensional array of ligand groups that stereochemically coordinates metal atoms. Protein folding can be conceptualized as a sequential pathway where short residue stretches of the sequence combine to form local "supersecondary" structures, which then combine or merge to form structural intermediates that eventually acquires a final structure (Dill, 1990; Kuwajima, 1989). In general, protein folding involves the presence of either a limited number of thermodynamically stable structural states, i.e., the folded and unfolded state (for small globular proteins < 20 kD)(Freire and Murphy, 1991; Privalov, 1979; Freire, 1990), or in the case of larger multidomain proteins, there may be a few partially folded intermediates in addition to the endpoint conformational states (Dill, 1990). In some instances, these multidomain intermediate states may correspond to some of the structural domains existing in a folded state, and other domains, an unfolded state (Privalov, 1982; Ramsay and Freire, 1990). If we assume that the PMMP protein possesses multiple binding domains for metal ions, then the possibility arises that (1) there may be specific metal-ion induced folding pathway(s) that lead to a stereochemically correct conformation, and (2) there may exist intermediate folding states for the PMMP molecule at various stages of metal ion loading. Identifying the pathway(s) for metal ion induced PMMP folding can provide useful information regarding PMMP-regulated nucleation or inhibition mechanisms, and, can be applied in designing effective synthetic template molecules based on the PMMP mechanism.

To date, very little is known regarding the metal ion-induced folding pathway for PMMP proteins. However, solution NMR spectroscopy experiments conducted by us on a PMMP protein, bovine dentine phosphophoryn (abbreviated as BDPP, MW=155 kD)(Evans and Chan, 1992a,b,c), have provided information regarding the divalent-cation depleted protein conformation, *viz.*: (A) the existence of three types of regions in the phosphophoryn molecule: (i) *polyelectrolyte cluster Ca (II) binding domains* (abbreviated as **PCBD**), containing $(Asp)_n$, $(PSer\text{-}Asp)_n$, and $(PSer)_n$ sequence blocks, (ii) *"hinge"* or *intervening domains* (abbreviated as **HD**) which are flanked on either side by a polyelectrolyte region, and (iii) a *hydrophobic cluster* region (abbreviated as **HC**) (Evans and Chan, 1992a,b). (B) At neutral pH, the global conformation of the divalent cation-depleted state is extended, and, at low pH, the global conformation transforms into a condensed or folded state. (C) The HD region mediates the folding of adjacent PCBD regions in response to charge repulsion or charge shielding, and, the HC is stabilized against pH-induced conformational change. The divalent cation depleted conformation can be thought of as the "initial" or "precursor" state of BDPP that will traverse a specific folding pathway to the "cation-loaded" state upon addition of Ca (II). It should then be possible to use multinuclear NMR spectroscopy to monitor conformational change in the three principle PMMP regions as a function of divalent cation addition. In this present manuscript, we report on preliminary $^{1}H/^{31}P/^{13}C/^{113}Cd$ NMR studies of BDPP in the presence of the Ca (II) analog, ^{113}Cd (II). As described herein, BDPP exhibits region-specific conformational transitions at different protein:ion stoichiometries.

Materials and Methods

Materials. The reagents used in the present study were obtained from the following sources: D_2O (99.9% atom) and ^{113}CdO (91.67% atom) from Cambridge Isotopes, deuterium chloride (DCl, 20% v/v solution, 98% atom), sodium deuteroxide (NaOD, 30% v/v solution, 98% atom)

from Aldrich Chemicals (Milwaukee, WI). ^{113}CdO (95% atom), from MSD Istopes, HCl and NaOH from Fisher Scientific, Chelex-100 resin was purchased from Bio-Rad (Richmond, CA). All buffers and protein solutions were prepared from 0.2 μm filtered deionized distilled water (DD_{water}). $^{113}CdCl_2$ stock solutions (2.5 M and 250 mM) were prepared by stoichiometric addition of concentrated HCl to ^{113}CdO solid, then diluting to the appropriate volume with DD_{water}. Cadmium stock concentrations and Cd (II) levels in BDPP samples were verified by ICP analysis. All glassware used in NMR sample preparation and analysis was first treated with EDTA/EGTA to remove unwanted divalent cations. The isolation and preparation of Ca (II) depleted BDPP from unerupted bovine third molars for the NMR experiments has been previously described (Evans and Chan, 1992a). For Cd (II) titration experiments, a 18.6 mg/ml BDPP sample was dissolved in 400 μl of DD_{water} (final protein concentration = 0.3 mM, based on a molecular weight of 155 kD (Stetler-Stevenson and Veis, 1983) containing 10% /v/ D_2O for field frequency stabilization. The pH was adjusted to 7.20 with NaOH stock solution, resulting in a low-ionic strength solution ($[Na^+]$ = 1 mM, verified by ICP analysis).

1-H,31-P, and 13-C NMR Cd (II) Titration Experiments. ^1H (500.138 MHz) /^{31}P (202.458 MHz)/^{13}C (125.759 MHz) NMR spectra were acquired at 300 K as described elsewhere (Evans and Chan, 1992a) on a Bruker AMX-500 Spectrometer and Aspect X32 computer, using a 5 mm inverse low-gamma probehead. The BDPP sample was titrated using microliter volumes (< 5 μl) of either the 250 mM or 2.5 M $^{113}CdCl_2$ stock solution. After ion addition, the pH was checked, using a Radiometer PHM83 Autocal pH Meter and Ingold 6030-02 glass microelectrode, and if necessary, readjusted to pH 7.20 ± 0.02 with NaOH stock solution. pH values are uncorrected for deuterium isotope effects. Standard ^1H NMR acquisition parameters for each ion titration point include a ^1H pulse width of 4.5 ms corresponding to a flip angle of 45°; a spectral window of 8 kHz, 16 K data points (corresponding to a digitial resolution of 0.429 Hz/pt), 120 scans, 2 dummy scans, and a recovery delay of 1 sec. Solvent suppression of the water signal was

achieved by use of a decoupler-generated presaturation pulse (60 dB decoupler attenuation) during the recovery delay. ^{31}P NMR acquisition parameters include a 45° flip angle of 8.2 μs, 5 kHz spectral window, 8 K data points (corresponding to a digital resolution of 0.610 Hz/pt), a recovery delay of 2 sec., and 1200 scans. Natural abundance proton-decoupled ^{13}C NMR experiments were carried out using the GARP composite-pulse decoupling sequence (Shaka *et al.*, 1985). ^{13}C NMR acquisition parameters include a ^{13}C 90° pulse of 17 μs, a decoupler 90° pulse of 125 μs at a decoupler power setting of 30 dB, a spectral window of 30 kHz, a recovery delay of 2.7 sec., 48,000 scans, and 32 K data points (corresponding to a digital resolution of 0.953 Hz/pt). Proton chemical shifts are referenced to external 2,2,2-trimethyl-2-silapentane sulfonic acid (TSS, 0.1 M in D_2O). ^{31}P chemical shifts are referenced to external ethylhydroxydiphos-phonate (EHDP, 0.5 M in D_2O) relative to 85% H_3PO_4. ^{13}C chemical shifts are referenced to external TMS. NMR processing parameters for each spectra are presented in the Figure legends. The term $\underline{\Delta\delta}$ is used in this report to denote upfield (positive) or downfield (negative) resonance frequency shift.

Homonuclear Hartmann-Hahn Correlation Spectroscopy. The assignment of ^1H NMR BDPP resonances during the course of the titration was achieved via the use of ^1H-^1H HOHAHA (Evans and Chan, 1992a; Bax and Davis, 1985a,b), and "clean" HOHAHA TPPI experiments (Griesinger *et al.*, 1988) where cross-relaxation is suppressed by the usage of two delays prior to and after the 180_y^o pulse in the MLEV-17 spin-lock sequence. Using the Bruker pulse sequence files MLEVPRTP and CLMLEVPRTP, the HOHAHA experiments utilized a spectral window of 6 kHz in both dimensions, 2 K data points (corresponding to a digital resolution of 2.94 Hz/pt in both dimensions), 240 scans/experiment, 512 experiments, 4 dummy scans, a 1 sec recovery period, a cross-relaxation suppression delay of 97.5 μsec, a t_1 value of 3 μs, a 90° excitation pulse of 10 μs, a SL_X trim pulse of 2.5 ms, MLEV 60°, 90°, and 180° decoupler pulses of 23 μs, 34 μs, and 68 μs, respectively, at a decoupler setting of 15 dB

attenuation. A spin-lock mixing time of 74 ms was used. FID transformation into the F1 and F2 frequency domain utilized 2 K data points in each dimension to create a 2 K x 2 K data matrix (4.883 Hz/pt in both dimensions), using a phase-sensitive squared sine-bell window function in both dimensions with no shift. Solvent suppression was obtained via a decoupler presaturation pulse (60 dB decoupler attenuation) during the recovery phase.

113-Cd NMR Solution Spectroscopy During the course of the titration, ^{113}Cd NMR (110.15 MHz) spectra were acquired at 300 K and 278 K on the Cd (II):BDPP sample. The lower temperature range was employed to slow any ^{113}Cd chemical exchange. Solution NMR acquisition parameters included a ^{113}Cd pulse width of 8 μs (45^{o} flip angle), a spectral window of 50 kHz, a recovery delay of 2 sec., and 64 K data points (corresponding to a digital resolution of 0.246 Hz/pt). ^{113}Cd chemical shifts in solution are referenced to external 2 M $Cd(ClO_4)_2$ in 90% H_2O/10%D_2O.

Graphic representation of data were processed in the KaleidaGraph software program, using the curve smoothing routine.

Results

Rationale for using ^{113}Cd (II) as a Probe for BDPP Conformational Change. The use of ^{113}Cd permits the observation of the metal ion environment, ligand composition and geometry, as well as the dynamics of ligand exchange [Kostelnik and Bothner-By, 1974; Holz *et al.*, 1976; Ackerman *et al.*, 1979; Summers, 1986, 1988; Jensen *et al.*, 1981]. Because the ionic radii of Cd (II) and Ca (II) differ by only 0.02 Å, and their coordination numbers are identical (L=6), ^{113}Cd NMR has proven to be a powerful tool in the study of metalloproteins, including Ca (II) binding proteins (e.g., calmodulin [Andersson *et al.*, 1983]; troponin C [Teleman *et al.*, 1983]; and on a number of Zn (II) containing proteins (e.g., alkaline phosphatase [Gettins and Coleman, 1984a,b], carboxypeptidase A [Gettins, 1986; Ellis, 1989], yeast

inorganic pyrophoshatase (Welsh and Cooperman, 1984), and HIV-1 nucleic acid binding protein (South *et al.*, 1989)). One extracellular mineral matrix Ca (II) binding protein which has been investigated by ^{113}Cd NMR is osteocalcin, or bone-GLA protein (Prigodich *et al.*, 1985). The Cd (II) bound form of the bone GLA-protein generates a ^{1}H NMR spectrum that is nearly identical to the Ca (II) form of the protein (Prigodich *et al.*, 1985), save for increased broadening of specific sidechain proton resonances induced by Cd (II) exchange that is more rapid than that of Ca (II). This difference in ion dissociation rates was also observed in equilibrium dialysis experiments conducted on the 30 kD rat phosphophoryn fragment by Zanetti *et al.*, (1981). Here, different dissociation constants ([Ca] K_d = 7.1 x 10^{-7} M; [Cd] = 2.5 x 10^{-5} M) were obtained for Ca (II) and Cd (II) in the presence of rat phosphophoryn. Therefore, care should be exercised in extrapolating phosphophoryn Cd (II) binding data to the Ca (II) situation. Our rationale for employing ^{113}Cd (II) in place of Ca (II) in this study was (1) to examine <u>divalent</u> cation-induced folding transitions in the phosphophoryn molecule, (2) utilize the ^{113}Cd nucleus to induce ^{31}P chemical shift effects (Gettins, 1988; Rhyu *et al.*, 1985; Gettins and Coleman, 1984a,b), and (3) assess the effects of BDPP binding on the exchange of the ^{113}Cd nucleus. An additional benefit derived from the use of ^{113}Cd (II) is the ability to determine the level of divalent cation saturation of BDPP binding sites via the appearance of the "unbound" ^{113}Cd (II) species (hydrated chloride form) resonating at 125 ppm from $Cd(ClO4)_2$.

Several precautions were taken during the course of the Cd (II) titration experiment. First, the Cd (II) titration was carried out at 7.20 to take advantage of the buffering capacity of BDPP PSer groups (Evans and Chan, 1992), thus minimizing the amount of NaOH needed to restore the pH after the addition of acidic $CdCl_2$ solution. By reducing the amount of NaOH added, sodium-induced ionic strength effects on BDPP were likewise reduced. Second, and most importantly, emphasis was placed on the low stoichiometry regime of the titration (^{113}Cd:protein $\leq$ 50:1 molar ratio). Phosphophoryn protein

precipitation and/or protein-protein condensation in the presence of Ca^{2+} has been witnessed at or above the level of 65% saturation (Marsh, 1989a,b). In our titration experiments, no changes in solution turbidity (monitored at $OD_{380\ nm}$) were observed at or below 100:1 Cd (II):protein stoichiometry. Theoretically, based on the known BDPP amino acid composition (Stetler-Stevenson and Veis, 1983; Rahima and Veis, 1988), a level of 300:1 or greater would be required to neutralize all of the negative charges on BDPP at pH 7.20, assuming that there is full exchange of Na^+ by Cd (II) on the protein surface. By limiting our study to the stoichiometry regime of 50:1 Cd (II):protein or less, precipitation problems are largely negated, and, conformational effects due to protein-protein interactions are minimized. As a sidenote, above the level of 250:1, we noted that the BDPP protein sample begins to exhibit slight turbidity, which transforms into an amorphous precipitate at 300:1 or above (data not shown).

Evidence for ^{113}Cd Chemical Exchange and Rapid Relaxation in the Presence of BDPP. In performing ^{113}Cd NMR experiments on BDPP at 25° C, and at 5° C, we were unable to detect any ^{113}Cd species in the presence of BDPP at any stoichiometry below 400:1. Above 400:1 stoichiometry, the labile CdX_n ($X = Cl^-$) species was observed at 125 ppm at room temperature, indicating that unbound Cd (II) ion was present in the sample. The low temperature experiments were employed to slow down the exchange of the ^{113}Cd spins between chemically distinct sites. We did not attempt above ambient temperature ^{113}Cd NMR experiments on BDPP out of a concern for thermally-induced conformational change effects arising from BDPP unfolding (Evans and Chan, 1992c). Linse *et al.*, in their ^{113}Cd study of polystyrene sulfonate polyelectrolytes, demonstrated that rapid exchange conditions occur in the presence of sulfonate and sulfate anionic groups (Linse *et al.*, 1981), and that if specific coordination or chelation effects are present, the exchange rate might shift to the intermediate or even the slow regime. Given that the K_d for Cd (II) in the presence of phosphophoryn is two orders of magnitude different from that of Ca (II), the affinity of Cd (II) for PSer and Asp sidechains is

significant; however, slow exchange conditions are unlikely. In addition, the Cd (II) species during its lifetime in the bound and free form is most likely undergoing multiple eqilibria with Cl^- and H_2O in aqueous solution. This can lead to profound changes in the shielding of the ^{113}Cd nucleus (up to 300 ppm)(Ackerman *et al.*, 1979). Therefore, the absence of any detectable ^{113}Cd resonance in the presence of BDPP up to the level of ion saturation may be due to intermediate chemical exchange between two or more sites, and/or, due to ^{113}Cd relaxation arising from high affinity ion binding at the surface of a large macromolecule.

BDPP Template Protein Folding in the Presence of Cd (II): Involvment of PCBD Regions. In a previous study, it was determined that the frequency range of the ^1H NMR BDPP spectrum from 2.40 ppm to 5.50 ppm represents the fingerprint region for sidechain protons in the HD regions (Peaks H, G, I, J, K, L, M, and 17-21) and the PCBD regions (Peaks A-F, 12-16, and Peak N) of the BDPP protein (Evans and Chan, 1992a,b). Similarly, the amide NH fingerprint region (7.50 ppm to 9.00 ppm) of BDPP has also been mapped, with assignments made to PSer and Asp residues in $(Asp)_n$, $(PSer)_n$, and $(PSer\text{-}Asp)_n$ polyelectrolyte sequence regions (Evans and Chan, 1992c). Therefore, proton backbone resonances (amide-NH, Gly α,α'-CH, Ser α-CH) and proton and phosphorus sidechain resonances that are several bonds removed from the peptide backbone (^1H NMR: Ser β,β'-CH_2; Asp β,β-CH_2; Pro δ-CH_2; ^{31}P NMR: PSer PO_4) provide sensitive indicators for BDPP secondary and tertiary conformational changes, respectively, in the presence of Cd (II).

An examination of the ^1H NMR NH amide fingerprint region of BDPP reveals that all of the PSer and Asp amide resonances experience small but discernable downfield chemical shift effects ($-\Delta\delta$ ranging from 0.01 to 0.05 ppm) between 0 and 1:1 Cd:protein stoichiometry (Figure 1A). Amide resonance frequency shifts are typically the result of backbone dihedral angle conformational change. During this same interval in the titration, the relative intensities of all amide resonances

decreases (Figure 1B), signifying changes in amide proton relaxation. Thus, the 0 to 1:1 titration interval represents the first conformational backbone transition in BDPP. A second conformational transition occurs between 5:1 and 10:1 for $(Asp)_n$ homopolymer and Asp residues residing in $(PSerAsp)_n$ heterpolymer sequence regions in the protein (Figure 1A, 1B, Peaks Δ, Ψ, Φ, and Ω). This folding event is evident in terms of relative amide intensities, but does not appear as a strong transition in terms of proton chemical shift effect. The third backbone conformational transition occurs in the10:1 to 50:1 Cd (II):protein interval of the titration, where all PSer and Asp amide resonances experience broadening effects, in some cases resulting in the disappearance of specific amide resonances (Figure 1A, B, note zero intensities). Collectively, the BDPP amide chemical shift and broadening effects in the presence of Cd (II) are a clear indication of stoichiometry-dependent Cd (II) interactions at specific PCBD cation binding regions. Cd (II) interactions at the anionic sidechains induce perturbations in the Asp and PSer backbone torsion angles, and lead to a reduction in the local backbone motion and hence in the transverse relaxtion times (T_2) of PSer and Asp amide protons in the BDPP protein.

BDPP Template Protein Folding in the Presence of Cd (II): Evidence for the Participation of Hinge Regions in the Folding Process. As the BDPP protein transitions from the Na^+ counterion form at low ionic strength to the fully protonated state, the molecule undergoes structural condensation (Evans and Chan, 1992). This process is mediated by HD regions of the BDPP protein that contain neutral amino acids (Pro, Gly, Ser). From our pH titration studies, we have categorized the hinge sidechain spin populations according to the type of polyelectrolyte sequence regions which flank them (i.e., $PSer_n$, $(PSerAsp)_n$, and Asp_n)(Evans and Chan, 1992b). Not surprisingly, as the BDPP molecule undegoes titration with Cd (II), the HD regions exhibit a response to Cd (II) interactions at flanking PCBD regions (Figures 2, 3A,B). As in the case of the amide resonances, HD conformational transitions are reflected in changes in relative resonance intensities, with major inflection points occurring in the 0 to 25:1 stoichiometry

range (Figure 3A). There also seems to be evidence for a minor conformational transition occurring between 1:10 and 1:1 stoichiometry as well (Figure 4A, 4B). The most significant change in relative intensities are observed for Peaks J and L (Gly α, α', Ser β, β', α protons) of HD regions which are flanked by $(PSer)_n$ and $(PSerAsp)_n$ PCBD regions (Evans and Chan, 1992b). Peaks J and G (Ser β,β', α/Gly α, α'), and K (Pro δ-CH_2), which represent HD regions of BDPP that are flanked by $(PSer)_n$ and $(PSerAsp)_n$-containing PCBD regions, are still undergoing conformational transitions above the 25:1 stoichiometry level. Peak A, which represents Ser β, β' CH_2 and Gly α, α' protons flanked by $(Asp)_n$ PCBD regions (Figure 4B), exhibits a major conformational transition centered near 10:1 stoichiometry.

During the course of the Cd (II) titration, aliphatic sidechain proton resonances (denoted as a-g, Figure 2) which were previously undetected at the zero stoichiometry level were identified in the ^1H NMR BDPP spectra. A "clean" 2-D HOHAHA experiment reveals that several of these resonances exhibit scalar connectivities with resonances at 1.275 and 1.130 ppm (Figure 5). In this same HOHAHA spectrum two other previously unidentified *J*-connectivities have been observed, arising from proton resonances at 1.210 and 4.253 ppm, and 1.393 and 4.300 ppm. These resonances may arise from the methyl protons of the Ala β→α AMX spin system. From the *J*-connectivity patterns and ^1H NMR chemical shifts (Basus, 1989; Jardetzky and Roberts, 1981), we have tentatively assigned resonances a-g to the δ, δ' protons comprising the $A_2(T_2)MPX$ spin system of the Pro pyrrolidine ring. Coincidentally, Pro δ, δ' diastereotopic ring proton resonances have been identified in this same 3.500 ppm frequency region during the course of the pH-induced folding transitions in BDPP (Evans and Chan, 1992a,b). The resonances at 1.130 ppm also demonstrate relayed scalar couplings with α-CH protons at 5.161 ppm. The absence of other Pro ring proton resonances in the BDPP spectrum at this stoichiometry may be explained by short T_2 values for the Pro protons which experience restricted motions on the peptide backbone. This would suppress spin-lock magnetization transfer for these protons

(Torchia *et al.*, 1989; Bax and Davis, 1985a,b). As shown in Figure 4B, the δ, δ' Pro resonances exponentially increase in intensity, synchronous with the decrease in relative intensities of the β, β'-CH_2 diastereotopic protons (Peaks 12-16) of Asp sidechains. It is likely that these Pro residues exist in HD regions that are flanked by Asp-containing PCBD regions that are interacting with Cd (II) ions.

The synchronicity of these HD conformational transitions with Cd (II) binding at flanking PCBD regions is strong evidence for the participation of these HD regions in the global folding of the BDPP molecule. The degree of conformational change is cation-dependent, i.e., many of the transitions have reached their endpoint at 25:1 stoichiometry (e.g., Peak A, L, M), while other conformational transitions, represented by Peaks a-g, 17-20, G, and J, are still in progress.

Evidence for Cd (II) Binding Site Preference on the BDPP molecule. In addition to conformational folding effects, we were interested in observing the Cd (II) loading of BDPP to ascertain if the divalent cation exhibits site preference for specific polyelectrolyte sequence regions. Preferetial site binding may play an important role not only in determining the sequence of folding events, but also in defining the remaining sites available for cation deposition in the nucleation process. Returning our attention to Figure 1A, the Cd (II) induced conformational folding effects occur more rapidly for PSer amide resonances P (null point attained) and Π (> 50% decrease in relative intensity) than for Asp amide resonances at stoichiometries less than or equal to 10:1. This suggests that Cd (II) at low stoichiometry binds to PSer sequence-containing PCBD regions preferentially over $(Asp)_n$ sequence-containing regions in BDPP. This is not surprising in light of the observation that high affinity binding sites on phosphophoryn are abolished after alkaline phosphatase dephosphorylation (Zanetti *et al.*, 1981). Further confirmation of Cd(II) preference for PSer binding sites is observed in Figure 5 and 6, where the ^{31}P resonances α, β, and χ feature downfield chemical shift effects ($-\Delta\delta = 0.25$ ppm) in the 0 to

10:1 stoichiometry range, which coincides with the first conformational backbone transition (see above). A larger downfield shift effect is observed above 10:1 ($-\Delta\delta$ = 1.2 to 1.5 ppm), which coincides with the second and third conformational transitions in the PCBD cation binding regions. This downfield ^{31}P NMR chemical shift effect is the result of ^{113}Cd interactions with the monophosphate ester groups; similar effects have been reported for other biological phosphoryl-cadmium coordination complexes (Gettins, 1988; Rhyu *et al.*, 1985). Within the total population of ^{31}P PSer spins, there is evidence for PSer binding site discrimination by Cd (II), as evidenced by the selective disappearance of Peak χ first from the ^{31}P NMR spectra above 10:1 stoichiometry (Figure 6, 7). From our previous work, we have assigned Peak χ to represent $(PSer\text{-}PSer)_n$ clusters whose high negative charge density results in pertubation of the monophosphate ester O-P-O bond angles and consequently a slight downfield ^{31}P resonance shift effect (Evans and Chan, 1992a). Peak β, which has been assigned to ^{31}P spins residing in $(PSerAsp)_n$ heteropolymer clusters, disappears later on in the course of the titration (> 10:1)(Figure 6,7). Interestingly, a small population of PSer ^{31}P spins (5%) are apparently insensitive to Cd (II) ion addition (Peak ζ, Figure 6), as evidenced by the absence of a downfield chemical shift effect for these residues at or above 25:1. It is presumed that this spin population does not interact with Cd (II) at this level of stoichiometry.

Amide resonance intensities for Asp NH protons residing in $(PSerAsp)_n$ polyelectrolyte clusters (Peaks Φ, Ψ, and Ω) also decrease in parallel with the PSer amide intensities, suggesting that Cd (II) exhibits a preference for the neighboring Asp residues in these heteropolymer anionic sequence stretches over $(Asp)_n$ homopolymer sequences. A more direct way of assessing interactions of Cd (II) with Asp carboxylate groups is through the use of ^{13}C NMR, where either chemical shift effects or spectral broadening at Asp C_γ resonances can be observed. In Figure 8, the proton-decoupled natural abundance ^{13}C NMR spectra of BDPP at zero Cd (II) and 50:1 Cd (II) in H_2O/D_2O, pH 7.21, is presented. At this pH, BDPP possesses six distinct Asp C_γ resonances,

denoted 1-6, with centerpeaks at 183.343, 183.244, 180.969, 176.593, 174.720, and 174.311 ppm, respectively. The $\Delta\nu_{1/2}$ values for these peaks range from 8 to 15 Hz, and Peaks 3-6 demonstrate shoulder regions. Although we are unable to assign the C_γ resonances to any specific Asp sidechains, it is important to note that at 50:1 stoichiometry Peaks 1 and 2 are no longer observed in the ^{13}C NMR spectra, while Peaks 3-6 exhibit spectral broadening ($\Delta\nu_{1/2}$ =100 -120 Hz) and a 35-50% reduction in peak area. The loss of ^{13}C intensities can be attributed to Cd (II) interactions at specific Asp carboxylate residues. These effects are less pronounced for other Asp carboxylate groups at 50:1, presumably due to lower affinity of Cd (II) at these sites. Thus, at low levels of ion saturation, the interaction of Cd (II) with Asp residues is unequal in affinity; certain carboxylate sites become occupied prior to other sites. We can summarize the information regarding the preference of Cd (II) for sites on the BDPP template as follows:

$$\text{Affinity:} \quad (\text{PSer-PSer})_n > (\text{PSerAsp})_n > (\text{Asp})_n$$

The preference of PSer residues as binding sites for divalent cations was also noted by Cookson *et al.* (1980), in which Mn (II) induced resonance broadening was observed primarily at PSer β-CH_2 resonances at low Mn:protein stoichiometry (< 15:1).

Discussion

This present investigation into the sequence of divalent cation induced folding events in BDPP strongly supports our original hypothesis regarding pH-induced folding: namely, that BDPP utilizies specific sequence regions or domains to organize the transition from the extended global conformational state to a more compact conformation that presumably acts as a template surface for nucleation of calcium phosphates. As shown in the pH-induced folding experiments (Evans and Chan, 1992a,b), and in our current divalent cation study, the two types of sequence regions -- PCBD and HD --

synergistically interact in the folding process. In the presence of divalent cations such as Cd (II), or, in the presence of H^+ which can more effectively shield sidechain charges, the BDPP molecule no longer experiences electrostatic repulsion forces in certain sequence regions. This would lead to a folding of the protein structure in these region. As more ion binding sites become progressively occupied, this negation of repulsive forces and structural condensation would propagate throughout the protein molecule according to the ion binding site preference.

Evidence for the Cation "Hopping" Hypothesis and Its Role in PMMP Protein Folding. A key element in the folding process is the ability to allow partially shielded PCBD regions to come closer together, since this would create a higher density of oxoanion atoms whose oxygen-oxygen distances would approximate those found in calcium hydroxylapatite. This would then result in the generation of a molecular complementary "face" of a forming crystal (Mann, 1988). One of the ways in which this higher negative charge density might be established over the BDPP molecule is effective charge neutralization of anionic sidechains via the "hopping" of divalent cations from site to site on the protein. Evidence for this type of "hopping" phenomenon in phosphophoryn was reported earlier by Cookson *et al.*, (1980), where Mn (II) exchange was noted via paramagnetic perturbation of the ^{31}P and ^{1}H spins of PSer at or below 15:1 Mn (II):phosphophoryn levels, and both the Asp and PSer β-CH_2 ^{1}H spins at higher levels of Mn (II). The authors claimed that the Mn (II) protein-bound species possessed a short lifetime (i.e., fast exchange) but a high affinity for the sites (i.e., $K_a = 10^6\ M^{-1}$). Thus, they perceived the Mn (II) species as continually "hopping" from one phosphate ester to the next while maintaining ligands to one or two other carboxylate and/or monophosphate ester ligands. In the present study, our observations suggest that the exchange rate of Cd (II) in the presence of BDPP is apparently in the intermediate regime. However, the affinity of Cd (II) for phosphophoryn is 100-fold lower than that of Mn (II), and, since the observed exchange rate may be influenced by the presence of multiple equilibria of the Cd (II) ion with halide or aquo

ligands (Ackerman *et al.*, 1979), it is possible that the Cd (II) ion may actually be exchanging rapidly between sites on the protein molecule. That being the case, Cd (II) "hopping" may very well be taking place between sites on the BDPP protein at low ion:protein stoichiometry. Further evidence for this "hopping" phenomenon is found in the following observations: (1) *^{113}Cd (II) induced perturbation of ^{31}P chemical shifts at low ion:protein stoichiometry levels* (Figure 5,6). Since the BDPP protein contains 512 PSer residues per molecule, the PSer sites on BDPP at 50:1 stoichiometry or lower are undersaturated by a factor of 10 or more. Thus, if 95% of the total ^{31}P spins are perturbed at 50:1, the Cd (II) must be exchanging on and off the monophosphate ester sites. (2) *Presence of Cd (II)-Asp COO interactions at low ion:protein stoichiometry levels* (Figure 1A, 2, 7). The ^{1}H NMR and ^{13}C NMR spectra of BDPP hint that the Cd (II) exchange process is occurring at Asp carboxylate sites as well (Figures 2,7). Thus, although divalent cations exhibit a preference for monophosphate ester sites (Zanetti *et al.*, 1981; see results, above), the dissociation constant (K_d) is large enough to permit some diffusion of the cation to other sites. It is not known to what extent the Cd (II) ion exchanges out into the bulk solution in the presence of BDPP. However, since it has been established that polyelectrolytes condense counterions at their surface (Manning, 1977, 1978), and that extra amounts of counterions will condense or bind on the polyion at almost all bulk ionic concentrations (Hao and Harvey, 1992; Anderson and Record, 1982, 1990), it is likely that counterions will cluster on BDPP at the low ion:protein stoichiometry range, and, that the cations will preferentially exchange at sites on the surface of the protein, rather than exchange out into the bulk solution. Therefore, an effective mechanism for folding the BDPP protein into a template surface may be the ability of divalent cations to exchange between anionic sites in a relatively rapid fashion, such that effective charge screening is achieved over a large portion of the protein molecule.

A Model for Divalent Cation Induced Protein Folding in PMMPs. Based on our observations in this report, we wish to refine our description of

BDPP global conformational change as was originally presented for a pH-induced mechanism. Figure 8 represents a conceptualization of the divalent cation-induced template folding process on a global scale. The first step of the folding process involves divalent cation interactions at high affinity sites on the BDPP molecule (Transition A → B). These are the $(PSer\text{-}PSer)_n$ repeat clusters which exhibit preferential binding to Cd (II) at low stoichiometries (Figure 1A, 5, 6). As the binding sites on these sequences become saturated, perhaps due to rapid cation exchange or "hopping," partial charge shielding occurs in these regions, and the BDPP molecule undergoes limited conformational folding. This folding may involve the participation of HD regions which are flanked by $(PSer)_n$ domains (indicated in Figure 8 by greyscale shading complementary to that of the PCBD). How these large conformational change movements are coordinated is not understood at this time. Conformational change in many proteins has been shown to involve hinge regions that are loop structures (White *et al.*, 1976); examples of hinge loop movements also include domain movements in immunoglobulins (Bennet and Huber, 1984; Lesk and Chothia, 1988) and lysozyme (Faber and Matthews, 1990). Alternatively, there have been reports of rigid-body motions of helices, such as those that occur in insulin (Chothia *et al.*, 1983). Following this initial folding step at $(PSer)_n$ domains, Cd (II) ions occupy binding sites in the $(PSerAsp)_n$ PCBD regions (Transition B → C, Figure 8), thereby initiating folding at these regions in BDPP by the same charge shielding, hinge-mediated mechanism. The binding sites in the $(PSer\text{-}Asp)_n$ repeat regions may involve both the PSer and Asp carboxylate groups at this stage. At or about this stage, PSer residues which are interdispersed in the sequence (represented by Peak α, Figure 5) and are not part of either the $(PSer)_n$ or $(PSer\text{-}Asp)_n$ repeat clusters, will also interact with Cd (II). Lastly, $(Asp)_n$ repeat sequence regions, which are apparently the low affinity sites on phosphophoryn (Zanetti *et al.*,1981) act as binding sites for Cd (II)(Transition C → D, Figure 8). The resultant structure has now collapsed to a smaller molecular volume that still possesses a sufficient number of unoccupied anion sites for cation binding, but now at a higher density (i.e., oxoanion ligands separated by 5 Å or less). The

protein could now act as an effective template surface for nucleation and formation of the metal atom array of a mineral lattice.

Significance of ^{31}P Spins Unaffected by Cd (II). As shown in our ^{31}P NMR data, there does exist PSer residues (5%) which are unaffected by ^{113}Cd (II) interactions up to 50:1 stoichiometry. One reason for this may be due to the tertiary structure of the region they are localized in. Marsh (1989a,b) demonstrated that in the presence of Ca (II), phosphophoryn creates protective "domains" where bound inorganic phosphate anions cannot readily dissociate from the protein-Ca (II) complex. It is possible that phosphophoryn posseses sequence motifs where a certain percentage of PSer residues are shielded from the solvent. Support for this theory has been demonstrated in our acetonitrile mixing experiment (Evans and Chan, 1992c), where two ^{31}P resonances representing ca. 5-7% of the total monophosphate ester species did not experience a solvent-induced chemical shift. One possible candidate for such a shielded motif would be the hydrophobic cluster (HC) region of BDPP, which in our earlier NMR studies exhibited insensitivity to pH induced folding effects (Evans and Chan, 1992b), and likewise in the present study demonstrated little if any perturbation by Cd (II). Further studies clarifying the role of the hydrophobic domain in the nucleation and folding process, as well as the kinetics of ion binding to BDPP, are currently underway.

Acknowledgments

We thank Mr. Gordon R. Bradford of the Citrus Research Center, University of California, Riverside, for performing the ICP analysis. JSE acknowledges a National Research Service Award from the National Institute of Dental Research (DE-05445-05), and a fellowship award from AMGEN Pharmaceuticals.

References

Ackerman, J.J.H, Orr, T.V., Bartuska, V.J., and Maciel, G.E. (1979) *J. Am. Chem. Soc.* 101 (2) 341-347.

Addadi, L., and Weiner, S. (1985) *Proc. Natl. Acad. Sci USA* 82 4110-4114.

Addadi, L., and Weiner, S. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York pp. 133-152.

Addadi, L., Moradian, J., Shay, E., Maroudas, N.G., and Weiner, S. (1987) *Proc. Natl. Acad. Sci USA* **84** 2732-2736.

Addadi, L., Moradian-Oldak, J., and Weiner, S. (1991) in *ACS Symposia: Surface Reactive Peptides and Proteins.* 13-27.

Addadi, L., Moradian-Oldak, J., Furedi-Milhofer, H., and Weiner, S. (1992) *Connect. Tiss. Res.* 27 (2-3) 93.

Anderson, C.F., and Record, M.T. (1982) *Annu. Rev. Phys. Chem.* 33 191-245.

Anderson, C.F., and Record, M.T. (1990) *Annu. Rev. Biophys. Biophys. Chem.* 19 423-445.

Andersson, A., Drakenberg, T., Thulin, E., and Forsen, S. (1983) *Eur. J. Biochem.* 134 459-465.

Basus, V. J. (1989) *Methods in Enzymology* 177 132-141.

Bax, A., and Davis, D. G. (1985a) *J. Am. Chem. Soc.* 107 2820-2821.

Bax, A., and Davis, D. G. (1985b) *J. Mag. Resonance* 65 355-360.

Bennet, W.S., and Huber, R. (1984) *Crit., Rev. Biochem.* 15 291-384.

Borbas, J.E., Wheeler, A.P., and Sikes, C.S. (1991) *Journ. Exptl. Zool.* 258 1-13.

Chothia, C., Lesk, A.M., Dodson, C.G., and Hodgkin, D.C. (1983) *Nature (London)* 302 699-702.

Constantz, B., and Weiner, S. (1986) *Journ. Exptl. Zool.* 248 253-258.

Cookson, D.J., Levine, B.A., Williams, R.J.P., Jontell, M., Linde, A., and de Bernard, B. (1980) *Eur. J. Biochemistry* 110 273-278.

Dill, K.A. (1990) *Biochemistry* 29 (31) 7133-7155.

Ellis, P.D. (1989) *J. Biol. Chem.* 264 (6) 3108-3110.

Evans, J.S., and Chan, S.I. (1992a) *Manuscript in preparation.*

Evans, J.S., and Chan, S.I. (1992b) *Manuscript in preparation.*

Evans, J.S., and Chan, S.I. (1992c) *Manuscript in preparation.*

Faber, H.R., and Matthews, B.W. (1990) *Nature (London)* 438 263-266.

Freire, E. (1990) *Commun. Mol. Cell Biophys.* 6 123-140.

Freire, E., and Murphy, K.P. (1991) *J. Mol. Biol.* 222 687-699.

George, A., Simonian, P., Tylzanowski, P., and Veis, A. (1992) *Connect. Tiss. Res.* 27 (2-3) 107.

Gettins, P. (1986) *J. Biol. Chem.* 261 (33) 15513-15518.

Gettins, P. (1988) *J. Biol. Chem.* 263 (21) 10208-10211.

Gettins, P., and Coleman, J.E. (1984a) *J. Biol. Chem.* 259 (8) 4987-4990

Gettins, P., and Coleman, J.E. (1984b) *J. Biol. Chem.* 259 (8) 4991-4997.

Gotoh, Y., Gernstenfeld, L.C., and Glimcher, M.J. (1990) *Eur. J. Biochem.* 187 49-58.

Griesinger, C., Otting, G., Wuthrich, K., and Ernst, R.R. (1988) *J. Am. Chem. Soc.* 110 7870-7872.

Grob, K-H., and Kalbitzer, H. R. (1988) *J. Mag. Resonance* 76 87-99.

Hao, M-H, and Harvey, S.C. (1992) *Macromolecules* 25 2200-2208.

Holz, M., Jordan, R.B., and Zeidler, M.D. (1976) *J. Mag. Res.* 22 47-52.

Jardetzky, O., and Roberts, G. C. K. (1981) *NMR in Molecular Biology* (eds. by Jardetzky and Roberts) Academic Press, New York, New York, pp. 143-186, 227-231, 234-270, 448-465.

Jensen, C.F., Deshmukh, S., Jakobsen, H.J., Inners, R.R., and Ellis, P.D. (1981) *J. Am. Chem. Soc.* 103 3659-3666.

Kostelnik, R.J., and Bothner-By, A.A. (1974) *J. Mag. Res.* 14 141-151.

Kuwajima, K. (1989) *Proteins: Structure, Function, Genetics* 6 87-103.

Lee, S.L., and Veis, A. (1980) *Int. J. Pept. Prot. Res.* **16** 231-240.

Lee, S.L., Veis, A., and Glonek, T. (1977) *Biochemistry* **16** (13) 2971-2979.

Lesk, A.M., and Chothia, C. (1988) *Nature.* 335 188-190.

Linse, P., Gustavsson, H., Lindman, B., and Drakenberg, T. (1981) *J. Mag. Res.* 45 133-141.

Lowenstam, H.A., and Weiner, S. (1989) in *On Biomineralization* Oxford University Press, New York, New York, pp.7-40.

Mann, S. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., eds) VCH Publishers, New York, New York,

Manning, G.S. (1977) *Quart. Rev. Biophys.* 11 179-210.

Manning, G.S. (1978) *Biophys. Chem.* 7 95-110.

Marsh, M.E. (1989a) *Biochemistry* 28 339-345.

Marsh, M.E. (1989b) *Biochemistry* 28 346-352.

Marsh, M.E., and Sass, R.L. (1984) *Biochemistry* 23 1448-1456.

Nancollas, G.H. (1988) in *Biomineralization: Chemical and Biochemical Perspectives* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 157-182.

Prigodich, R.V., O'Connor, T., and Coleman, J.E. (1985) *Biochemistry* 24 6291-6298.

Privalov, P.L. (1979) *Advances in Protein Chemistry* 33 167-241.

Privalov, P.L. (1982) *Advances in Protein Chemistry* 35 1-104.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* 42 104-112.

Ramsay, G., and Freire, E. (1990) *Biochemistry* 29 8677-8683.

Rhyu, G.I., Ray, W.J., and Markley, J.L. (1985) *Biochemistry* 24 2536-2541.

Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) *Biochem. J.* 276 699-707.

Shaka, A.J., Barker, P.P., and Freeman, R. (1985) *J. Mag. Res.* 64 547-552.

Shiraga, H., Min, W., VanDusen, W.J., Clayman, M.D., Miner, D., Terrell, C.H., Sherbotie, J.R., Foreman, J.W., Przysiecki, C., Neilson, E.G., and Hoyer, J.R. (1992) *Proc. Natl. Acad. Sci USA* 89 426-430.

South, T.L., Kim, B., and Summers, M.F. (1989) *J. Am. Chem. Soc.* 111 395-396.

Stetler-Stevenson, W.G., and Veis, A. (1983) *Biochemistry* 22 4326-4335.

Stetler-Stevenson, W.G., and Veis, A. (1987) *Calcif. Tiss. Int.* 40 97-102.

Summers, M.A. (1986) *J. Am. Chem. Soc.* 108 4254-4258.

Summers, M.F. (1988) *Coordination Chem Rev.* 86 43-134.

Swift, D.M., Sikes, C.S., and Wheeler, A.P. (1986) *Journ. Exptl. Zool.* 240 65-73.

Teleman, I., Drakenberg, T., Forsen, S., and Thulin, E. (1983) *Eur. J. Biochem.* 134 453-457.

Torchia, D. A., Sparks, S. W., Young, P. E., and Bax, A. (1989) *J. Am. Chem. Soc.* 111 8315-8317.

Veis, D.J., Albinger, T.M., Clohisy, J., Rahima, M., Sabsay, B., and Veis, A. (1986)*J. Expt. Zool.* 240 35-46.

Weiner, S. (1983) *Biochemistry* 22 4139-4145.

Weiner, S., and Hood, L. (1975) *Science* 190 987-989.

Weiner, S., and Traub, W. (1984) *Phil. Trans. R. Soc. London Ser. B* 304 421-438.

Welsh, K.M., and Cooperman, B.S. (1984) *Biochemistry* 23 4947-4955.

White, J., Hackert, M.L., Buehner, M., Adams, M.J., Ford, G.C., Lentz, P.J., Smiley, I.E., Steindel, S.J., and Rossmann, M.G. (1976) *J. Mol. Biol.* 102 759-779.

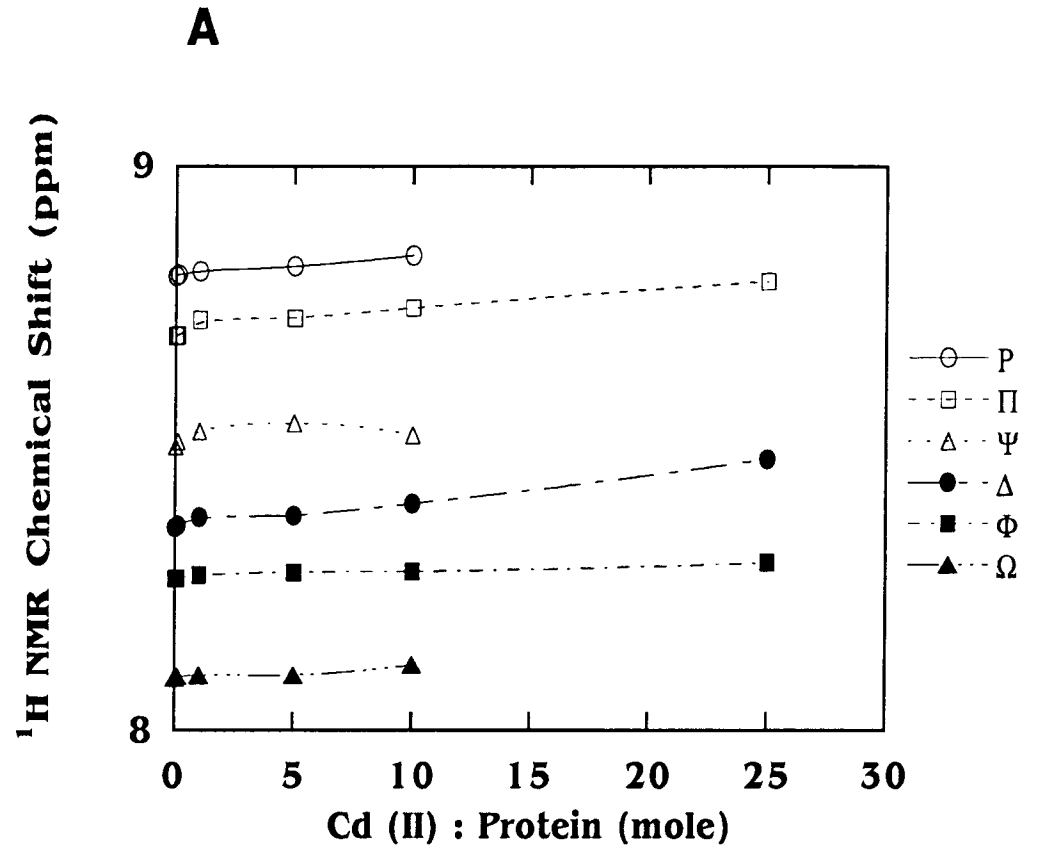
Worms, D., and Weiner, S. (1986) *Journ. Exptl. Zool.* 237 11-20.

Zanetti, M., de Bernard, B., Jontell, M. and Linde, A. (1981) *Eur. J. Biochem.* 113 541-545.

Zhang, Q., Domenicucci, C., Goldberg, H.A., Wrana, J.L., and Sodek, J. (1990) *J. Biol. Chem.* 265 (13) 7583-7589.

Figure 1

(A): The dependence of BDPP PSer and Asp amide proton chemical shifts on Cd (II) addition. **(B)**: Relative BDPP PSer and Asp amide proton intensities as a function of Cd (II) addition. The *y*-axis is given in arbitrary units. Curve fitting was performed using a smoothing function in the KALEIDAGRAPH software program.


A
^{1}H NMR Chemical Shift (ppm)
9
8
0
5
10
15
20
25
30
Cd (II) : Protein (mole)
P
Π
Ψ
Δ
Φ
Ω

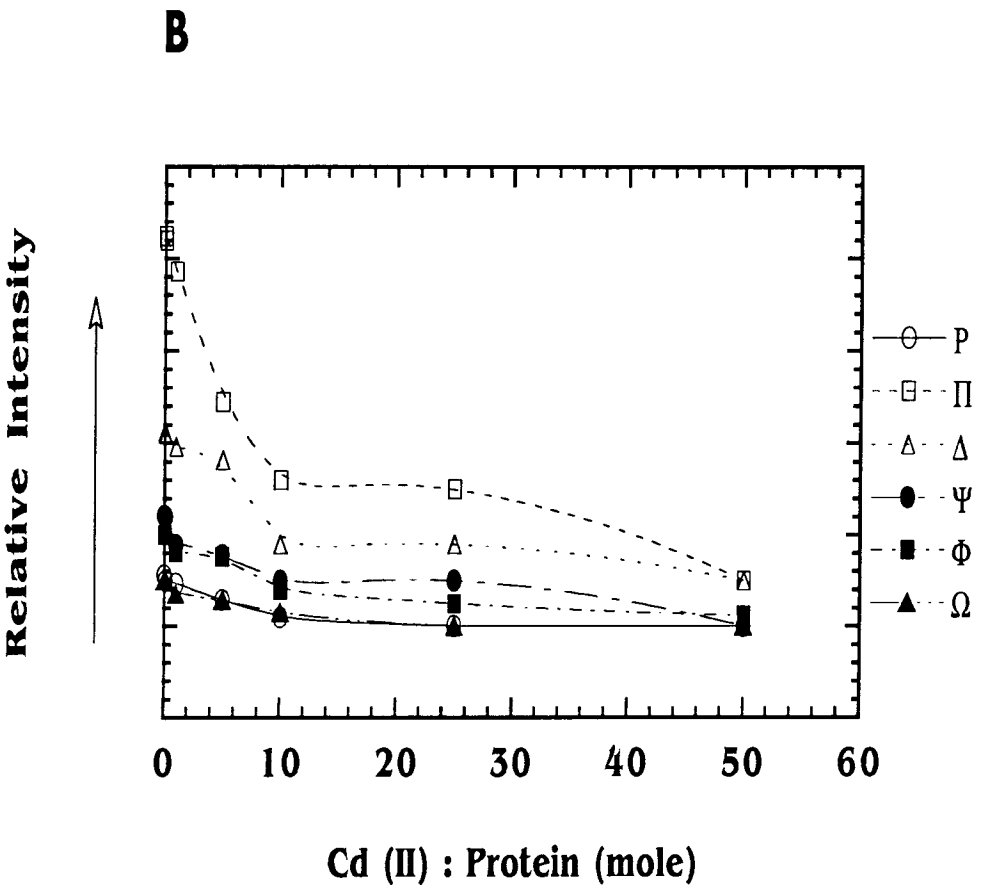

B
Relative Intensity
0
10
20
30
40
50
60
Cd (II) : Protein (mole)
P
Π
Δ
Ψ
Φ
Ω

Figure 2: *500.138 MHz ^{1}H NMR spectra of BDPP as a function of ^{113}Cd (II) addition.*

The non-exchangeable sidechain proton fingerprint region for PSer, Asp, and neutral amino acid sidechains is shown. All spectra are plotted at absolute intensities relative to Cd:Protein = 1:10. Proton chemical shifts are referenced from 0.1 M TSS in D_2O.

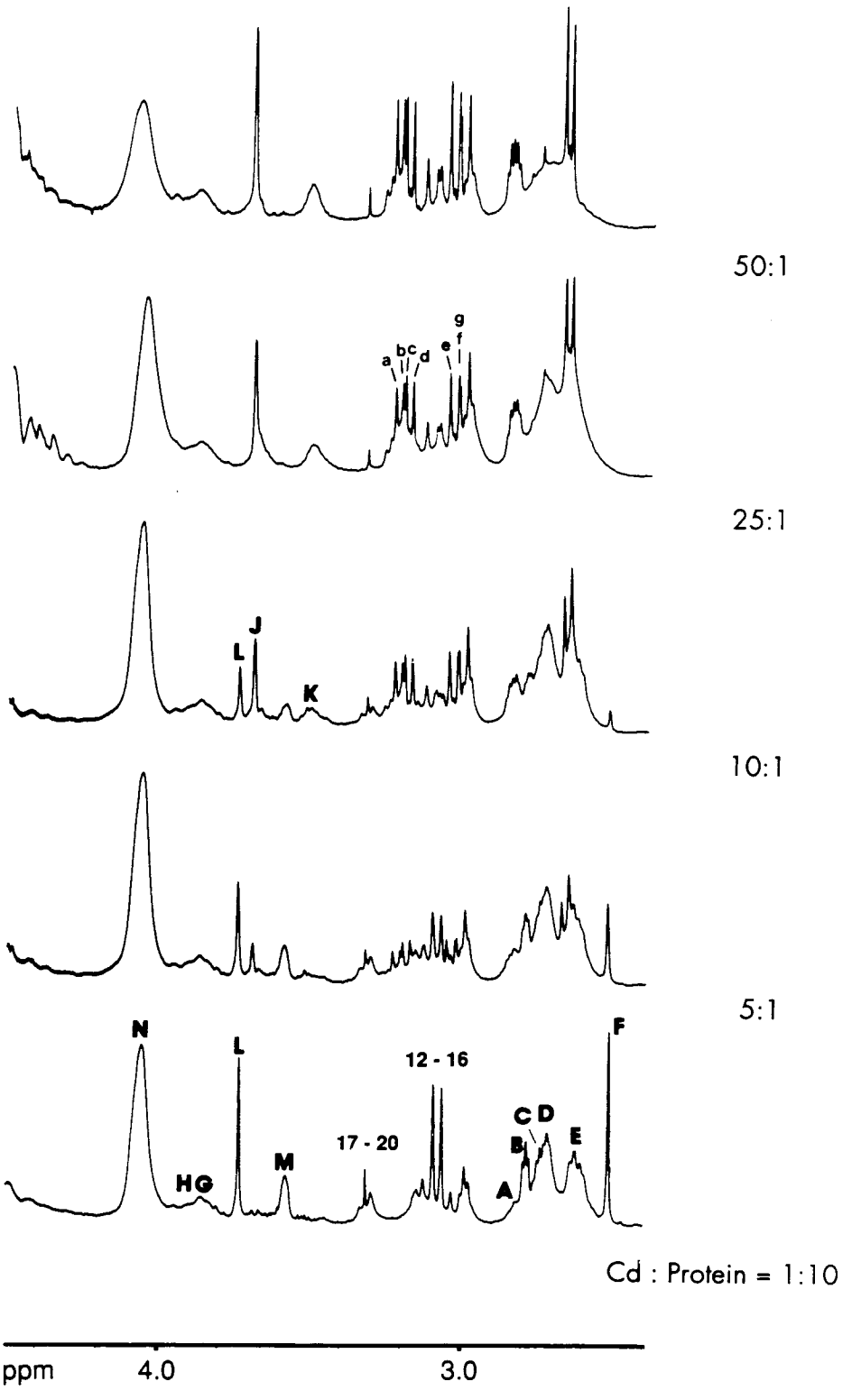
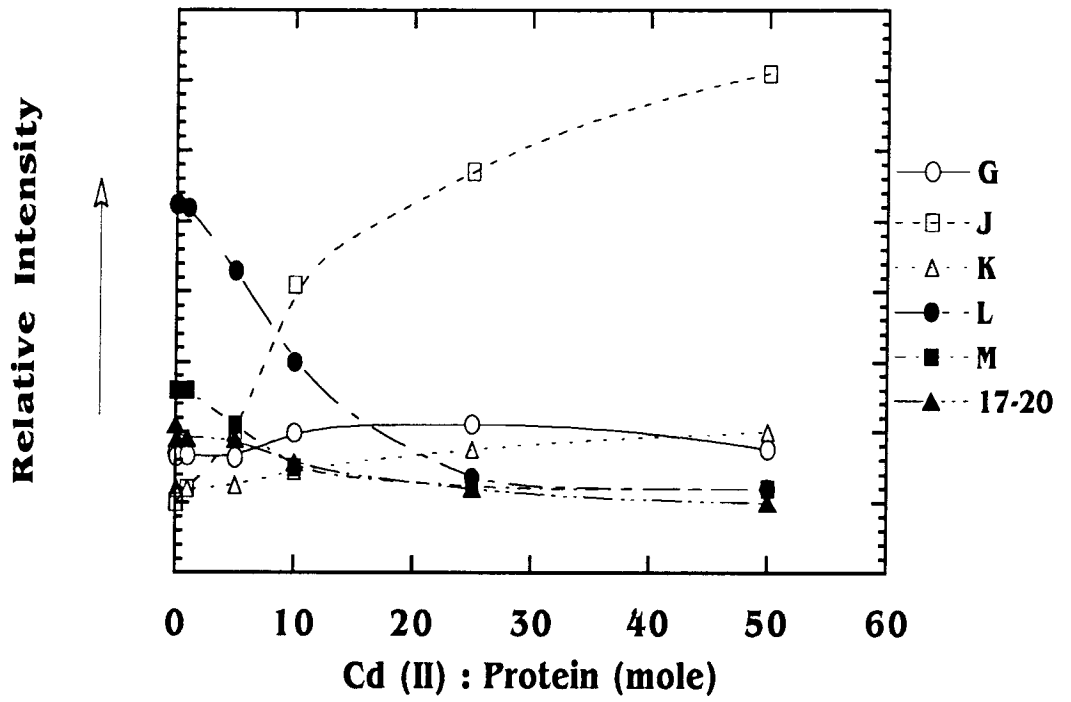

50:1
a
b
c
d
e
f
g
25:1
L
J
K
10:1
5:1
N
L
12 - 16
F
C
D
B
E
17 - 20
M
HG
A
Cd : Protein = 1:10
ppm
4.0
3.0

Figure 3: *The dependence of non-exchangeable BDPP hinge region sidechain resonance intensities on Cd (II) addition.*

The *y*-axis is given in arbitrary units. Curve fitting was performed using a smoothing function in the KALEIDAGRAPH software program.


Relative Intensity
G
J
K
L
M
17-20
0
10
20
30
40
50
60
Cd (II) : Protein (mole)

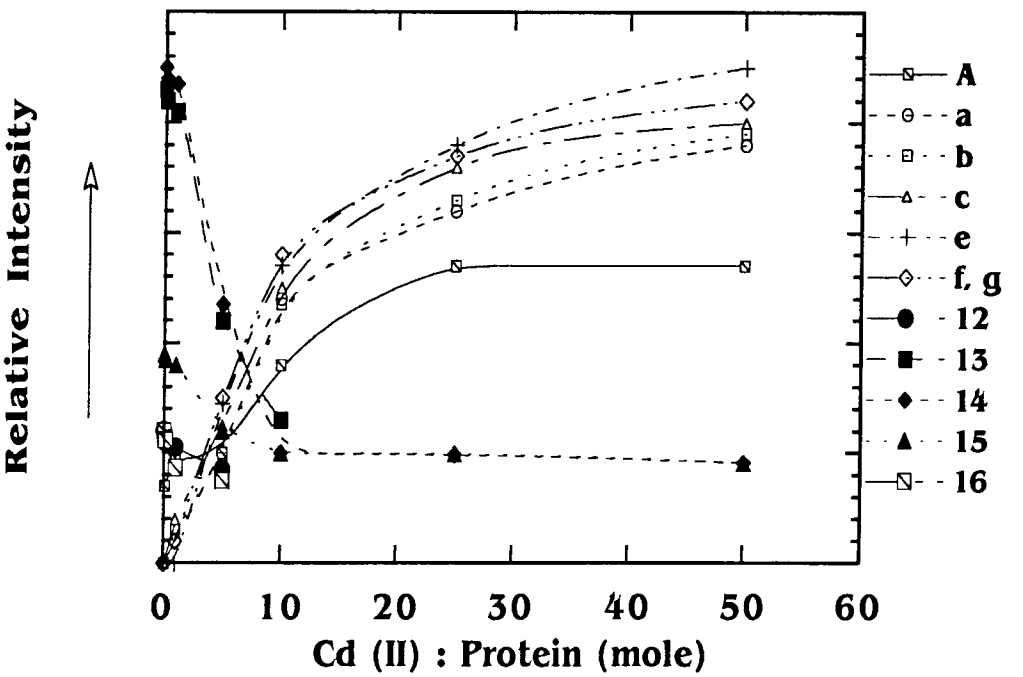

Relative Intensity
A
a
b
c
e
f, g
12
13
14
15
16
0
10
20
30
40
50
60
Cd (II) : Protein (mole)

Figure 4: *"Clean" 2-D HOHAHA spin-lock correlation contour plot of the upfield aliphatic frequency region of BDPP at pH 7.21, 50:1 Cd (II):Protein stoichiometry.*

Acquisition and processing parameters are given in the text. ^{1}H NMR chemical shifts are referenced from 0.1 M TSS in D_2O.

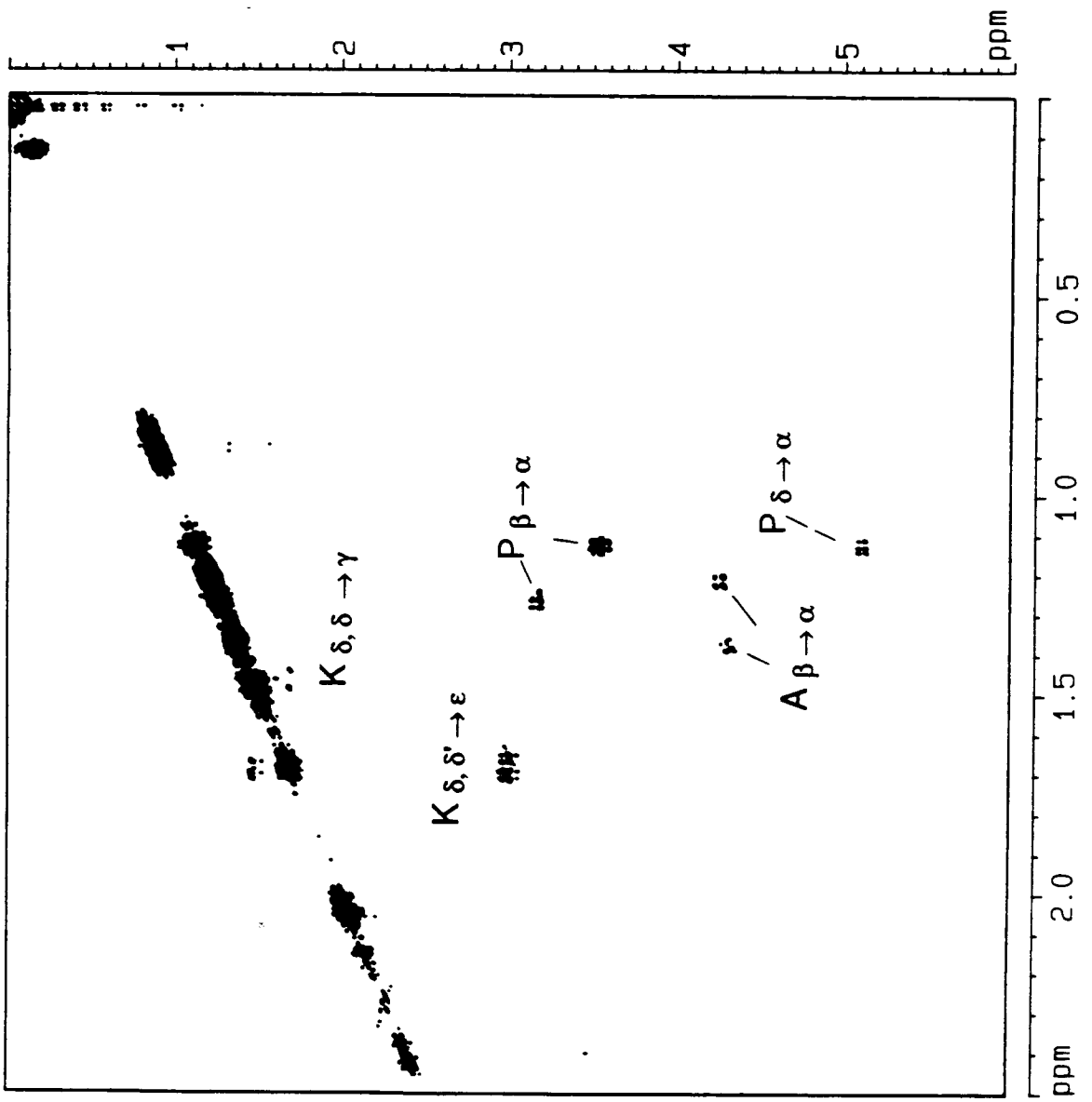

1
2
3
4
5
ppm
K δ, δ → γ
K δ, δ' → ε
P β → α
A β → α
P δ → α
0.5
1.0
1.5
2.0
ppm

Figure 5: *202.458 MHz ^{31}P NMR spectra of BDPP PSer residues as a function of Cd (II) addition.*

A Lorentzian apodization function of 4 Hz was utilized. For clarity, the spectra at 50:1 is displayed at 1.5 x the vertical scaling of the 1:10 spectra. All other spectra intensities are scaled relative to the 1:10 spectrum. Other acquisition and processing parameters are given in the text.

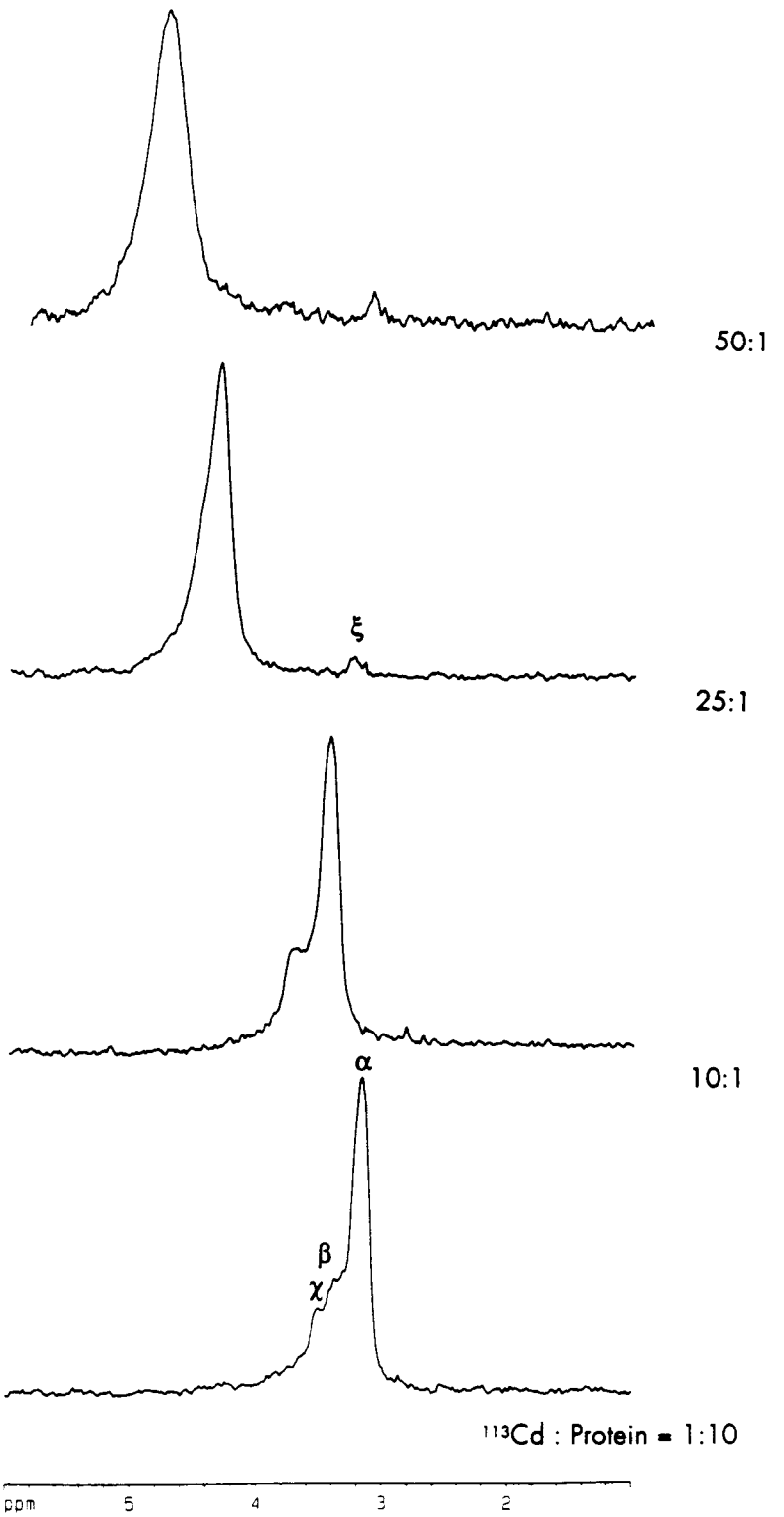

50:1
ξ
25:1
α
10:1
β
χ
^{113}Cd : Protein = 1:10
ppm
5
4
3
2

Figure 6: Dependence of ^{31}P PSer chemical shifts on Cd (II) addition.

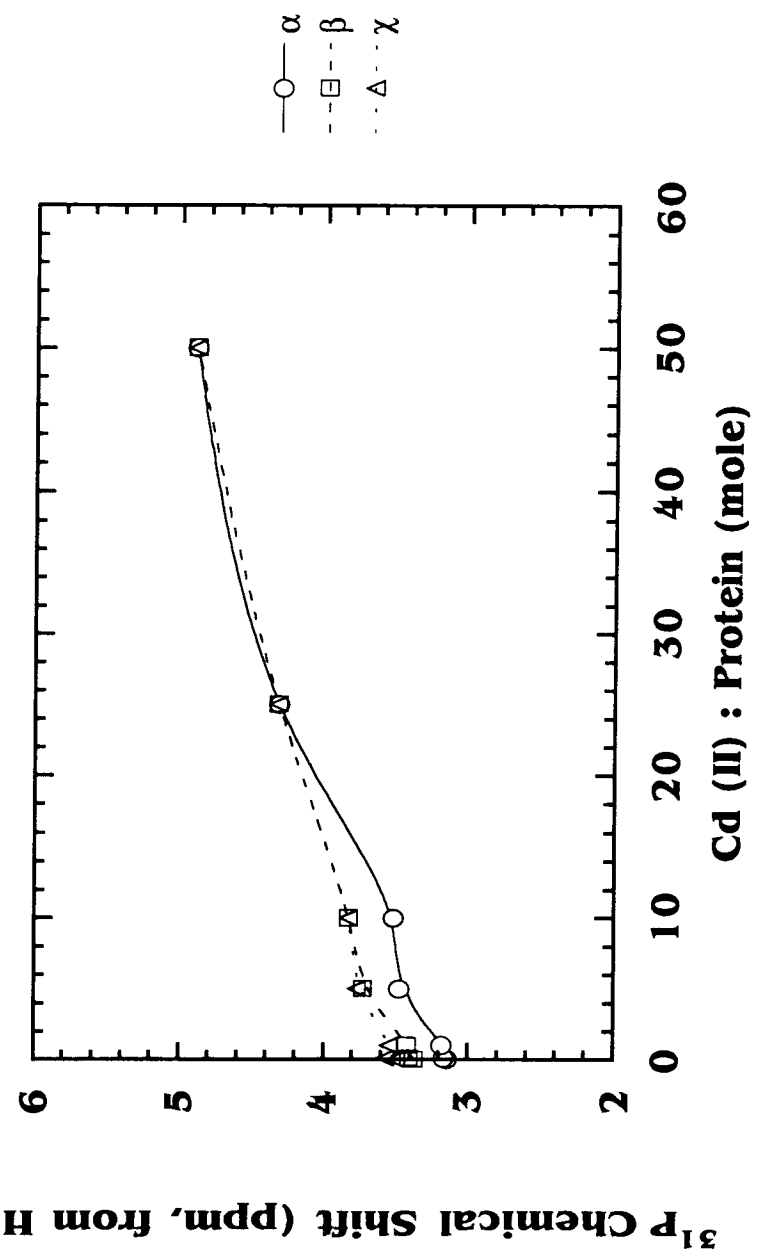

α
β
χ
6
5
4
3
2
0
10
20
30
40
50
60
Cd (II) : Protein (mole)
^{31}P Chemical Shift (ppm, from H_3PO_4)

Figure 7: *125.759 MHz natural abundance ^{13}C GARP-^{1}H decoupled NMR spectra of BDPP at 0 and 50:1 Cd (II):Protein stoichiometries.*

The downfield Asp C_γ carboxylate resonances are featured. A Lorentzian apodization function of 6 Hz was applied during exponential multiplication.

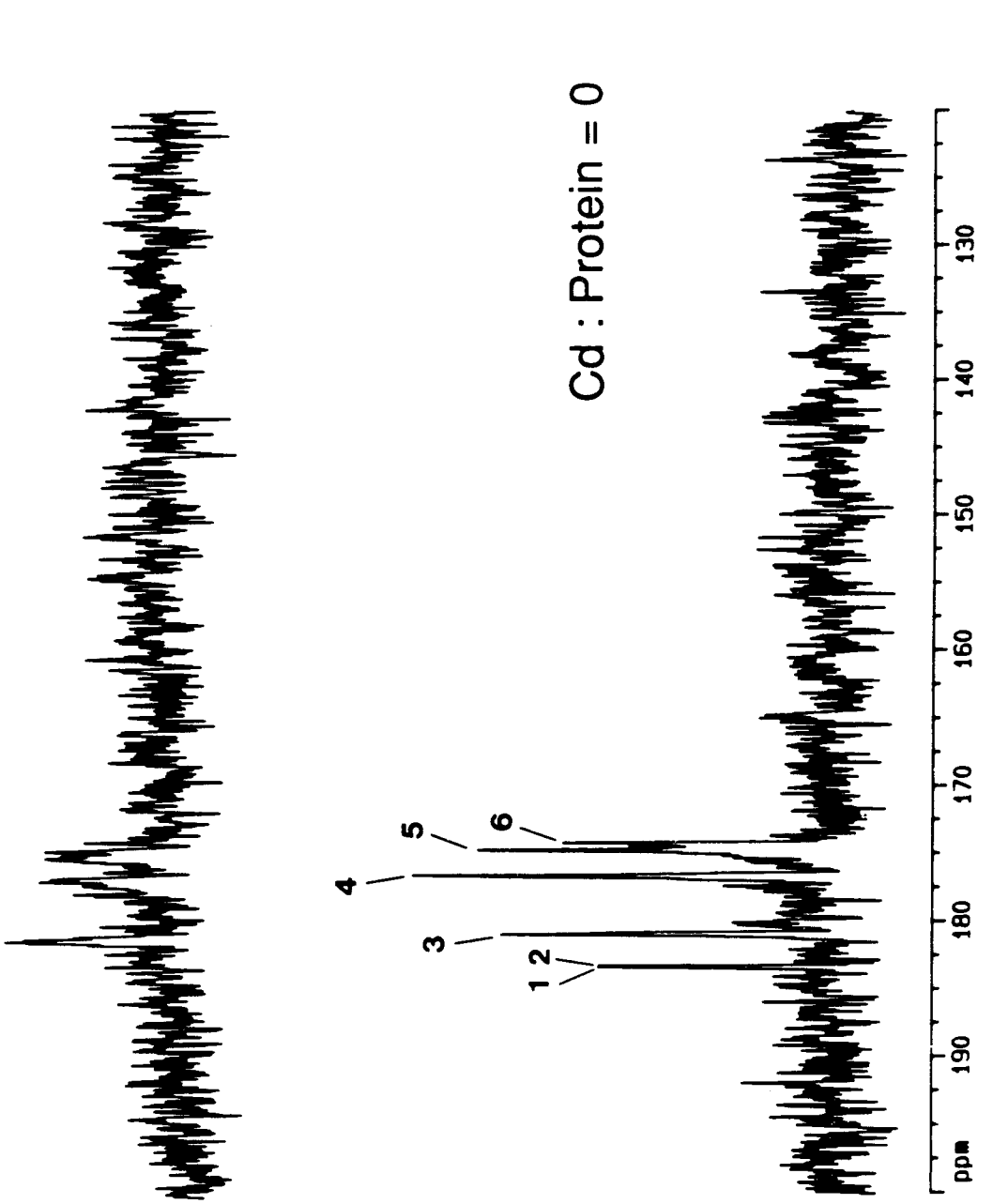

Cd : Protein = 50 :1
Cd : Protein = 0
1
2
3
4
5
6
130
140
150
160
170
180
190
ppm

Figure 8: *Proposed mechanism for Cd (II) induced folding of BDPP at low ionic strength, under conditions of divalent cation undersaturation.*

Legend to Figure: **Sawtooth structures**: PCBD regions; **Hinge structures**: HD regions; **Ellipsoial structures**: HC regions; **Spheres**: divalent cations. Here, the protein is arbitrarily represented as having paired PCBD domains of $(PSer)_n$, $(PSerAsp)_n$, and $(Asp)_n$, with each PCBD pair interrupted by a corresponding HD region. Structure A represents the unloaded or divalent cation-depleted conformational state; subsequent Structures detail the folding process as a function of sequential divalent cation loading.

A B C D

Nine

Synthesis and Characterization of a Ca (II)-Binding Polyelectrolyte Peptide (INTERV-1) Which Mimics a Template Macromolecule.

INTRODUCTION

The repeating polyelectrolyte-hinge-polyelectrolyte structure [Evans and Chan, 1992a-d] comprises an unknown number of regions in a particular class of template macromolecule proteins known as the Polyelectrolyte Mineral Matrix Protein, or PMMP. This polyelectrolyte-hinge-polyelectrolyte motif, abbreviated as PHP, can be conceptualized as two regions: (1) The *polyelectrolyte cation binding domain* (abbreviated PCBD), which is a continuous sequence stretch of anionic amino acids, such as Asp, Glu, and *o*-phosphoserine. This region is believed to bind and coordinate divalent cations with high affinity [Cookson *et al.*, 1980; Evans and Chan, 1992c,d]. (2) The *hinge* or *intervening* domain (abbreviated as HD), a sequence stretch of primarily neutral amino acids that is flanked on either side by PDCB regions. This region is hypothesized to function as a true "hinge" and permit the flanking PCBD regions to fold relative to one another [Evans and Chan, 1992e]. The mimimum sequence length requirements for each unit of the PHP motif is not known. However, given the amino acid composition of the known PMMP proteins [Rahima and Veis, 1988; Weiner and Hood, 1975; refs.], and recent preliminary sequence data [Sabsay *et al.*, 1991; Evans and Chan, 1991], it can be estimated that the PCBD regions could vary from 4 to 100 residues in length.

The role of the PMMP proteins, that of a nucleation center for inorganic solid formation, requires that the protein adopt a stereochemically correct arrangment of charged ligand groups along its exterior and interior surface [Mann, 1988]. Divalent cations then recognized this surface as an array of "active sites", not unlike one of the crystallographic planes of anions that exists in a biomineral. The cations then "sit" and bind to the template surface, thus creating a fixed array of metal atoms. Apposition then proceeds off of this fixed metal atom array. In a similar manner, the template PMMP protein can recognize the metal atom array of a forming crystal face and "sit" on this array, thus preventing further apposition at this crystallographic face. It has been postulated that the protein creates this stereochemical ligand or high charge density arrangement through the process of protein folding [Evans and Chan,

1992]. This would most likely involve the alignment of the PCBD regions relative to one another via the hinge regions. How this alignment occurs is not known, but it is thought that divalent cations, through either electrostatic shielding and/or ligand coordination, may mediate this folding process [Evans and Chan, 1992; Evans and Chan, unpublished observations].

A straightforward approach towards the study of the PHP motif folding problem would involve the creation of a PHP peptide model, i.e., a hinge region flanked on either side by two defined sequence length PDCB regions. Recently, preliminary peptide mapping data was obtained for rat α-phosphophoryn, a Ca (II)-binding PMMP protein isolated from the calcium phosphate matrix of tooth dentine [Sabsay *et al.*, 1991]. A 17-residue trypsin digest fragment, -DDDDDDYSDSDSSDSDD-, was identified and sequenced, using liquid-phase acid sequencing methods. Given that the Asp-Asp dipeptide linakge is extremely susceptible to peptide bond cleavage via decarboxylation and β-elimination in the presence of strong acid [Weiner and Hood, 1975; Weiner, 1983; Sabsay *et al.*, 1991; Borbas *et al.*, 1991], it is likely that the C-terminal region of this peptide fragment may have possessed a longer sequence stretch of Asp residues which were subsequently degraded. This 17 residue fragment most likely represents a potential PHP motif, where the 9-residue stretch, YSDSDSSDS, is the hinge region, flanked on either side by a poly-L-$(Asp)_n$ PCBD region.

This report describes the synthesis, preliminary isolation, and conformational characterization of INTERV1[1], a 21 residue capped polypeptide:

***N-acetyl*-DDDDDDYSDSDSSDSDDDDDD-C-amide (I)**

INTERV1 is a 4 Asp residue extension at the C-terminus of the 17 residue rat a-phosphophoryn fragment identified by Sabsay *et al.* [1991]. This

[1]INTERV1 is an acronym for INTervening Electrostatic pRotein MotiVes, fragment 1.

extension was performed to make each PDCB region symmetric in sequence length. Using ^{1}H NMR spectroscopy, we compare the conformational folding of this peptide at low and high pH, and in the presence of the Ca (II) ion analogue, ^{113}Cd (II), and pose the question: does the folding of this PHP peptide model mimic the pH/metal ion induced-conformational folding of the parent protein molecule, phosphophoryn?

MATERIALS AND METHODS

Reagents. For FMOC peptide synthesis, peptide synthesis grade solvents [dichloromethane (DCM), *N*-methyl-2-pyrrolidone (NMP), methanol (MeOH)]; piperidine, trifluoroacetic acid (TFA), N,N-dicyclohexylcarbodiimide (DCC) in NMP (1 M); 1-hydroxybenzotrazole (HOBt) in NMP (1 M); and acetic anhydride were obtained from Applied Biosystems (Foster City, California). Phenol, 1,2-ethanedithiol, thioanisole, 2,4-butanedithiol, and hexamethylphosphoramidate (HMPA) were purchased through Janssen Chemicals; ethanol and glacial acetic acid from Fisher Chemicals; ninhydrin, deuterium chloride (DCl)(20% v/v solution; 98% atom-*d*), sodium deuteroxide (NaOD)(30% v/v solution, 98% atom-*d*) and potassium cyanide (KCN) from Aldrich Chemicals (Milwaukee, WI). N-α-9-fluorenylmethoxycarbonyl (FMOC)-t-butyl ester (oBu^{t}) protected amino acids L-Asp, L-Ser, and L-Tyr from Bachem (Torrance, California); N-α-FMOC-a-pentafluorphenyl ester oBu^{t} derivatized amino acids L-Asp, L-Ser, and L-Tyr, and FMOC-L-Asp-oBu^{t}-dimethoxyalkoxybenzylamine resin (100-200 mesh, polystyrene 1% divinylbenzene (DVB) crosslinked, 0.41 mEq/g) from Penninsula Laboratories (Belmont, California); Ethylenediaminetetraacetic acid (EDTA), potassium chloride (KCl), and ethylene glycol-O,O'-bis(2-aminoethyl)-N, N, N',N'-tetraacetic acid (EGTA) from Fluka Chemicals; Chelex-100 resin was purchased from Bio-Rad (Richmond, CA). Deuterium oxide (D_2O)(99.9% atom-*d*) and cadmium oxide (^{113}CdO, 91.67% atom113-Cd) were purchased from Cambridge Isotope Labs. ^{13}C-(methyl)-tetramethylsilane (TMS) (33.7% atom ^{13}C) was obtained from MSD Isotopes. Tris-HCl was obtained from Bethesda Research Laboratories (Bethesda, MD). $^{113}CdCl_2$ stock solutions (2.5 M

and 250 mM) were prepared by stoichiometric addition of concentrated HCl to ^{113}CdO solid, then diluting to the appropriate volume with DD_{water}. Cadmium stock concentrations were verified by ICP analysis. 0.2 μm filtered deionized distilled water (DD_{water}) was used in the solubilization of the resin-cleaved peptide and in all subsequent purification steps; all glassware used in NMR sample preparation and analysis was first treated with saturated EDTA/EGTA to remove unwanted divalent cations.

FMOC Solid-Phase Synthesis of INTERV1. A modified orthogonal scheme for FMOC synthesis [Chang and Meienhofer, 1978; Fields and Noble, 1990] of the 21-residue INTERV1 peptide was performed manually at 25° C using a filter frit reaction vessel. The aminoacylation cycle of FMOC-protected amino acids to FMOC-L-Asp-oBu^t-dimethoxyalkoxybenzylamine resin (0.2 mmole scale synthesis) was conducted according to a modification of established protocols [Chang and Meienhofer, 1978] as follows: (i) flow wash with NMP solvent (20 column volumes). NMP was found to be a better solvent for peptide-bound resin solubilization than dimethylformamide (DMF)[Fields and Fields, 1991]. (ii) piperidine FMOC deprotection (50% piperidine/NMP, 4 mL), 1 min wash, for a total volume of 70 mL. Aliquots withdrawn from expended piperidine washes were spectrophotometrically monitored for the decrease in *N*-(9-fluorenyl)methylpiperidine formation (OD_{301nm} reading), using 50% piperidine/NMP was as a blank. (iii) flow wash with NMP; (iv) Aminoacylation (residues 20-11) was performed in NMP, mixing 4 equivalents of each FMOC amino acid (0.8 mmole in 2.0 mL NMP) and 2 equivalents each of HOBt and DCC (both 1 M in NMP) in glass vials, then transferred to the resin solution. Two coupling cycles (30 min/60 min duration) were performed for each residue for the first 5 residues; this was increased to 3 cycles (30 min/60 min/60 min) from residue 5 to 11. (v) Completeness of each acylation was monitored quantitatively using the ninhydrin reaction (570 nm spectrophotometric monitoring) on the free amine (using 2-5 mg resin sample)[Sarin *et al.*, 1981]. (vi) When incomplete aminoacylations were encountered, the resin was subjected to NMP flow wash and recoupled under the following conditions: (a) the substitution of FMOC pentafluorophenyl ester-activated L-amino acids [Atherton and

Sheppard, 1985, Atherton *et al.*, 1988] for FMOC-L-amino acids and DCC/HOBt coupling, and (b) employing HMPA instead of NMP as the coupling solvent. Two successive coupling cycles (30 min/60 min) were used per residue. Midway in the synthesis, it was discovered that standard HOBt/DCC coupling in NMP gave poor coupling efficiencies (< 95%), most likely as a result of unfavorable sequence-dependent secondary structure formation in NMP as peptide elongation progressed past 10 residues in length [Van Woerkom and Van Nispen, 1991; Milton *et al.*, 1990], changes in resin swelling [Milton *et al.*, 1990], or, due to pendant peptide solubility problems [Fields and Fields, 1991]. (vii) When necessary after the addition of each residue, the pendant peptide was *N*-acetylated by the addition of 10 equivalents of acetic anhydride in 2 mL DCM (30 min, 25° C). The N-terminal FMOC group was removed on completion of the target sequence, and the finished peptide was capped with acetic anhydride.

Sidechain Deprotection and Cleavage From the Resin. The N-α-acetyl-INTERV1 peptide was deprotected and cleaved from the resin as the C-terminal amide using a slight modification of the Reagent K protocol of King *et al.* [1990]. Cleavage reactions of washed, filtered, dry peptide resin were carried out at a concentration of 10 mg resin/1 mL volume containing 82.5% v/v TFA, 5% w/v phenol, 5% v/v DD_{water}, 5% v/v thioanisole, 1.25% v/v 1,2-ethanedithiol, and 1.25% v/v 2,4-butanedithiol. The reaction was carried out at 25° C for 30 min. The reaction mixture was then filtered on ice; the supernatant was neutralized with the addition of 1 N NaOH to a pH of 6.0, and the mixture was then extracted 3 x with diethylether, discarding the organic layer and re-extracting the aqueous layer. The resulting aqueous layer was then dialyzed vs. 0.5 M EDTA/0.1 M EGTA for 12 hrs at 4° C, using Spectra/Por CE cellulose ester dialysis tubing (MWCO = 2kD). The crude peptide mixture was then dialyzed vs. DD_{water} at 4° C for 1 week (dialysis volume ratio: 2000:1), concentrated using a Amicon YM-2 membrane (MWCO = 1 kD), and lyophilized.

Peptide Purification. The cleaved, dialyzed, lyophilized INTERV1 peptide mixtue was applied to a Pharmacia FPLC Mono-Q column at 25° C in

5 mM Tris-HCl, pH 7.1, at a flow rate of 0.5 mL/min, with absorbance monitoring at 254 nm. After a ten-minute wash with the starting buffer, a linear gradient of 2.0 M NaCl in 5 mM Tris-HCl, pH 7.1, was applied from 0% to 100% over a 50 minute period. A single peak eluted at approximately 1.5 M NaCl concentration. The peak was collected and applied to a FPLC Fast Desalt G25 column, using a running buffer of 10 mM $(NH_4)HCO_3$, pH 7.0 A single peak was detected at 254 nm, eluted, collected, and dialyzed against DD_{water} for 7 days at 4° C using Spectra/Por 7 membrane (MWCO=1.0 kD), and lyophilized. The lyophilized peptide was dissolved in DD_{water}, then treated with Chelex-100 resin to remove unwanted divalent cations. The solution was filtered (0.2 μm), re-lyophilized, and the peptide was stored at -80° C until needed.

Peptide Characterization. After Chelex-100 treatment, the lyophilized peptide sample was analyzed for amino acid content via 6 N HCl hydrolysis at 160° C and subsequent HPLC mapping on an Applied Biosystems ABI 130A Separation System linked to a ABI 420A Derivatizer and Model 900A Control Data Analysis Module. A parallel, non-acetylated peptide sample was dissolved in DD_{water} (10 μg/mL) and sequenced using liquid-phase methods on an Applied Biosystems ABI 477A Liquid Phase Peptide Sequencer linked to a ABI 120A PTH Analyzer. A total of 23 cycles were performed on a single sample. Cycles were manually analyzed and compared to computer analysis of the cycles. For sequence and amino acid composition determinations, three runs on parallel samples were performed.

The acetylated peptide was analyzed by capillary zone electrophoresis (CZE)(100 μg/mL INTERV1 sample in DD_{water}) in electrophoretic runs at pH 2.5 [20 mM sodium phosphate buffer (NaP), normal polarity mode] and at pH 11.0 [20 mM sodium 3-cyclohexyl-amino-1-propane-sulfonic acid (CAPS) buffer, reverse polarity mode], using an Applied Biosystems 270A Capillary Electrophoresis System running at 30° C, 25 kV constant applied voltage, electrokinetic loading, with monitoring of eluted components via absorbance at 200 nm. Capillaries utilized for separation were made of fused silica glass, and possessed dimensions of 72

cm total length, and an inner diameter of 75 mm. Applied sample volume ranged from 10 nL (pH 2.5 run) to 42 nL (pH 11.0 run).

Molecular weight determination was performed on an Applied Biosystems BioPolymer time-of-flight Mass Spectrometer BIO-ION 20R. The instrument was equipped with a reflecton flight tube (24.3 cm path), with the capability of a 270 cm path in the reflected mode. The lyophilized peptide sample was reconstituted in 50% EtOH/1% v/v TFA in DD_{water} (50 mL) and applied to nitrocellulose sprayed aluminized mylar foil grids. The foils were washed gently with DD_{water} to remove excess solute ions and dried under vacuum. Time-of-flight was determined in the normal mode (positive ion detection, +15 kV, under 5×10^{-7} mbars vacuum) and in the reverse polarity mode (negative ion detection, -11 kV, under 5×10^{-7} mbars vacuum)

Multinuclear NMR Spectroscopy of INTERV1 in the Protonated and Deprotonated Folding States. The INTERV1 acetylated peptide was dissolved in 400 ml of 99.9% D_2O (final peptide concentration: 1.2 mM; no buffer solution utilized), and the pH was adjusted with microliter aliquots of DCl or NaOD, using an Ingold 6030-02 glass microelectrode and Radiometer PHM83 Autocal pH meter. ^1H and ^{13}C NMR spectra were recorded at pH 7.08 and 2.15, uncorrected for deuterium isotope effects. All NMR experiments were conducted at 300 K, using a Bruker AMX-500 NMR Spectrometer/Aspect X-32 System with variable temperature control, and a 5 mm inverse mode low-gamma probehead.

One-dimensional ^1H NMR (500.138 MHz) and proton-decoupled natural abundance-^{13}C NMR (125.771 MHz) spectra of INTERV1 were obtained at pH 2.15 and 7.08, using the following acquisition parameters: (1) for ^1H NMR, a spectral window of 8 kHz, a recovery delay of 1.5 sec, a pulse width of 5 μs (^{a} 45° pulse), receiver gain of 40 dB, 16 K data points (corresponding to a digital resolution of 0.492 Hz/pt), 2 dummy scans, and 120 scans were utilized. The residual HOD signal was suppressed via a presaturation pulse during the recovery phase, using a decoupler setting of 60 dB attenuation. Further information on acquisition and processing

parameters can be found in the figure legends for each spectra. Proton chemical shifts are reported relative to external 2,2,2-trimethyl 2-silapentane sulfonic acid (TSS, 0.1 M in D_2O).

^{113}Cd Titration and NMR spectroscopy. The INTERV1 peptide, at pH 7.16, in 99.9% D_2O, was titrated with microliter quantities of $^{113}CdCl_2$ stock solution (see above), and the pH of the peptide sample was checked, and if necessary, the pH was adjusted to 7.16 via addition of DCl or NaOD stock solutions. The concentration of Cd (II) and Na (I) in the sample were determined via ICP analysis. After pH titration, the total content of Na (I) was found to be approximately 22 mM, as determined by ICP analysis.

Multidimensional NMR Experiments. ^1H amino acid spin system assignments at pH 2.15, 7.08, and in the presence of $^{113}Cd(II)$ were obtained via "clean" TPPI phase-sensitive ^1H - ^1H 2-D TOCSY spin-lock correlation experiments [Griesinger *et al.*, 1988; Davis and Bax, 1985; Bax and Davis, 1985a] where cross-relaxation is suppressed by the usage of two delays before and after the 180_y^o pulse in the MLEV-17 spin-lock sequence, using the Bruker microprogram CLMLEVPRTP. Typical acquisition parameters include a spectral window of 6 kHz in the F1 and F2 dimension, a 90° ^1H pulse of 10.2 μs, 2 dummy scans per experiment, 48 scans per experiment, a total of 512 experiments, a cross-relaxation suppression delay of 97.5 μsec, a t_1 incrementation of 3 μsec, 2 K data points in both the F1 and F2 dimension (corresponding to a digital resolution of 2.94 Hz/pt), a recovery delay of 1.5 sec, a SL_x trim pulse of 2.5 msec, and an MLEV spin-lock mixing time of 74 msec. Solvent suppression was achieved by the use of a presaturation decoupler pulse (corresponding to a decoupler power setting of 60 dB of 1.5 sec duration) during the recovery phase of each experiment. ^1H NMR chemical shifts are referenced from external TSS (0.1 M in D_2O).

RESULTS

Characterization of INTERV1 Peptide Generated by FMOC Solid-Phase Peptide Synthesis. Our rationale for utilizing the FMOC solid-phase synthesis route for INTERV1 rests on the observation that Asp-Asp peptide linkages are notoriously labile in the presence of strong acid [Weiner and Hood, 1975]. This being the case, it was necessary to avoid the use of t-BOC chemistry, due to the application of strong, volatile acids such as anhydrous HF or TFMSA for peptide cleavage from the resin. It was found despite the presence of scavengers, both acids were deleterious to the nascent t-BOC synthesized Poly-L-(Asp)10 peptide sequence [unpublished observations]. The FMOC chemistry utilizes milder acid hydrolysis (TFA), and with the use of Reagent K scavengers [King *et al.*, 1991], we were able to minimize acid-generated free radical attack on the labile Asp and Tyr residues in the INTERV1 sequence. In addition, immediately after TFA deprotection and cleavage from the resin, the peptide mixture is immediately neutralized with the addition of NaOH to minimize acid-catalyzed peptide bond cleavage. Table 1 presents the results of the amino acid composition, liquid-phase sequencing, and mass spectrometry- determined molecular weight of INTERV1, along with the predicted values for each category. A comparison of the theoretical amino acid composition values with the experimentally determined values indicates that the synthesized peptide is homogeneous in its composition. This is further underscored by the amino acid sequencing data [Table 1], wherein the N-terminal portion of purified INTERV1 up to residue 11 and from residue 16 to 21 is consistent with the target sequence. The letter "X" denotes uncertainties as to whether the residue is either S or D. The premature Asp-Asp dipeptide cleavage phenomenon does hinder accurate sequence determination under acid conditions, as revealed by uncertainties in the sequence data after 11 cycles in the sequencer. This was also observed in our N-terminal sequencing of full-length BDPP [Evans and Chan, 1991]. As a further check, time-of-flight mass spectrometery data [Table 1] was performed on the purified INTERV1 peptide. The molecular weight of the peptide sample is calculated to be 2661.5 amu, which is above the theoretical molecular weight of 2341.66 amu for the N-acetyl-C-terminal amide peptide in the

acid form. However, since the peptide sample is in aqueous solution as the sodium salt form, if one takes into account the presence of counterions (in this case, Na^+), then adjusting for condensation (15 Na^+: peptide, MW Na = 23 amu), the theoretical molecular weight thus becomes 2686.66 amu, which is within 1% of experimentally determined value. It should be noted that this charged peptide does not give rise to a sharp m/z peak in the time-of-flight spectrometer, but rather, a broad peak (width = 400 amu). This "broadening" phenomenon has also been observed for time-of-flight spectra obtained for Poly-L-$(Asp)_{10}$ polyelectrolyte peptides (unpublished observations).

Amino Acid Spin Assignments and Monitoring pH-Induced Peptide Folding in INTERV1 by 500 MHz ^1H NMR Spectroscopy. The one- dimensional ^1H NMR spectra for INTERV1 D_2O is presented in Figure 1. Table II gives the exchangeable and non-exchangeable ^1H spin system assignments for INTERV1 based on the "clean" HOHAHA spectra of INTERV1 in D_2O (Figures 2, 3, 4) and the integration of the 1-D ^1H NMR spectrum. Several problems are encountered in determining the solution structure of INTERV1. Given the redundancy of the peptide sequence, the NH resonances exhibit little if any chemical shift dispersion due to secondary structure formation (i.e., intraresidue hydrogen bonding) as seen in helical peptide structures [Klaus and Moser, 1992][Figure 2]. If stabilized intraresidue hydrogen bonding is not extant in the INTERV1 peptide, then a large percentage of the peptide NH atoms will be easily accessible to solvent hydrogen-deuterium acid and base catalyzed exchange at neutral pH [Blout *et al.*, 1961; Wada, 1960; Lurmy *et al.*, 1964]. For a hydrophilic peptide with random coil structure, one would predict a fast exchange rate for the amide NH atoms [Blout *et al.*, 1961; Wada, 1960; Lurmy *et al.*, 1964]. However, if there are different rates of exchange for each site, then one might expect a superpositioning of narrow and broad NH resonances. In fact, this is what is observed for INTERV1 amide resonances [Figure 2]: narrow ($\Delta\nu_{1/2} \approx$ 12-24) and broad linewidths ($\Delta\nu_{1/2} \approx$ 24-60 Hz) are observed for the majority of the INTERV1 amide NH resonances. The resulting exchange broadening (i.e., $\Delta\nu_{1/2}$) can be directly interpreted in terms of average lifetime of the proton or deuteron between successive exchanges

(Meiboom, 1960; Berger and Linderstrom-Lang, 1957; Bryan and Nielsen, 1960]. Thus, chemical shift overlap and solvent-induced exchange broadening effects result in weak amide proton resonance intensities in the INTERV1 peptide fingerprint region. This makes sequence-specific assignments via homonuclear interresidue cross-relaxation magnetization transfer (NOESY, ROESY) impossible at neutral pH or even at low pH (data not shown).

A second difficulty in determining sequence-specific assignments is the significant ^{1}H NMR β, β'-CH_2 chemical shift overlap for Asp and Ser residues in INTERV1 [Figure 3,4]. This problem is aggrevated by the near chemical equivalency of the PCBD regions, resulting from identical sequence length and composition. For this reason it is impossible to obtain sequence-specific amino acid spin assignments in INTERV1 by ^{1}H NMR spectroscopy alone. However, it <u>is</u> possible to identify resonances in the ^{1}H NMR spectrum that arise from Asp residues that reside in the hinge region and collectively in both of the PCBD regions. An examination of the TOCSY spectra at pH 7.08 in D_2O [Figure 3] reveals that there are Asp sidechains that experience a downfield shift of their α-CH resonances (centered at 5.640 ppm) relative to the bulk Asp α-CH ^{1}H NMR chemical shifts centered at 4.600 ppm. Another set of Asp α-CH resonances, centered at 4.400 ppm, are shifted upfield from the bulk Asp α-CH resonances. Since α-CH ^{1}H NMR shifts are sensitive indicators of backbone secondary structure in peptides and proteins [Grob and Kalbitzer, 1988; Pastore and Saudek, 1990; Wishart *et al.*, 1991; Szilagyi and Jardetzky, 1989; Williamson, 1990], it is clear that these Asp residues exist in a region of INTERV1 where the secondary structure is unique. This would most likely occur in the -YSDSDSSDS- sequence stretch, and not in either of the $(Asp)_6$ stretches, where one would predict a coil structure [Evans *et al.*, 1992]. This is supported in the corrsponding HOHAHA-H_2O spectrum [Figure 2], where a set of amide NH resonances (centered at 8.351 and 8.201 ppm, respectively) exhibit relayed $^3J_{HH}$ and $^4J_{HH}$ scalar coulpings with these unique Asp α-CH/β, β'-CH_2 resonances [Figure 3]. In contrast, 80% of the Asp NH resonances exist as two centerpeaks

(8.235, 8.224 ppm), with a downfield shoulder region at 8.247. Based on the unique α-CH and NH chemical shifts, it is believed that these resonances originate from Asp^{9}, Asp^{11}, and Asp^{14} in the hinge domain of INTERV1, and are not part of the PCBD domains.

pH and Cd (II) Dependent Folding of INTERV1 as Monitored by Two-Dimensional ^{1}H NMR Spectroscopy. By lowering the pH and protonating the anionic sidechain groups in BDPP, we were able to effect global conformational change through most of the protein molecule [Evans and Chan, 1992]. As shown in Figure 3, the INTERV1 peptide molecule also demonstrates pH-dependent backbone conformational effects. At pH 7.08 in the fully deprotonated state, the INTERV1 ^{1}H NMR "clean" HOHAHA spectrum exhibits a 0.22 ppm chemical shift dispersion for Hinge Ser a-CH resonances (Figure 3, left spectra). Upon lowering the pH below the pK_a of Asp COOH groups ($pK_a \approx 4 - 5$ in BDPP poly-L-(Asp) sequences [Evans and Chan, 1992]), the Ser α-CH resonances exhibit a greater degree of chemical shift overlap (0.12 ppm) [Figure 3, right spectra]. For Asp residues in the Hinge domain, the α-CH resonances downfield of the HOD solvent peak experience a downfield chemical shift effect of 0.08 ppm as the peptide experiences H^{+} neutralization. Simultaneously, the HD Asp α-CH resonances that are upfield of the solvent peak are no longer discernable in the HOHAHA spectrum. These change reflect conformational perturbation at the peptide backbone in the Hinge region of INTERV1. The PCBD regions also demonstrate pH-dependent conformational change, as evidenced by the following [Figure 3,4]: (1) As the pH decreases from 7.08 to pH 2.15, PCBD Asp β, β'-CH_2 diastereotopic protons exhibit increased chemical shift overlap and linewidth broadening [Figure 4]. At pH 2.15 the α-CH resonances for PCBD Asp residues exist in two populations: one, centered at 4.700 ppm, and a smaller population centered at 4.550 ppm [Figure 3]. (2) Approximately 60-70% of the α-CH resonances for PCBD Asp residues exhibit a downfield shift of 0.01 ppm from pH 7.08 to pH 2.15 [Figure 3].

Collectively, these observations are indicative of perturbation in the secondary structure of the anionic cluster PCBD and HD regions in INTERV1. Thus, as H^+ charge neutralization of Asp sidechains occurs, the INTERV1 molecule undergoes conformational change in both the Hinge and PCBD regions. This has also been observed for the intact BDPP protein [Evans and Chan, 1992].

In the presence of $^{113}Cd^{2+}$ divalent cation at low ionic strength, the INTERV1 peptide demonstrates complex interactions with Cd (II) as a function of stoichiometry. Figure 3, in which the cation-loading process is monitored at the level of ^1H NMR, clearly demonstrates that the INTERV1 peptide undergoes conformational change in the presence of increasing levels of Cd (II). An examination of the "clean" HOHAHA spectra of the INTERV1 α-CH proton fingerprint region reveals that the INTERV1 peptide has undergone a folding process in the presence of 17.5 mM Cd (II), or 15:1 ion:peptide stoichiometry at pH 7.16. At this stoichiometry, there is a 1:1 molar ratio between Cd (II) and Asp residues in the peptide. As was the case at pH 2.15, in the presence of saturating levels of Cd (II), the HD Ser α-CH resonances exhibit an identical degree of chemical shift overlap (0.13 ppm). Simultaneously, the HD Asp α-CH resonances upfield of the HOD solvent peak disappear from the HOHAHA spectrum [Figure 3], and, the downfield set of HD Asp α-CH resonances are weak in intensity. Thus, charge neutralization via divalent cation interactions results in a folding process that involves peptide backbone conformational change in the HD region. This divalent cation induced folding appears qualitatively similar to that generated in the presence of H^+.

DISCUSSION

Peptide mimetics and *de novo* has evolved into an important tool for understanding protein folding and domain function [Ruan *et al.*, 1990; Ghadiri *et al.*, 1992; Shaw et al., 1992]. A study of two synthetic helix-loop-helix peptide mimetics of the chicken skeletal troponin C revealed that in the presence of Ca (II), the two peptide sequences formed heterodimeric complexes with Ca (II) and underwent conformational change from

random-coil to b-sheet structure [Shaw *et al.*, 1992]. Recently, several investigations into peptide and protein folding employed the use of metal-peptide ligand formation to induce polypeptide assembly into simple helices [Ruan *et al.*, 1990] or into more complex triple-helix bundles [Ghadiri *et al.*, 1992]. This so-called "cross-linking" or self-organizing folding process is similar in spirit to the biomineralization template, where divalent cations induce conformational folding of a protein, thereby creating domain organization of anion atoms that mimics the complementary anion atom array in mineral crystals [Evans and Chan, 1992c,d]. Ironically, the self-assembly peptide mimetic model has received scant attention with regard to template-regulated controlled crystal growth. To date, all of these template mimetic studies have relied exclusively on peptides that are constructed from *hypothetical*, finite-length homopolymer $(Asp)_n$, $(PSer)_n$, $(Glu)_n$ sequence stretches [Addadi *et al.*, 1987, 1991; Lee and Veis, 1980; Sikes, 1991]. Not surprisingly, these anionic mimetics demonstrate Ca (II) ion binding and affinity for Ca (II) arrays on forming crystal faces [Addadi *et al.*, 1987; 1991; Sikes, 1991]. However, let us note the following: (1) Depending on the sequence, these polyanionic homopolymers exhibit differing nucleation and crystal growth regulatory capacities under <u>in vitro</u> conditions (e.g., Poly-L-$(Asp)_n$ is more efficient than Poly-L-$(Glu)_n$ in nucleation of calcium carbonates [Addadi *et al.*, 1987; 1991]), and (2) Polyanionic sequences alone do not always catalyze cation nucleation effectively (e.g., sometimes a neutral C-terminal sequence attached to the polyanionic peptide results in more effective binding to crystal surfaces [Sikes and Wheeler, 1988a,b]). Although it is acknowledged that polyanionic sequences bind divalent cations with high affinity, their "hit-and-miss" effectiveness as "mini-templates" suggests that some other factors are required. What would be more apropos is the testing of a *known template protein sequence* that features PCBD domains. In our present work, we investigate a known 17-residue sequence taken from phosphophoryn, a PMMP template macromolecule. The attractiveness of this sequence arises from the 9-residue -YSDSDSSD- stretch that contains predominantly neutral sidechains, in accordance with our postulate that PMMP proteins require neutral amino acids for proper template folding [Evans and Chan, 1992c,e]. With some subtle sequence modifications of this 17 residue

sequence, we have shown that this 21 residue peptide, INTERV1, undergoes conformational change as a function of charge neutralization either by H^+ or by Cd (II) addition. What is interesting about this process is that charge neutralization of Asp residues by either H^+ or Cd (II) addition results in backbone conformational change not only in the PCBD regions, but also in the 9-residue Hinge sequence region [Figures 3,4]. This property was observed in the pH and Cd(II) induced folding studies conducted on bovine dentine phosphophoryn [Evans and Chan, 1992d], and suggests that the role of the HD or neutral amino acids is to assist in the folding of the protein. It should be understood that the folding of INTERV1 is being assessed on a qualitative level, i.e., that of backbone peptide conformational change occurring in the HD domain. We are unable at this time to identify individual amino acids in either the HD domain sequence or in either of the PCBD regions. Thus, important parameters, such as changes in backbone dihedral angles or nOe/rOe connectivities, must await the use of stable isotope-labeling (i.e., $^{13}C/^{15}N$) and multinuclear 2-D, 3-D, and 4-D NMR experiments.

In light of our findings regarding the folding of INTERV1, we are interested in how the folding process might regulate the binding of metal ions to the template surface. In earlier ^{113}Cd NMR experiments with BDPP, it was observed that at levels of ion undersaturation relative to the total number of anionic sites on the protein, a large percentage of these sites were experiencing ^{113}Cd ion interactions [Evans and Chan, 1992d]. This was attributed to the "hopping" phenomenon originally described in earlier NMR experiments with a 30 kD rat phosphophoryn fragment. Unfortunately, direct observations of divalent cation interactions with BDPP via ^{113}Cd NMR spectroscopy was not possible, due to the presence of intermediate chemical exchange phenomena and rapid relaxation effects on the ^{113}Cd nuclei. However, as shown in the accompanying paper, ^{113}Cd species are observable in the presence of INTERV1, and in conjunction with our folding studies, provide interesting insights into the divalent cation binding process.

Summary. The successful use of FMOC solid-phase peptide synthesis in the creation of a template peptide mimetic, INTERV1, has led to the characterization of this peptide by two-dimensional ^1H NMR spectroscopy. With this peptide mimetic, we have confirmed our initial observations regarding BDPP folding in the presence of H^+ and divalent cations: namely, that the protein possesses a number of sequence regions, termed hinge domains (HD), that possess neutral amino acid residues. In turn, these regions are flanked by polyelectrolyte cation binding domains (PCBD), which act as sites of cation and/or crystal binding. Upon charge neutralization, INTERV1 undergoes backbone conformational change that involves the participation of the HD region, similar to that seen in the global folding of BDPP [Evans and Chan, 1992c,d].

ACKNOWLEDGEMENTS

We wish to thank John Racs, Tamara Bauer, and Jane Sanders for their technical expertise and assistance in the amino acid analysis, CZE, sequencing, and biomass spectrometry utilized in this report.
Thanks also to Ken Graham for his advice and input in the planning of the synthesis strategy of INTERV1. This project was supported (in part) by BRSG Grant RR07003 awarded by the Biomedical Research Support Grant Program, Division of Research Resources, National Institutes of Health. JSE acknowledges a Postdoctoral National Research Service Award (NIDR DE-05445-05) from the National Institute of Health, and a scholarship award from AMGEN Pharmaceuticals.

REFERENCES

Addadi, L., Moradian, J., Shay, E., Maroudas, N.G., and Weiner, S. (1987) *Proc. Nat. Acad. Sci USA.* 84 2732-2736.

Addadi, L., Moradian-Oldak, J., and Weiner, S. (1991) *ACS Symposium Surface Reactive Peptides and Polymers.* American Chemical Society, pp. 13-27.

Atherton, E., and Sheppard, R.C. (1985) *J. Chem. Soc. Chem. Commun.* 165-168.

Atherton, E., Cameron, L.R., and Sheppard, R.C. (1988) *Tetrahedron* 44 843-847.

Bax, A., and Davis, D.G. (1985a) *J. Mag. Res.* 65 355-360.

Bax, A., and Davis, D.G. (1985b) *J. Mag. Res.* 63 207-213.

Berger, A., and Linderstrom-Lang, K. (1957) *Arch. Biochem. Biophys.* 69 106-112.

Blout, E.R., de Loze, C., and Asadourian, A. (1961) *J. Am. Chem. Soc.* 83 1895-1896.

Borbas, J.E., Wheeler, A.P., and Sikes, C.S. (1991) *J. Expt. Zool.* 258 1-13.

Bothner-By, A.A., Stephens, R.L., Lee, J-m., Warren, C.D., and Jeanloz, R.W. (1984) *J. Am. Chem. Soc.* 106 811-813.

Bryan, W.P., and Neilsen, S.O. (1960) *Biochem Biophys. Acta* 42 552-558.

Cavanagh, J., and Keeler, J. (1988) *J. Mag. Res.* 80 186-194.

Chang, C.D., and Meienhofer, J. (1978) *Int. J. Peptide Prot. Res.* 11 246-249.

Cookson, D.J., Levine, B.A., Williams, R.J.P., Jontell, M., Linde, A., and de Bernard, B. (1980) *Eur. J. Biochem.* 110 273-278.

Davis, D.G., and Bax, A. (1985a) *J. Am. Chem. Soc.* 107 2820-2821.

Evans, J. S., and Chan, S.I. (1991) in *Proteins: Structure, Dynamics, and Design* (Renugopalakrishnan, V., Carey, P.R., Smith, I.C.P., Huang, S.G., and Storer, A.C. Eds) ESCOM Press, Leiden, Netherlands, pp. 251-260.

Evans, J.S., and Chan, S.I. (1992a) Manuscript in preparation.

Evans, J.S., and Chan, S.I. (1992b) Manuscript in preparation.

Evans, J.S., and Chan, S.I. (1992c) Manuscript in preparation.

Evans, J.S., and Chan, S.I. (1992d) Manuscript in preparation.

Fields, G.B., and Fields, C.G. (1991) *J. Am. Chem. Soc.* 113 4202-4207.

Fields, G.B., and Noble, R.L. (1990) *Int. J. Peptide Prot. Res.* 35 161-214.

Ghadiri, M.R., Soares, C., and Choi, C. (1992) *J. Am. Chem. Soc.* 114 825-831.

Griesinger, C., and Ernst, R.R. (1987) *J. Mag. Res.* 75 261-271.

Griesinger, C., Otting, G., Wuthrich, K., and Ernst, R.R. (1988) *J. Am. Chem. Soc.* 110 7870-7872.

Grob, K.H., and Kalbitzer, H.R. (1988) *J. Mag. Res.* 76 87-99.

Kaiser, E., Colescott, R.L., Bossinger, C.D., and Cook, P.I. (1970) *Anal. Biochem.* 34 595-598.

King, D.S., Fields, C.G., and Fields, G.B. (1990) *Int. J. Pept. Prot. Res.* 36 255-266.

Klaus, W., and Moser, R. (1992) *Protein Engineering* 5 (4) 333-341.

Lee, S.L., and Veis, A. (1980) *Int. J. Pept. Prot. Res.* 16 231-240.

Lumry, R., Legare, R., and Miller, W.G. (1964) *Biopolymers* 2 489-498.

Mann, S. (1988) in Biomineralization: Chemical and Biochemical Perspectives (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 35-62.

Meiboom, S. (1960) *Z. Elektrochem.* 64 50.

Milton, R.C. de L., Milton, S.C.F., and Adams, P.A. (1990) *J. Am. Chem. Soc.* 112 6039-6046

Neuhaus, D., and Keeler, J. (1986) *J. Mag. Res.* 68 568-574.

Pastore, A., and Saudek, V. (1990) *J. Mag. Res.* 90 165-176.

Rahima, M., and Veis, A. (1988) *Calcif. Tiss. Int.* 42 104-112.

Ruan, F., Chen, Y., and Hopkins, P.B. (1990) *J. Am. Chem. Soc.* 112 9403-9404.

Sabsay, B., Stetler-Stevenson, W.G., Lechner, J.H., and Veis, A. (1991) *Biochem. J.* 276 699-707.

Sarin, V.K., Kent, S.B.H., Tam, J.P., and Merrifield, R.B. (1981) *Anal. Biochem.* 117 147-157.

Shaka, A.J., Barker, P.B., and Freeman, R. (1985) *J. Mag. Res.* 64 547-552.

Shaw, G.S., Findlay, W.A., Semchuk, P.D., Hodges, R.S., and Sykes, B.D. (1992) *J. Am. Chem. Soc.* 114 6258-6259.

Sikes, C.S., and Wheeler, A.P. (1988a) in *Chemical Aspects of Regulation of Mineralization.* (Eds C.S. Sikes and A.P. Wheeler). University of South Alabama Publication Services, pp. 15-20.

Sikes, C.S., and Wheeler, A.P. (1988b) in *Biomineralization: Chemical and Biochemical Perspectives.* (Mann, S., Webb, J., and Williams, R.J.P., Eds.) VCH Publishers, New York, New York, pp. 95-131.

Szilagyi, L., and Jardetzky, O. (1989) *J. Mag. Res.* 83 441-449.

Van Woerkom, W.J., and Van Nispen, J.W. (1991) *Int. J. Pept. Prot. Res.* 38 103-113.

Wada, A. (1960) *Mol. Phys.* 3 409-415.

Weiner, S., and Hood, L. (1975) *Science* 190 987-989.

Weiner, S. (1983) *Biochemistry* 22 4139-4145.

Williamson, M. P. (1990) *Biopolymers.* 29 1423-1431.

Wishart, D.S., Sykes, B.D., and Richards, F.M. (1991) *J. Mol. Biol.* 222 311-333.

Table I: *Physical Properties of INTERV1 Synthesized by FMOC Solid-Phase Peptide Synthesis.*

Analysis of INTERV1 FMOC Synthesis

	Experimental	Predicted
Amino Acid Composition	Asp: 12.5/18[1] Ser: 5/7 Tyr: 0.66/1	Asp: 15 Ser: 5 Tyr: 1
Sequence	DDDDDDYSDSDXXXXDDDDDD[2]	DDDDDDYSDSDSSDSDDDDDD
Molecular Weight (a.m.u.)	2661.5[3]	2341.66

[1]Normalized value, relative to Tyr residue.

[2]X denotes uncertainty in sequence; may be either Ser or Asp.

[3]Determined by biomass spectrometry. Broad peak centered at 2661.5 amu, ± 400 amu

Figure 1: *500.138 MHz ^1H NMR spectra of INTERV1 in D_2O, pH 7.08, at low ionic strength.*

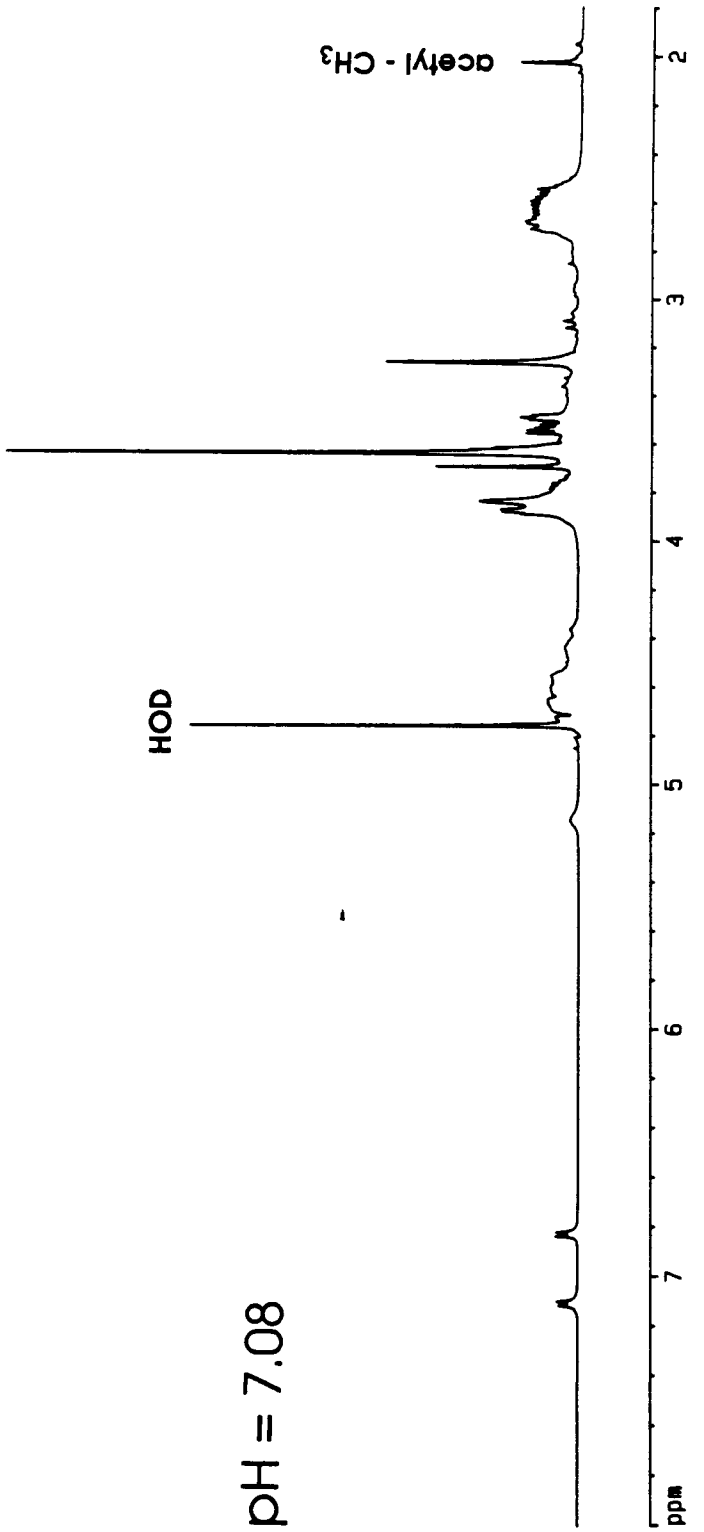

pH = 7.08
HOD
acetyl - CH3
ppm
7
6
5
4
3
2

Figure 2: *"Clean" ^1H - ^1H NMR HOHAHA spin-lock correlation spectroscopy of the INTERV1 amide fingerprint region.* A 7.8 mM INTERV1 sample in 90% H_2O/10% D_2O, at pH 7.1, was utilized in this experiment. ^1H NMR chemical shifts are referenced from 0.1 M TSS.

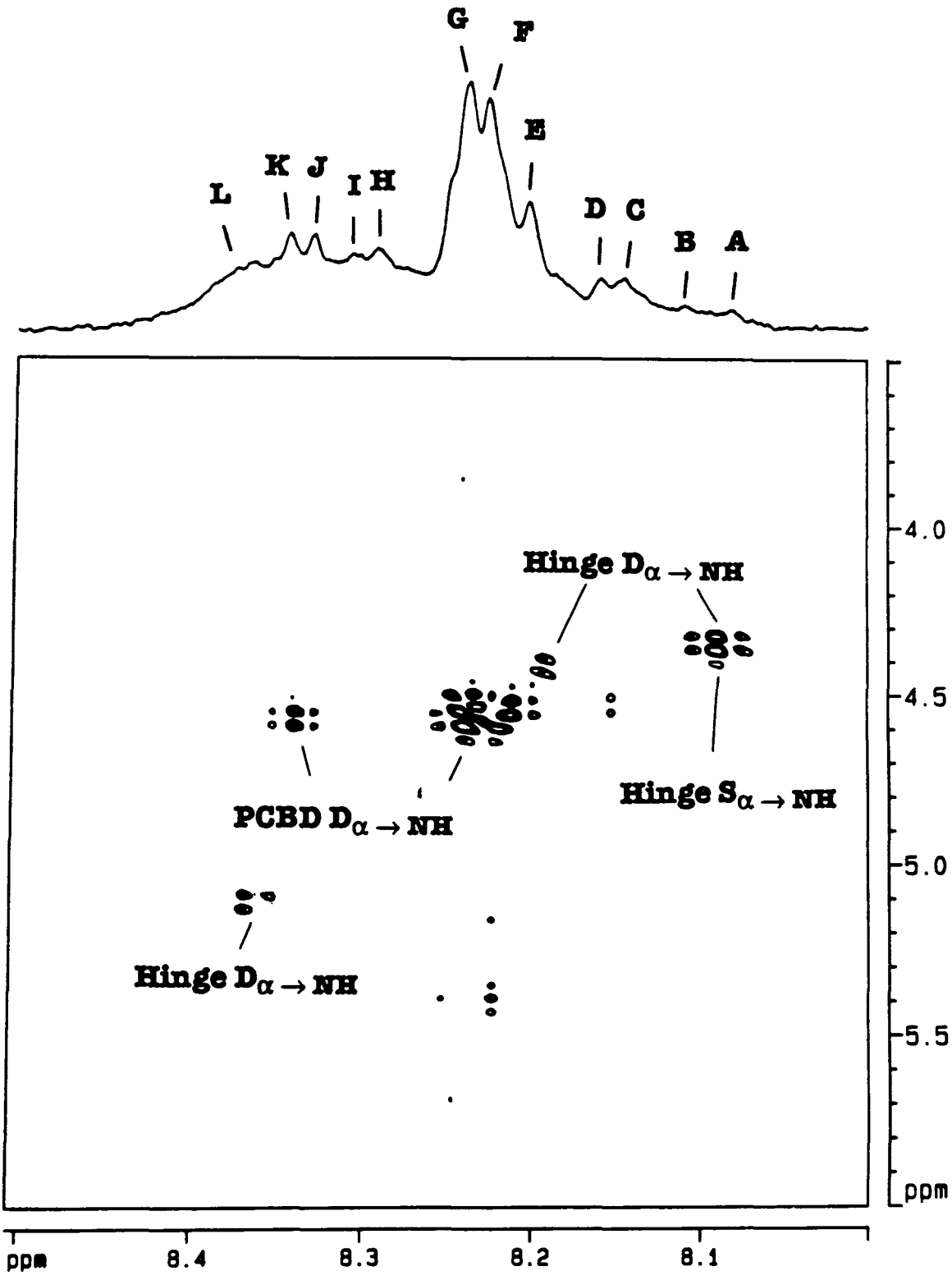

G
F
E
L
K
J
I
H
D
C
B
A
Hinge $D_\alpha \rightarrow NH$
PCBD $D_\alpha \rightarrow NH$
Hinge $S_\alpha \rightarrow NH$
Hinge $D_\alpha \rightarrow NH$
4.0
4.5
5.0
5.5
ppm
ppm
8.4
8.3
8.2
8.1

Figure 3: *"Clean" ^{1}H - ^{1}H NMR HOHAHA (500.138 MHz) spin-lock correlation spectroscopy of α-CH fingerprint region of INTERV1 (1.2 mM) in D_2O at* (**A**) pH 7.08, low ionic strength, divalent-cation depleted (**B**) as in A, but at pH 2.15, (**C**) pH 7.16, in the presence of 17.5 mM ^{113}Cd (II)

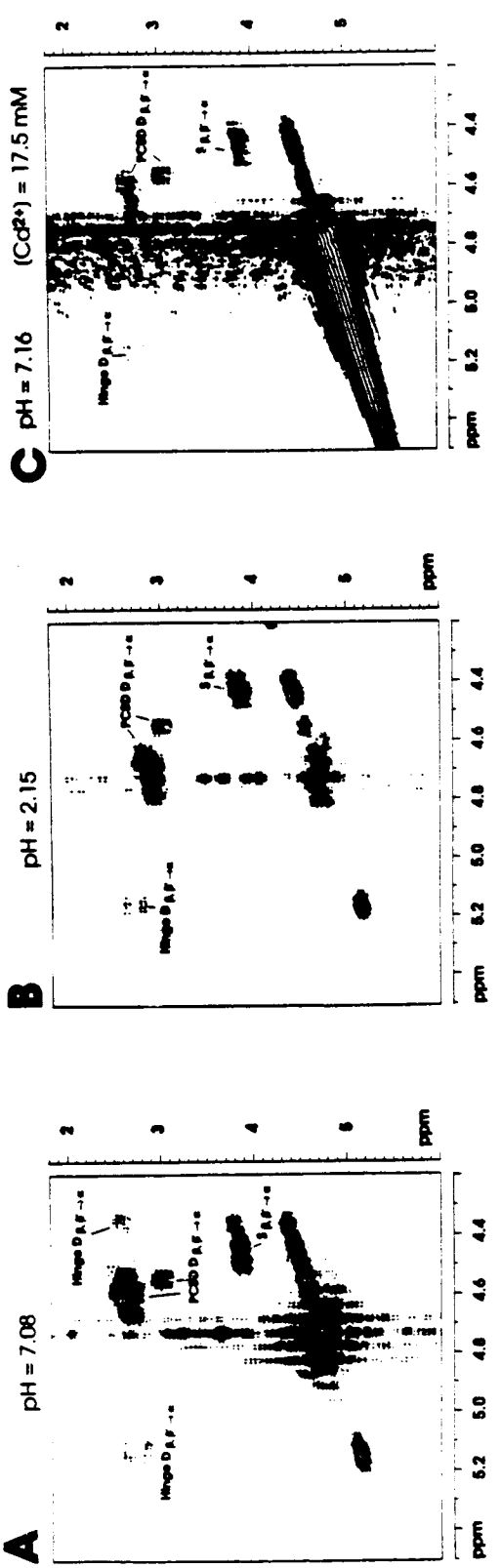

A
pH = 7.08
B
pH = 2.15
C
pH = 7.16
(Cd2+) = 17.5 mM
ppm

Table II: *^1H NMR amino acid spin assignments for INTERV1.* All ^1H chemical shifts are reported from external 0.1 M TSS in D_2O.

Table II

^{1}H NMR Amino Acid Spin Assignments for Non-ExchangeableProtons in INTERV1, pH 7.08, in D_2O

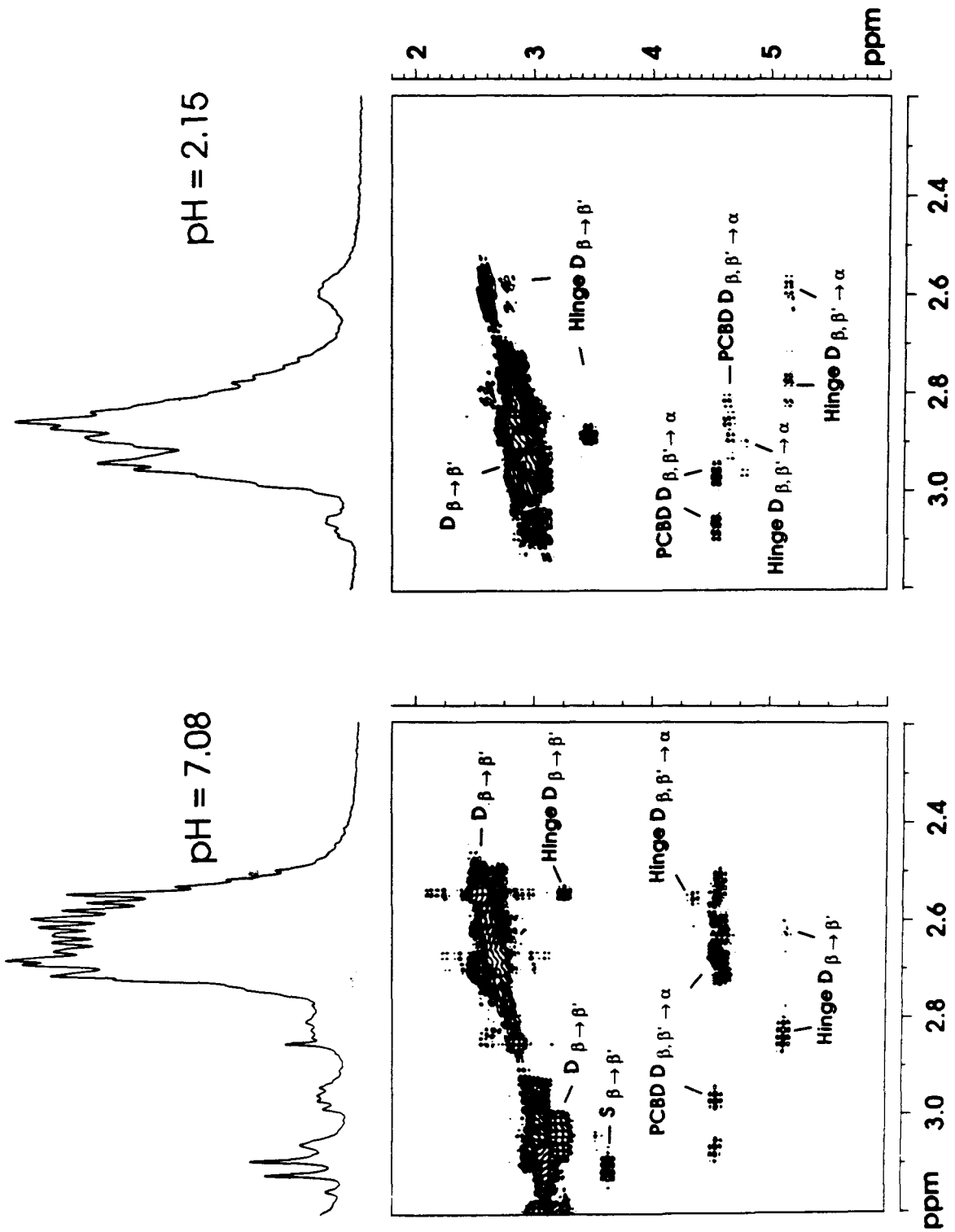
Peak	δ, ppm	Assignment	Peak	δ, ppm	Assignment	Peak	δ, ppm	Assignment
methyl	2.027	acetyl-CH_3	21	2.855	H $D_{\beta,\beta'}$	42	3.872	$S_{\beta,\beta'}$
1	2.499	$D_{\beta,\beta'}$	22	2.907	$D_{\beta,\beta'}$	43	3.881	$S_{\beta,\beta'}$
2	2.505	$D_{\beta,\beta'}$	23	2.943	$D_{\beta,\beta'}$	44	3.890	$S_{\beta,\beta'}$
3	2.517	$D_{\beta,\beta'}$	24	2.959	$D_{\beta,\beta'}$	45	4.364	α-CH
4	2.531	$D_{\beta,\beta'}$	25	2.972	$D_{\beta,\beta'}$	46	4.439	α-CH
5	2.546	H $D_{\beta,\beta'}$	26	2.988	$D_{\beta,\beta'}$, $Y7_{\beta,\beta'}$	47	4.556	α-CH
6	2.563	$D_{\beta,\beta'}$	27	3.061	$S_{\beta,\beta'}$	48	4.655	α-CH
7	2.579	$D_{\beta,\beta'}$	28	3.093	$S_{\beta,\beta'}$	50	5.136	H D_α
8	2.596	$D_{\beta,\beta'}$	29	3.123	$S_{\beta,\beta'}$	51	6.821	Y7
							6.837	C3H, C5H
9	2.613	$D_{\beta,\beta'}$	30	3.153	$S_{\beta,\beta'}$	52	7.100	Y7
							7.116	C2H, C6H
10	2.629	$D_{\beta,\beta'}$	31	3.265	H $D_{\beta,\beta'}$			
11	2.643	H $D_{\beta,\beta'}$	32	3.329	$S_{\beta,\beta'}$			
12	2.662	$D_{\beta,\beta'}$	33	3.365	$S_{\beta,\beta'}$			
13	2.679	$D_{\beta,\beta'}$	34	3.399	$S_{\beta,\beta'}$			
14	2.688	$D_{\beta,\beta'}$	35	3.493	$S_{\beta,\beta'}$			
15	2.712	$D_{\beta,\beta'}$	36	3.535	$S_{\beta,\beta'}$			
16	2.720	$D_{\beta,\beta'}$	37	3.626	$S_{\beta,\beta'}$			
17	2.742	$D_{\beta,\beta'}$	38	3.697	$S_{\beta,\beta'}$			
18	2.810	H $D_{\beta,\beta'}$	39	3.763	$S_{\beta,\beta'}$			
19	2.828	H $D_{\beta,\beta'}$	40	3.839	$S_{\beta,\beta'}$			
20	2.841	H $D_{\beta,\beta'}$	41	3.881	$S_{\beta,\beta'}$			

Table II

^1H NMR Amino Acid Spin Assignments for Exchangeable Protons in INTERV1, pH 7.08, in H_2O

Peak	δ, ppm	Assignment
A	8.081	Hinge Asp/Ser-NH
B	8.109	Hinge Asp/Ser-NH
C	8.143	NH (?)
D	8.159	NH (?)
E	8.201	Hinge Asp-NH
F	8.224	PCBD Asp-NH
G	8.235	PCBD Asp-NH
H	8.290	NH (?)
I	8.304	NH (?)
J	8.327	PCBD Asp-NH
K	8.341	PCBD Asp-NH
L	8.362	Hinge Asp-NH

Figure 4: *"Clean" ^{1}H-^{1}H NMR HOHAHA (500.138 MHz) spin-lock correlation spectroscopy of INTERV1 (1.2 mM) β-CH_2 fingerprint region at* (**A**) pH 7.08 and (**B**) pH 2.15 in D_2O, low ionic strength.


pH = 7.08
pH = 2.15
D β→β'
Hinge D β→β'
S β→β'
PCBD D β,β'→α
Hinge D β,β'→α
PCBD D β,β'→α
ppm
3.0
2.8
2.6
2.4
2
3
4
5

Ten

Summary

Bovine Dentine Phosphophoryn (BDPP) is a member of the "Asp-rich" superfamily of template macromolecules known as polyelectrolyte mineral matrix proteins (PMMPs). In the preceding chapters, a substantial amount of information regarding BDPP conformational folding in solution has been obtained, which supports the molecular complementarity theory. The main points are the following:

I. BDPP sequence organization can be conceptualized as three types of domains: *polyelectrolyte calcium binding domains* (**PCBD**), *hinge domains* (**HD**), and the *hydrophobic domains* (**HC**). The PCBD regions can be further subdivided into $(Asp)_n$, $(PSer)_n$, and $(PSerAsp)_n$-containing homopolymer and heteropolymer sequence stretches. The HD contain predominantly neutral or uncharged amino acids such as Ser, Gly, and Pro. These HD domain sequences are flanked on either side by PCBD regions.

II. BDPP contains a number of Lys sidechains (44/1000 residues) which are believed to form ion pairs with either Asp, Glu, or PSer residues in the protein. The function of this salt-bridging is not understood at this time, but it may be responsible for maintaining the protein template molecule in a conformation that retains a high negative charge density.

III. At low pH, or in the presence of divalent cations, BDPP assumes a global conformation that is condensed in particle size. At neutral pH under conditions of low ionic strength and in the absence of divalent cations, this global conformation converts to an extended form. The folding transition between these two conformers is mediated by conformational change in the HD regions which are flanked by specific PCBD sequence regions.

IV. The PCBD regions possess some degree of tertiary and secondary structural organization in the absence of divalent cations at low ionic strength. This folding permits the surface charge density of these regions to remain high, relative to that of a random coil conformer.

V. BDPP exhibits a selectivity in terms of divalent cation binding sites. Under conditions of low ionic strength and divalent cation depletion, the addition of Cd (II) to BDPP leads to binding at various PCBD sequence stretches, according to the following order:

$$(PSer)_n > (PserAsp)_n > (Asp)_n$$

VI. Modeling studies conducted on PCBD sequence peptides [$(Asp)_{20}$, $(PSer)_{20}$, and $(PSerAsp)_{10}$] indicate that there is a sequence preference for certain conformers in the presence of Na^+, i.e., $(Asp)_{20}$ forms "supercoils", $(PSer)_{20}$ forms "hairpins", and $(PSerAsp)_{10}$ forms "spirals" or "distorted" hairpins. Each of these conformers features some degree of sidechain folding and/or peptide backbone secondary structure, in support of Mann's hypothesis, as well as the experimental data obtained for BDPP.

VII. A peptide mimetic which represents a PCBD-Hinge-PCBD motif of rat α-phosphophoryn, was constructed using FMOC solid-phase peptide synthesis. Using 2-D NMR spectroscopy and monitoring α-CH and sidechain β-CH_2 proton chemical shifts, we can demonstrate that this peptide mimetic folds or collapses under conditions of low pH and divalent cation addition, in a manner similar to that observed for intact BDPP.