The Ice-Binding Site of Sea Raven Antifreeze Protein Is Distinct from the Carbohydrate-Binding Site of the Homologous C-Type Lectin†

Michèle C. Loewen,[‡,§] Wolfram Gronwald,[||,⊥] Frank D. Sönnichsen,[@] Brian D. Sykes,[||] and Peter L. Davies*[,‡]

Department of Biochemistry, Queen's University, Kingston, Ontario K7L 3N6, Canada, Protein Engineering Network Centers of Excellence and Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2S2, Canada, and Department of Physiology and Biophysics, Case Western Reserve University, Cleveland, Ohio 44106-4970

Received August 24, 1998; Revised Manuscript Received October 26, 1998

ABSTRACT: Antifreeze proteins lower the freezing point of their solution by binding to ice and inhibiting its growth. One of several structurally different antifreeze proteins in fishes (type II) is homologous to the carbohydrate-recognition domain of Ca^{2+}-dependent lectins and adopts the same three-dimensional fold. Type II antifreeze proteins from herring and smelt require Ca^{2+} for binding to ice, whereas this same antifreeze protein in sea raven binds to ice in the absence of Ca^{2+} and has only two of the five Ca^{2+}-liganding amino acids that are present in the lectin. To locate the ice-binding site, site-directed mutants of the 15 kDa, globular, disulfide-bonded sea raven antifreeze protein were produced by secretion from *Pichia pastoris*. Pairs of amino acid replacements, insertions, and a peptide loop swap were made in the region equivalent to the sugar-binding site of the lectin that encompasses loops 3 and 4 and β-sheets 7 and 8. Even the most extensive mutation caused only a 25% decrease in antifreeze activity and demonstrated that the residues corresponding to the Ca^{2+}-binding site are only peripherally involved in ice binding. When adjacent surface residues were mutated, the replacement of one residue, Ser120 by His, caused a 35% decrease in activity by itself and an 80% loss in conjunction with the peptide loop swap mutation. This pivotal sea raven antifreeze protein amino acid does not coincide with the herring ice-binding epicenter, but is located within the region corresponding to the proposed $CaCO_3$-binding surface of a third homologue, the pancreatic stone protein. Intron and exon structure of the sea raven AFP gene also suggests that it might be more closely related to the stone protein gene than to the lectin gene. These results support the notion that this family of proteins has evolved more than one binding surface from the same protein scaffold.

Sea raven type II antifreeze protein (SRAFP)[1] is a member of a growing class of proteins, known to protect isothermic organisms from freezing to death at sub-0 °C temperatures (*1*, *2*). These antifreeze proteins (AFP) function through an adsorption/inhibition mechanism, in which proteins bind to specific ice surfaces and inhibit their growth (*3*). According to the Kelvin effect, the bound AFPs cause the ice between the protein molecules to grow with an increased radius of curvature and, therefore, with a higher surface free energy (*4*, *5*). The resulting inhibition of the energetically favorable *a*-axis ice growth causes formation of static hexagonal bipyramidal ice crystals (*6*). The exact shape of the hexagonal bipyramid is dependent upon the type of AFP bound and reflects the AFP's affinity for different ice surfaces (*7*). The overall effect is depression of the nonequilibrium freezing point below the melting point. This difference is termed thermal hysteresis and is a function of AFP concentration.

AFPs present a structurally diverse class of proteins, of which the fish AFPs (types I−IV) have been the most widely studied (*8*). Solution and/or crystal structures are currently available for AFP types I, II, and III. They demonstrate no sequence or structural homology. Type I is comprised of a single α-helix (*9*). Type II is a 14 kDa globular protein with a mixed α/β structure (*10*). Type III is a 7 kDa compact angular protein, comprised primarily of short β-strands (*11*, *12*). The common antifreeze activity of these proteins can be reconciled with structural diversity through their ability to each bind a different plane of ice (*13*). Extensive mutagenic work has been carried out on both types I (*14*, *15*) and III (*16*, *17*), leading to a number of models depicting possible interactions at the ice−protein interface (*9*, *11*, *12*, *18*, *19*). These interactions are thought to be mediated by a variety of forces, including hydrogen-bonding networks, hydrophobic burial, and a specific contour fit leading to ideal

† This work was supported by grants from the Medical Research Council of Canada to P.L.D. and B.D.S., as well as by an MRC studentship to M.C.L.

* To whom correspondence should be addressed: Department of Biochemistry, Queen's University, Kingston, ON K7L 3N6, Canada. Telephone: (613) 545-2983. Fax: (613) 545-2497. E-mail: daviesp@post.queensu.ca.

‡ Queen's University.

§ Present address: Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139-4307.

|| University of Alberta.

⊥ Present address: Universitaet Regensburg, Institut fuer Biophysik und Physikalische Biochemie Postfach, 93040 Regensburg, Germany.

@ Case Western Reserve University.

[1] Abbreviations: SRAFP, sea raven antifreeze protein; rSRAFP, recombinant sea raven antifreeze protein; AFP, antifreeze protein; CRD, carbohydrate-recognition domain; MBP, mannose-binding protein; PSP, pancreatic stone protein; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis.

10.1021/bi9820513 CCC: $15.00
Published on Web 12/03/1998

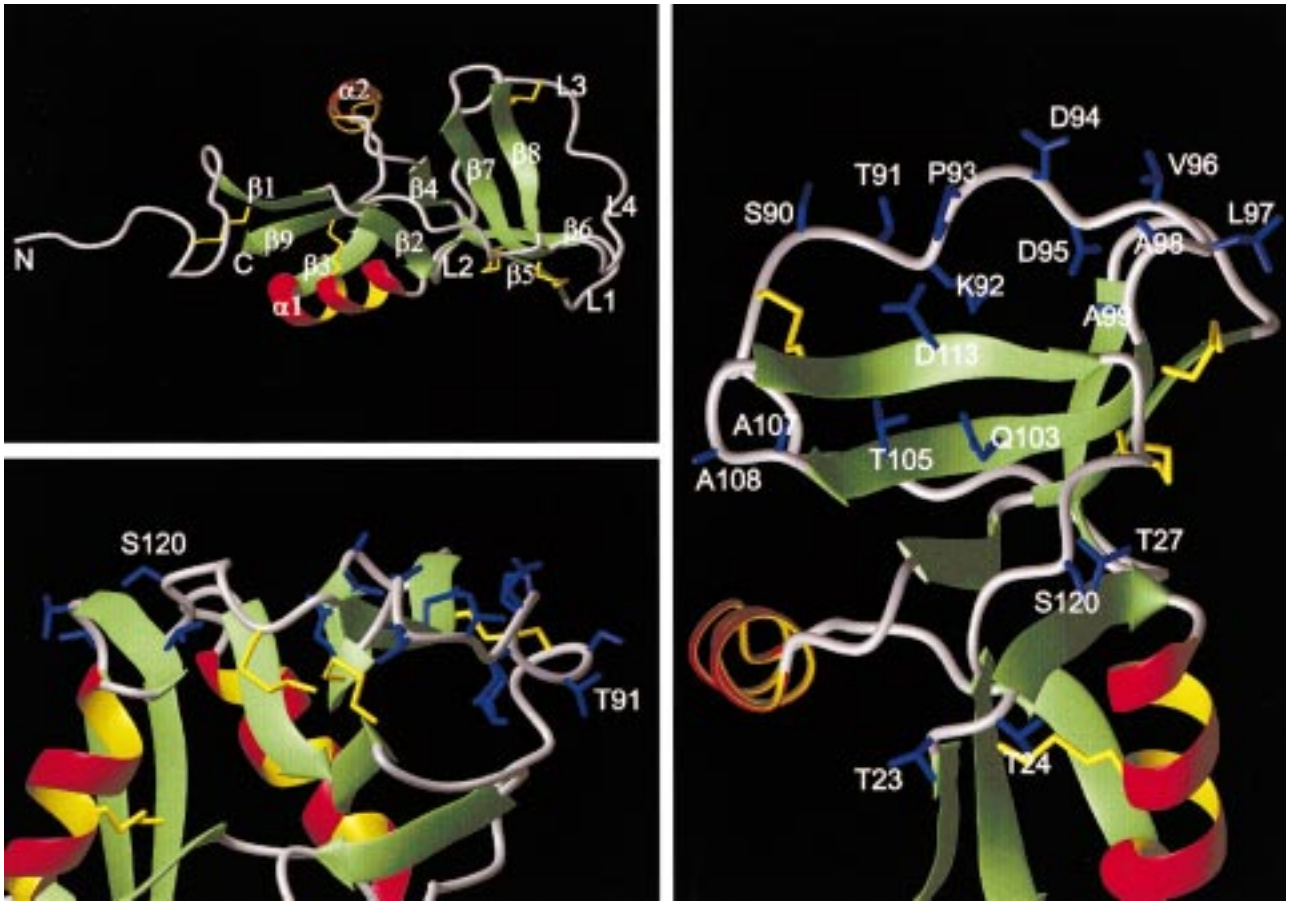

FIGURE 1: Views of the 3D NMR structure of recombinant sea raven antifreeze protein. (Top left) Backbone representation with secondary structure elements highlighted as follows. β-Strands are green. α-Helices are red and yellow. Loop regions and random chain are gray. Disulfide bridges are yellow, and the N and C termini are indicated with N and C, respectively. The region corresponding to the CRD Ca^{2+}-binding site includes β-strands 7 and 8 and loops L3 and L4. (Right) Presentation of the portion of the SRAFP NMR structure which corresponds to the Ca^{2+}-binding site of the CRD of C-type lectins. Side chain conformations were not very well defined by the NMR structure (*10*). This structure is color coded as described above, but with side chains in blue. (Bottom left) Same as in the right panel, but the molecule is rotated to display the flatness of the region corresponding to the Ca^{2+}-binding site of the CRD of C-type lectins as well as the separation between the T91 and S120 regions.

van der Waals contacts. The relative contribution of these different interactive forces and the existence of a common mechanism of binding for these structurally diverse proteins remains to be determined.

Type II AFPs have been isolated from three different fish: sea raven (*20*), herring (*21*), and smelt (*22*). The sea raven isoform is 40% identical to the herring and smelt isoforms, which are 83% identical to one another. The earliest indications of structure−function relationships for the type II AFPs came from their observed homology (up to 30% identity) to the lectin superfamily (*21*−*23*). This family includes the carbohydrate-recognition domains (CRD) of Ca^{2+}-dependent (C-type) lectins, which bind sugar moieties in a Ca^{2+}-dependent manner (*24*), and pancreatic stone proteins (PSP), thought to interact with $CaCO_3$ in some manner (*25*).

Three-dimensional (3D) structures of sea raven (*26*) and herring (*27*) AFPs have been modeled from the X-ray crystal structure of the CRD of the rat mannose-binding protein (MBP) (*24*). A solution structure of SRAFP was recently determined by NMR (*10*) which confirmed its similarity to the lectin protein fold. The SRAFP structure shares the conserved disulfide bridges, the two α-helices, a similar β-strand structure, and the extensive loop regions (Figure 1). The NMR structure also confirmed the additional disulfidebridges proposed by the model (*26*). Unfortunately, the NMR structure provides little additional information on side chain positioning which is critical for determining the basis of the match between the {11−21} plane of ice and SRAFP (*13*).

There is a potentially important distinction to be made between SRAFP and the Ca^{2+}-dependent type II AFPs of herring and smelt. Specifically, the smelt and herring AFPs conserve four of the five Ca^{2+}-binding residues of the MBP (*21*) and show metal ion (Mn^{2+}, Ba^{2+}, Zn^{2+}, and Ca^{2+})-dependent structure and activity (*27*). These observations strongly imply a role for this Ca^{2+}-binding region [corresponding to β7 and β8 and surrounding loops 3 and 4 in the SRAFP structure (Figure 1)] in ice binding, analogous to sugar binding in the C-type lectins. In fact, a recent mutagenic study of herring AFP conclusively demonstrated that amino acid residues of the Ca^{2+}-binding site are the primary ice-binding ligands (*28*). While SRAFP conserves only two of the five MBP Ca^{2+}-binding residues and does not require Ca^{2+} for activity (*21*, *22*, *26*), the evolutionary link between these proteins makes the region a logical candidate for the ice-binding site. As well, the overall flat (Figure 1) and hydrophilic nature of this corresponding SRAFP surface led to its proposal as a putative ice-binding site (*26*).

This study probes the mechanistic and evolutionary relationships between type II AFPs and the lectin superfamily. Initially, we characterized purified recombinant SRAFP secreted from *Pichia pastoris*. Subsequently, we made a series of site-specific and domain-swap mutants in the general region corresponding to the CRD Ca^{2+}-binding site to determine which amino acids mediate SRAFP ice binding. The effect of the mutagenesis on AFP activity and structure was determined by thermal hysteresis measurements, observation of altered ice crystal morphology, and NMR analysis. Our results suggest that the epicenter of the SRAFP ice-binding surface is situated around residue Ser120, which is distinctly different from the herring isoform ice-binding center (*28*). The significance of these results is discussed with respect to SRAFP's evolutionary relationship to CRDs, PSPs, and herring and smelt type II AFP isoforms.

EXPERIMENTAL PROCEDURES

Production of Recombinant SRAFP. SRAFP cDNA, including a six-amino acid C-terminal His tag, was cloned into the pPIC9 vector (Invitrogen, San Diego, CA) immediately downstream of the yeast α-mating factor signal sequence, where it was under the control of the yeast alcohol oxidase (AOX1) promoter. The resulting expression construct, pPIC9-SRm-CTHT (described in ref *29*), was linearized with *Bgl*II and integrated into the *P. pastoris* genome by homologous recombination at the AOX1 gene site. This produced clones with a His^+, "methanol utilization slow" (Mut^s) phenotype. The linearized vector was introduced into *P. pastoris*, either using the spheroplast transformation procedure (*30*) or by electroporation (*31*). Positive transformants were grown in 1 L cultures of buffered minimal glycerol-complex medium (BMGY) at 30 °C for 48 h. BMGY contained the following (per liter): yeast extract (10 g), meat peptone (20 g) (Sigma, St. Louis, MO), yeast nitrogen base without amino acids (13.4 g) (Difco, St. Louis, MO), biotin (400 μg) (Sigma), and glycerol (10 mL) in 100 mM potassium phosphate buffer (pH 6.0). The glycerol in this medium inhibited the AOX1 promoter. Cells were harvested by centrifugation and resuspended in the same volume of buffered minimal methanol-complex medium (BMMY), which contained methanol for induction of the AOX1 promoter, and again incubated for 48 h at 30 °C. BMMY contained the same components as BMGY except that glycerol was replaced by methanol (5 mL/L). Cells were removed by centrifugation, leaving recombinant SRAFP to be purified from the medium.

Purification and Analysis of Recombinant SRAFP. The yeast medium containing recombinant SRAFP (rSRAFP) was loaded directly onto a nickel−agarose column (10 mm × 50 mm) (Qiagen, Chatsworth, CA). The loaded column was washed extensively with buffer N [20 mM Tris-HCl (pH 7.6), 0.5 M NaCl, 2% glycerol, and 5 mM imidazole] prior to elution of rSRAFP by an imidazole gradient (5 to 125 mM) over the course of 75 min at a flow rate of 1.5 mL/min. Fractions containing rSRAFP were pooled and dialyzed against exchange buffer [25 mM Tris-HCl (pH 9.0)] and were subsequently chromatographed on an FPLC anion exchange column [Q-Sepharose 16/10 (16 mm × 100 mm); Pharmacia, Uppsala, Sweden] pre-equilibrated in exchange buffer. Bound AFP was eluted by a linear 0.2 to 0.4 M NaCl gradient in exchange buffer at a flow rate of 2.5 mL/min. Fractions containing AFP were pooled and dialyzed into 100 mM NH_4HCO_3 and lyophilized.

Proteins were analyzed by SDS−PAGE on 17% polyacrylamide gels, containing 0.1 M sodium phosphate and 4 M urea at pH 6.8. These gels either were stained with Coomassie blue (0.25%) and amido black (0.1%) and then destained (10% methanol and 10% ethanol) or were Western blotted onto a nylon membrane. The membrane was incubated with rabbit anti-SRAFP antiserum and then horseradish peroxidase-linked goat anti-rabbit IgG (Promega, Madison, WI). Detection was carried out using enhanced chemiluminescence (Amersham, Uppsala, Sweden). AFP amounts were estimated from Western blot analysis by visual inspection and comparison of band intensities generated by known quantities of SRAFP. Dry weight, A_{280nm} [using the predetermined standard for SRAFP of 4.5 mg/mL = 10.0 $_{OD}$)] and Bradford protein detection (*32*) were also used to quantify purified AFP.

One-Dimensional (1D) ^1H NMR. Freeze-dried protein was dissolved in 500 μL of 90% H_2O/10% D_2O (v/v) to give 0.5−1.0 mM protein solutions. In addition, 10 mM imidazole was added as a buffer and 2,2-dimethyl-2-silapentane-5-sulfonic acid was used for internal referencing. The pH of the samples was adjusted to 7.0. NMR experiments were carried out at 25 °C on a Varian Unity 500 MHz NMR spectrometer. Ca^{2+} titrations were carried out by 10 stepwise additions of $CaCl_2$ (4 μL of a 108 mM solution), increasing the Ca^{2+} concentration by 0.85 mM in each step. The starting Ca^{2+} concentration was 0 mM, and the final Ca^{2+} concentration was 7.65 mM.

Antifreeze Activity Measurements and Photomicroscopy. Thermal hysteresis was measured by placing nanoliter volumes (0.1−0.5 μL) of an AFP solution within oil droplets on a cooling stage beneath a microscope for observation of ice crystals. The temperature of the stage was controlled by a nanoliter osmometer (Clifton Technical Physics, Hartford, NY) (*33*). The growth of a single ice crystal in the presence of AFP was observed as the temperature is lowered from the melting point (where ice and its melt are at equilibrium) to the freezing point (rapid ice growth and/or burst). Some AFP mutants permitted slow growth of the ice crystal prior to the freezing point. For the purposes of this study, slow growth was ignored and the assay was continued until a rapid ice burst was observed. The measured difference between the melting point and freezing point is termed thermal hysteresis. All measurements were made in 100 mM NH_4HCO_3. Ice crystal morphology was observed through a Leitz Dialux 22 microscope and recorded by a Panasonic CCTV camera linked to a JVC Super VHS video recorder. Still images were obtained from a Silicon Graphics INDY terminal using IRIS Capture version 1.2. For ice growth rate analysis, samples were held at a fixed degree of undercooling and images were captured at 0 and 10 min time points.

Primer-Directed Mutagenesis. All mutants were made by primer-directed mutagenesis (*34*). Single-stranded uracil-containing template from either the pT7-7F-SRm or pET20b+-SRm vectors (*29*) was produced using the helper phage M13KO7 and was recovered by standard procedures (*35*). All mutagenic primers were designed to have DNA lengths on either side of the missmatch, which gave a T_{ms} of at least 40 °C. These are listed in Table 1. Positive mutant clones

Table 1: Antisense Mutagenic Primers

Mutation	Sequence[a]
L-Swap[b]	pCGTTCCTGGTGTTCTGACCAACCTAACGATCACGGTTCCGGTGAATGCTGTATGCAGATG
S90A	pATCATCAGGTTTGGTAGCACACCAGGAACGAAA
D113A	pGGACAAGGCAAGTCAGCCCAGCATTGGTC
T105A/Q103A	pGGTCAGCTGCAGCAGCCATAGCCATACAGCACGCG
T23N/T24N/T27A	pGCCAGAGCCCAAGCCATCGCATTATTCTCATAATAGATACAGC
S120H	pGCAGACTGATTTGTGATGTGCCGGACAAGGCAA
T23H	pCAAGTCATCGCTGTATGCTCATAATAGATACAG
T24H	pCCCAAGTCATCGCATGTGTCTCATAATAGATA
Inserts:	
94/95	pCGGCCAGTACCATCACCACCTTGGTACCCACCATCAGGTTTGGTAGA
107/108	pCAGCATTGGTCAGCACCACCTTGGTACCCACCTGCAGCAGTCATCT

[a] All sequences written 5′ to 3′. p is a site of phosphorylation. [b] Sense primer. Sea raven AFP nucleotide sequence located in Genbank under accession numbers J05100 and J02593.

identified by dideoxy sequencing (*36*) were digested with *Bst*EII and *Not*I and subcloned in the pPIC9-SRm-CTHT vector for transformation into yeast and expression. Mutants that included two single-site changes at different positions along the cDNA were mutated by the addition of two mutagenic primers to the same mutagenic reaction. Other double mutants where a single site change was made in addition to the loop-swap mutation were produced using ssDNA of the pET20b+-SRm-CTHT-L-Swap mutant.

Alignment of Genomic Structures. The intron and/or exon boundaries of the SRAFP genomic sequence were previously determined by comparison with SRAFP cDNA (*37*). The amino acid sequences of all related proteins, including SRAFP, were aligned using the conserved WIGL sequence (*38*).

RESULTS

Expression, Purification, and Characterization of Recombinant SRAFP. Using the methylotrophic yeast *P. pastoris* as a host, a recombinant mature form of SRAFP was secreted into the medium at a level of about 3 mg/L as estimated by Western blotting (Figure 2A). The rSRAFP was isolated by nickel−agarose affinity chromatography followed by Q-Sepharose ion-exchange chromatography with a final yield of 65%. When the purified protein was examined by staining after SDS−PAGE, it showed a single, broad band with an apparent M_r of ∼15000 (Figure 2B). N-Terminal Edman-degradation showed minor heterogeneity at the N terminus, with the majority of the rSRAFP lacking the first Gln residue. The 1D ^{1}H NMR solution spectrum of the rSRAFP was consistent with the NMR data obtained from SRAFP purified from sea raven serum (Figure 3) (*10*, *26*). The resonances were well dispersed over the entire ^{1}H resonance range, typical for folded proteins, and the spectra showed similar relative intensities. Of note was the ^{1}H resonance line at −0.8 ppm in both wild-type and recombinant spectra. This specific resonance line corresponds to the methyl group of Ala68. On average, the resonance line for alanine methyl groups in proteins and peptides appears around 1.4 ppm (*39*). This unusual shift of 2.2 ppm was due to the specific orientation

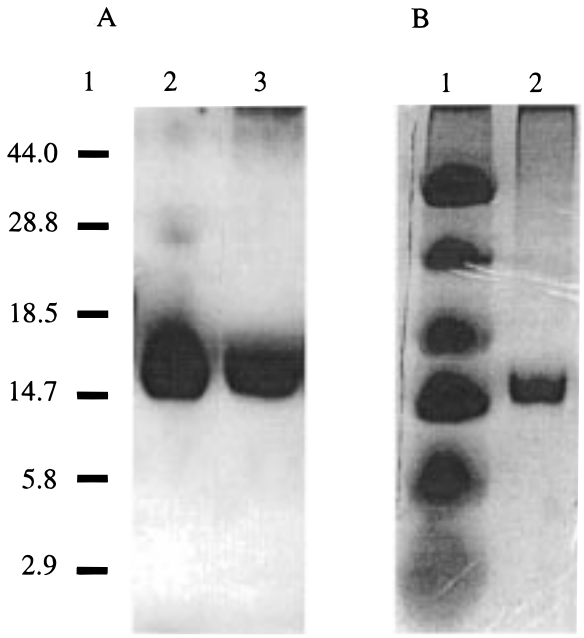

FIGURE 2: SDS−PAGE analysis of recombinant sea raven antifreeze protein. (A) Western blot: lane 1, molecular mass markers (kilodaltons); lane 2, wild-type SRAFP (∼200 ng) purified from sea raven serum as previously described (*20*); and lane 3, recombinant SRAFP in yeast media (20 μL). (B) Stained gel: lane 1, molecular mass markers (masses correspond to numbers in lane 1 of panel A) (kilodaltons); and lane 2, rSRAFP purified by nickel affinity and ion exchange chromatography (∼70 ng).

of the methyl group directly above the center of the aromatic ring of Trp75 in the hydrophobic core of the protein (*26*). The presence of this line in the recombinant spectra provided strong evidence of a fold identical to that of the wild type. The thermal hysteresis activity profile of the recombinant purified protein also matched that of the native protein, demonstrating that the C-terminal His tag addition did not affect antifreeze activity (*10*). Comparison of ice crystal morphology also indicated virtually identical activity, with both forms producing the same stubby, somewhat rounded hexagonal bipyramidal ice crystals (Table 2) (*10*, *29*).

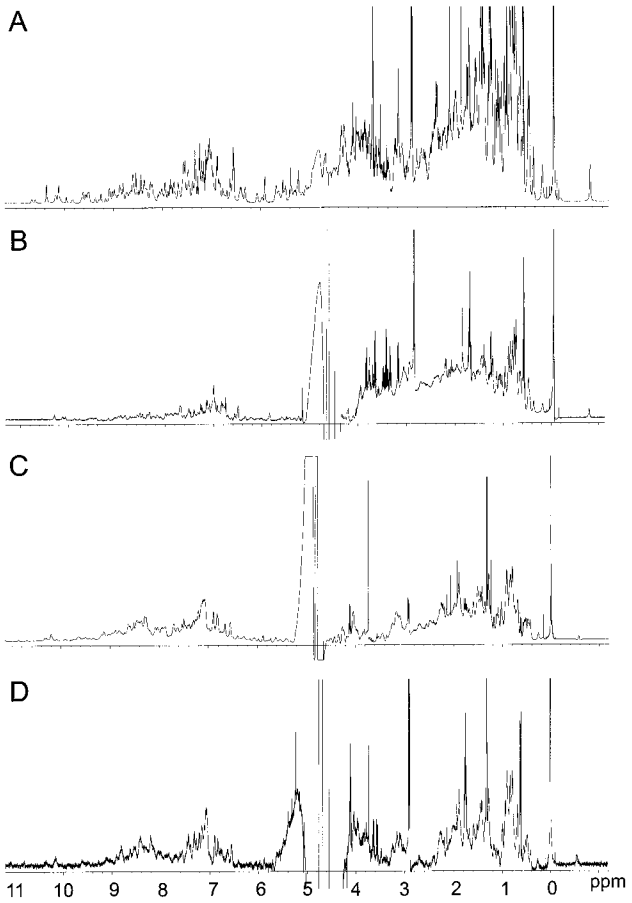

FIGURE 3: 1D ^{1}H NMR analysis of recombinant and mutant sea raven antifreeze proteins. 1D ^{1}H NMR spectra for (A) wild-type SRAFP purified from sea raven serum, (B) rSRAFP purified from yeast media, (C) rSRAFP L-Swap mutant purified from yeast media, and (D) rSRAFP S120H and L-Swap mutant also purified from yeast media.

Individual Replacement Mutations. The region of rSRAFP corresponding to the CRD Ca^{2+}-binding site includes strands β7 and β8 as well as adjacent loops-3 and 4 (Figure 1) (*26*). The right panel of Figure 1 shows an expansion of this region in the NMR structure and includes the positions of side chains. Initially, two double mutant proteins were made, T105A/Q103A and S90A/D113A. These four residues were chosen on the basis of their surface accessibility, hydrogen-bonding potential, and their proximity to the area corresponding to the CRD Ca^{2+}-binding region. Alanine was chosen as a benign substitute that would eliminate hydrogen-bonding potential. Alanine replacements had previously been used to identify ice-binding residues in type III AFP (T18A and T15A) (*12*, *40*). Surprisingly, thermal hysteresis values in the double mutants were virtually identical to those of rSRAFP (Figure 4). Only the T105A/Q103A mutant variants showed a slightly altered crystal morphology, which was more faceted, with better defined, angular edges, than those obtained with rSRAFP or the S90A/D113A mutant (Table 2). Neither mutant allowed any ice crystal growth over time. These results suggested that while T105 and Q103 might play some role in mediating AFP−ice interactions, the primary binding residues are located elsewhere.

Insertion Mutations. Because the double mutants had little effect on activity, more intrusive mutations were designed that involved peptide insertions. Two mutants, one with an insertion at positions 94 and 95 in loop 4 (94/95 insertion variant) and the second with an insertion at positions 107 and 108 between β-strands 7 and 8 (107/108 insertion variant) (Figure 1), were produced on the basis of the

Table 2: Thermal Hysteresis Activity and Ice Crystal Morphology of rSRAFP Mutants

* 1- minute time point — 0.1 mm

Mutation	% Thermal Hysteresis (at 3 mg/mL AFP)	Ice Crystal 0 min	Morphology 10 min
Wild Type	100		
rSRAFP	100		
T105A/Q103A	99		
S90A/D113A	100		
94/95 insert	72		
107/108 insert	85		
L-Swap	75		
S120H	65		
S120H+L-Swap	18		*
T23H	92		
T23H+L-Swap	56		

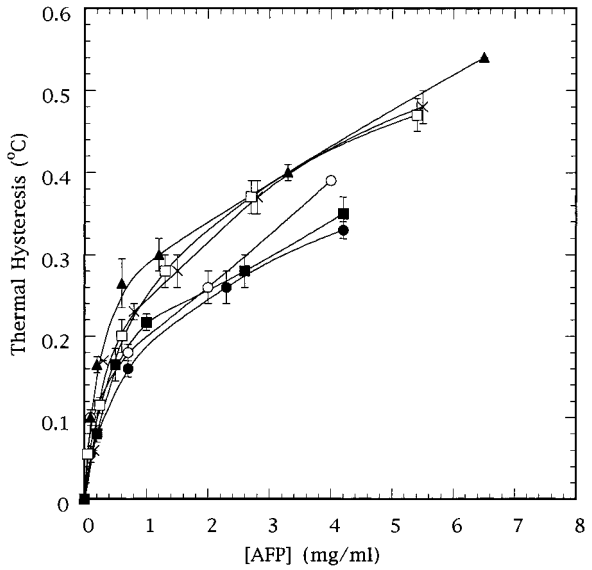

FIGURE 4: Thermal hysteresis activity of recombinant sea raven antifreeze protein and mutants. Thermal hysteresis measurements were carried out as described in Experimental Procedures. Measurements were made in triplicate (error bars represent the standard deviation of measurements) over a range of concentrations for each purified protein and plotted as follows: rSRAFP (▲), T105A/Q103A (×), S90A/D113A (□), 94/95 insert (●), 107/108 insert (○), and L-Swap (■).

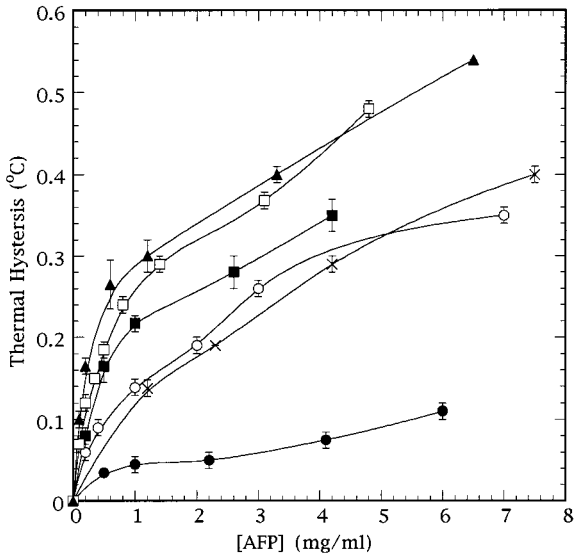

FIGURE 5: Thermal hysteresis activity of recombinant sea raven antifreeze protein and mutants. Thermal hysteresis measurements were carried out as described in Experimental Procedures. Measurements were made in triplicate (error bars represent the standard deviation of measurements) over a range of concentrations for each purified protein and plotted as follows: rSRAFP (▲), L-Swap (■), S120H and L-Swap (●), S120H (○), T23H (□), and T23H and L-Swap (×).

hypothesis that a peptide insertion near the ice-binding site might sterically inhibit interactions with ice without destabilizing the AFP structure. The insertion sequence was Gly-Gly-Tyr-Gln-Gly-Gly, where Gly's were chosen to allow a flexible link between the protein and the insert, and Tyr and Gln were chosen for their bulkiness and hydrophilic nature, the latter to minimize solubility problems. The exact sites of insertion on the AFP molecule were chosen on the basis of their proximity to the region corresponding to the CRD Ca^{2+}-binding site and their likelihood of structurally accepting an insertion (Figure 1). These mutant proteins were expressed, secreted, and purified at levels of >1 mg/L. The 94/95 insertion variant lost ~25% of the wild-type rSRAFP thermal hysteresis activity (Figure 4), and its ice crystal morphology showed more sharply defined facets, with slow growth over time (Table 2). The 107/108 insertion variant protein exhibited an only 15% loss of thermal hysteresis activity (Figure 4), and although the ice crystals were somewhat larger than wild-type crystals, they remained rounded and grew less over time than those formed with the 94/95 insertion variant (Table 2). These results suggested that the 94/95 and 107/108 locations are close to the ice-binding site.

Loop-Swap Mutation. Sönnichsen et al. (*26*) proposed that the SRAFP model deviates from the X-ray structure of the MBP in the orientation of the loop region (loop 4) preceding β7 (*26*). The corresponding loop 4 of the MBP structure is flipped up such that it produces a pocket in which the Ca^{2+} binds (*24*). A similar loop conformation was observed for the PSP (*41, 42*). In contrast, this loop region is flipped down in the SRAFP model similar to that described for the E-selectin CRD (*10, 43*), such that the surface is relatively planar, and potentially ideal for ice interactions (*26*), assuming that planarity is an essential feature (*44*). The NMR solution structure of rSRAFP confirmed that the loop was flipped down (*10*). The loop-swap mutation (L-Swap) is a domain swap of the nine amino acids (91-Thr-Lys-Pro-Asp-Asp-Val-Leu-Ala-Ala-99) making up loop 4 of the SRAFP structure (Figure 1), with the corresponding 10 residues in the MBP (Asp-Glu-Pro-Asn-Asp-His-Gly-Ser-Gly-Glu). This swap was designed to delete several potential ice-binding residues while causing a radical change in conformation (loop flip up) of the putative ice-binding surface. The yield of the L-Swap mutant protein from shake flask expression improved to 5 mg/L of yeast medium, compared to 3 mg/L for rSRAFP. 1D NMR of this mutant protein indicated no extensive alteration in protein fold compared to rSRAFP (Figure 3). A 25% decrease in thermal hysteresis activity compared to that of rSRAFP was observed (Figure 5), and ice crystal morphology showed an even more dramatic change from that obtained with the similarly located 94/95 insert. The characteristic small, rounded hexagonal bipyramid was replaced by a much more elongated, faceted crystal, which grew more rapidly over time (Table 2). Despite the re-introduction of four of the five MBP Ca^{2+}-binding residues (the ones present in herring and smelt type II AFPs), there was no significant change in activity, or in 1D NMR spectra, in the presence or absence of Ca^{2+} (data not shown).

Mutations of the Surrounding Surface. Because even a large domain swap had relatively modest effects on activity, areas surrounding the region corresponding to the CRD Ca^{2+}-binding site were probed. Specific sites were chosen for mutagenesis on the basis of the close relationship (as determined by gene structure) between SRAFP and the PSP (Figure 6 and Discussion). The putative $CaCO_3$-binding site of PSP is somewhat removed from the lectin Ca^{2+}-binding surface (*41*). Some of the PSP-binding site residues are also equivalent to the putative hyaluronate binding site of the proteoglycan tandem-repeat domain of Link protein, which maintains a similar overall fold (*45*). After extensive primary

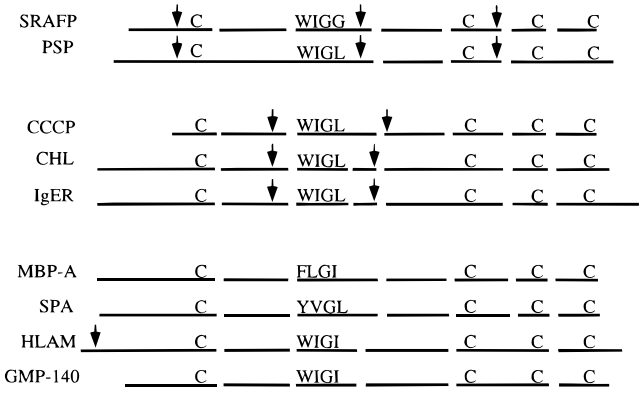

FIGURE 6: Comparison of genome structure for sea raven antifreeze protein and homologues. The amino acid sequences of SRAFP, PSP, and CRDs were aligned using the tetrapeptide WIGL (Trp-Ile-Gly-Leu) as a reference (described in ref *38*): SRAFP (*37*), rat pancreatic-associated protein (PSP) (*38*), core protein of chicken cartilage proteoglycan (CCCP (*49*), chicken hepatic lectin (CHL) (*50*), lymphocyte IgER Fc receptor (IgER) (*51*), human mannose binding protein (MBP-A) (*52*), pulmonary surfactant apoprotein (SPA) (*53*), human lymphocyte homing receptor (HLAM) (*54*), and platelet granule protein (GMP-140) (*55*). The arrows indicate intron positions. Gaps in the amino acid line represent gaps in the sequence alignment. The conserved Cys residues are indicated by C.

and tertiary structure comparisons, rSRAFP residues T23 and T24 (situated immediately prior to α-helix 1 and in close proximity to T27 within this helix) as well as S120 (localized to the loop region between β8 and β9) were tested (Figure 1).

A number of mutants were designed to probe these sites, including S120H+L-Swap, T23H+L-Swap, S120H, and T23H. Histidine was chosen for substitution because it is a large residue that might cause steric hindrance between the AFP and ice. Other mutations at positions T24 and T27 were also attempted, but no protein was recovered from the expression system for characterization. The S120H mutant protein retained 65% of wild-type activity (Figure 5) and produced a clearly faceted ice crystal that grew significantly over time (Table 2). This was by far the most deleterious single mutation made in SRAFP. When combined with the L-Swap mutation, thermal hysteresis dropped to ∼15−20% of wild-type thermal hysteresis activity (Figure 5). Ice crystal morphology was again less rounded and more angular, similar to the crystals formed by the L-Swap mutant protein and the S120H mutant. There was significant growth over time to the extent that the apexes of the bipyramid were sometimes not formed (Table 2). 1D ^{1}H NMR analysis indicated that the mutant protein maintained a fold identical to that of wild-type SRAFP (Figure 3). The T23H mutation by itself had little effect on either thermal hysteresis or ice crystal morphology (Table 2 and Figure 5). However, the T23H+L-Swap mutant protein exhibited a depressed level of thermal hysteresis in relation to both single mutations (Figure 5) and a significantly greater rate of ice crystal growth activity compared to the L-Swap mutant (Table 2).

DISCUSSION

This structure−function study was initially guided by the hypothesis that the ice-binding site of SRAFP would be close to the sugar-binding site of the CRD homologue and might involve equivalent residues. The failure of amino acid replacements to significantly reduce antifreeze activity was instructive in light of the marked activity losses seen with AFP types I and III when only one or two residues were replaced (*12*, *14*−*16*).

The 94/95 and 107/108 insertion probes were designed as radical steric mutations to test the involvement of local regions of the SRAFP in ice binding. Even modest steric mutations at the ice-binding site of type III AFP, in which amino acids with small side chains were substituted with bulkier ones (e.g., A16H and T18N), were particularly effective in reducing antifreeze activity to a small fraction of the wild-type activity (*16*, *46*). The failure of the 94/95 and 107/108 insertions to seriously disrupt ice binding raised the suspicion that the region on SRAFP corresponding to the CRD Ca^{2+}-binding site either is peripheral to the ice-binding region or mediates only a small part of the ice-binding interactions. This suspicion was confirmed by the L-Swap mutation. With this and the other mutations, four hydrophobic and seven hydrophilic residues in the region, including four of the CRD equivalent Ca^{2+}-binding residues and all the Thr and Ser residues, were mutated, resulting in a maximum 25% loss of thermal hystersis activity. It would seem that the SRAFP region corresponding to the CRD Ca^{2+}-binding site does not play the major role in ice binding, and therefore, we considered an alternative model.

A comparison of the genomic organization of SRAFP, PSP, and CRD suggested that SRAFP is more closely related to PSP than to the CRD (Figure 6). Furthermore, the proposed $CaCO_3$-binding site of PSP is comprised of residues outside of the Ca^{2+}-binding surface of the CRD (*41*), consistent with the hypothesis that the SRAFP ice-binding residues are not necessarily limited to the region corresponding to the CRD Ca^{2+}-binding site. Subsequently, residues of SRAFP corresponding to the proposed PSP $CaCO_3$-binding surface, which were surface accessible in the SRAFP model, were targeted for change. These mutants were initially made in the L-Swap mutant background so that the ice-binding activity would be at least partially inactivated, and visualization of the effect of the change would be more easily detected. The low level of activity (18%) observed for S120H+L-Swap mutant protein was the first indication of an ice-binding surface distinct from the CRD model. Subsequently, S120 was shown to be a pivotal side chain. It is the only residue where a single site mutation resulted in a significant loss of activity (>35% loss). While the T23H mutation by itself gave no indication of being involved in ice binding, the T23H+L-Swap mutant protein suggested that T23 mediates some ice-binding activity. Overall, these data demonstrate that residues outside of the region corresponding to the CRD Ca^{2+}-binding site are involved in mediating ice−protein interactions.

Even as the CRD model has proved to be inadequate in explaining SRAFP binding to ice, it has been validated as a good model for the ice-binding site of herring type II AFP. In contrast to SRAFP, herring and smelt AFPs have maintained the key Ca^{2+}-binding residues, and Ca^{2+} is required for their antifreeze activity (*21*, *22*, *27*). The ice crystal shape obtained with herring type II AFP is elongated and strongly faceted in contrast to the stubbier more rounded crystal produced by SRAFP. The central role for Ca^{2+} liganding in herring type II antifreeze activity was illustrated by a series of divalent metal ion substitutions (*27*). Not only did these change the level of antifreeze activity, but they

also had dramatic effects on the ice crystal morphology. Ca^{2+} was also shown to stabilize herring type II AFP against proteolysis.

The most convincing evidence for the similarity between the sugar-binding site of the CRD and the ice-binding site of herring AFP has come from a double mutant (Gln-Pro-Asp to Glu-Pro-Asn) in the latter protein. This mutant retained Ca^{2+} binding, but had lost all thermal hysteresis activity (*28*). The analogous mutation in the CRD of rat MBP is associated with the switch between galactose- and mannose-binding specificity. These mutations correspond to changes made in the SRAFP L-Swap mutation (including K92E and D94N producing the Glu-Pro-Asn CRD mannose-specific motif) which led to only a 25% loss of activity. These data demonstrate that these isoforms of type II AFP have different binding sites. The epicenter of the herring (and possibly the smelt) form(s) lies with the Ca^{2+}-binding site, while the sea raven AFP epicenter lies nearby S120. The later site corresponds more closely to the proposed $CaCO_3$-binding surface of PSP than to the CRD Ca^{2+}-binding site. This concept of different binding sites carrying out the same function in related proteins is not unique to this group of AFPs. Sculpin and flounder have related type I AFPs, but they contain different hydrophilic residues and have been shown to bind different ice planes (*13*). In one example from the literature, three "hydrophobic-molecule-binding proteins" with up to 90% amino acid identity and identical folds have very different ligand specificities and also use different surfaces of the molecules to bind the ligand (*47*).

From an alignment of the genomic structures of these related proteins and domains, it becomes evident that this family of homologous proteins can be separated into three evolutionary groups (Figure 6) (*38*). This alignment suggests that the SRAFP and the PSP are more closely related to each other than to the CRD. Unfortunately, the genome structures of the herring and smelt AFPs are not known. However a definite subdivision exists between these and SRAFP based on the Ca^{2+}, dependence and the low level of amino acid identity. The demonstration, herein, of a different ice-binding site for SRAFP further supports the independent evolution of SRAFP from a common ancestor shared with the CRD and herring and smelt AFP isoforms (*26*, *48*). Whereas sea raven, herring, and smelt type II AFPs are all derived from the superfamily of C-type lectin homologues , the timing of their speciation suggests that these fish developed their AFPs independently, and in the case of sea raven versus Ca^{2+}-dependent herring and smelt AFPs, from different members of the superfamily.

ACKNOWLEDGMENT

We gratefully acknowledge Dr. C. L. Hew, of the University of Toronto (Toronto, ON), for the gifts of partially purified SRAFP (from serum) and rabbit anti-SRAFP antiserum. We also thank Professor Fontecilla-Camps, of the Institut de Biologie Structurale J. B. Ebel in Grenoble (France), for early release of the PSP X-ray coordinates.

REFERENCES

1. Davies, P. L., and Hew, C. L. (1990) *FASEB J. 4*, 2460−2468.
2. Yeh, Y., and Feeney, R. E. (1996) *Chem. Rev. 96*, 601−617.
3. Raymond, J. A., and DeVries, A. L. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2589−2593.
4. Knight, C. A., Cheng, C. C., and DeVries, A. L. (1991) *Biophys. J. 59*, 409−418.
5. Wilson, P. W. (1993) *Cryo-Lett. 14*, 31−36.
6. DeVries, A. L. (1984) *Philos. Trans. R. Soc. London, Ser. B 304*, 575−588.
7. Chao, H., DeLuca, C. I., and Davies, P. L. (1995) *FEBS Lett. 357*, 183−186.
8. Davies, P. L., and Sykes, B., D. (1997) *Curr. Opin. Struct. Biol. 7*, 828−834.
9. Sicheri, F., and Yang, D. S. C. (1995) *Nature 375*, 427−431.
10. Gronwald, W., Loewen, M. C., Lix, B., Sönnichsen, F. D., Daugulis, A. J., Davies, P. L., and Sykes, B. D. (1998) *Biochemistry 37*, 4712−4721.
11. Sönnichsen, F. D., DeLuca, C. I., Davies, P. L., and Sykes, B. D. (1996) *Structure 4*, 1325−1337.
12. Jia, Z., DeLuca, C. I., Chao, H., and Davies, P. L. (1996) *Nature 384*, 285−288.
13. Cheng, C. C., and DeVries, A. L. (1991) in *Life Under Extreme Conditions* (diPrisco, G., Ed.) pp 1−14, Springer-Verlag, Berlin.
14. Wen, D., and Laursen, R. A. (1992) *J. Biol. Chem. 267*, 14102−14108.
15. Chao, H., Houston, M. E., Hodges, R. S., Kay, C. M., Sykes, B. D., Loewen, M. C., Davies, P. L., and Sönnichsen, F. D. (1997) *Biochemistry 36*, 14652−14660.
16. Chao, H., Sonnichsen, F. D., DeLuca, C. I., Sykes, B. D., and Davies, P. L. (1994) *Protein Sci. 3*, 1760−1769.
17. DeLuca, C. I., Chao, H., Sönnichsen, F. D., Sykes, B. D., and Davies, P. L. (1996) *Biophys. J. 71*, 2346−2355.
18. Houston, M. E., Chao, H., Hodges, R. S., Sykes, B. D., Kay, C. M., Sönnichsen, F. D., Loewen, M. C., and Davies, P. L. (1997) *J. Biol. Chem. 273*, 11714−11718.
19. Wen, D., and Laursen, R. A. (1992) *Biophys. J. 63*, 1659−1662.
20. Slaughter, D., Fletcher, G. L., Ananthanarayanan, V. S., and Hew, C. L. (1981) *J. Biol. Chem. 256*, 2022−2026.
21. Ewart, K. V., and Fletcher, G. L. (1993) *Mol. Mar. Biol. Biotechnol. 2*, 20−27.
22. Ewart, K. V., Rubinsky, B., and Fletcher, G. L. (1992) *Biochem. Biophys. Res. Commun. 185*, 335−340.
23. Ng, N. F., and Hew, C. L. (1992) *J. Biol. Chem. 267*, 16069−16075.
24. Weis, W. I., Crichlow, G. V., Murtly, H. M., Hendrickson, W. A., and Drickamer, K. (1991) *J. Biol. Chem. 266*, 20678−20686.
25. DeReggi, M., Gharib, B., Patard, L., and Stoven, V. (1998) *J. Biol. Chem. 273*, 4967−4971.
26. Sönnichsen, F. D., Sykes, B. D., and Davies, P. L. (1995) *Protein Sci. 4*, 460−471.
27. Ewart, K. V., Yang, D. S. C., Ananthanarayanan, V. S., Fletcher, G. L., and Hew, C. L. (1996) *J. Biol. Chem. 271*, 16627−16632.
28. Ewart, K. V., Li, Z., Yang, D. S. C., Fletcher, G. F., and Hew, C. L. (1998) *Biochemistry 37*, 4080−4085.
29. Loewen, M. C., Liu, X., Davies, P. L., and Daugulis, A. J. (1997) *Appl. Microbiol. Biotechnol. 48*, 480−486.
30. Cregg, J. M., Barringer, K. J., and Hessler, A. Y. (1985) *Mol. Cell Biol. 5*, 3376−3385.
31. Scorer, C. A., Clare, J. J., McCrambie, W. K., Ramanos, M. A., and Sreekrishna, K. (1994) *Bio/Technology 12*, 181−184.
32. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
33. Chakrabartty, A., and Hew, C. L. (1991) *Eur. J. Biochem. 202*, 1057−1063.
34. Kunkel, T. A., Roberts, J. D., and Zarkour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
35. Mead, D. A., Sczesna-Skarupae, E., and Kemper, B. (1986) *Protein. Eng. 1*, 67−74.
36. Sanger, F., Nicklen, S., and Carlson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.

37. Hayes, P. H., Scott, G. K., Ng, N. F. L., Hew, C. L., and Davies, P. L. (1989) *J. Biol. Chem. 264*, 18761−18767.
38. Dussetti, N. J., Frigerio, J. M., Keim, V., Dagorn, J. C., and Iovanna, J. L. (1993) *J. Biol. Chem. 268*, 14470−14475.
39. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
40. DeLuca, C. I. (1997) *Ph.D. Dissertation*, Queen's University, Kingston, ON.
41. Bertrand, J. A., Pignol, D., Bernard, J.-P., Verdier, J.-M., Dagorn, J.-C., and Fontecilla-Camps, J. C. (1996) *EMBO J. 15*, 2678−2684.
42. Patard, L., Stoven, V., Gharib, B., Bontems, F., Lallemand, J.-Y., and DeReggi, M. (1996) *Protein Eng. 9*, 949−957.
43. Bradford, J. G., Crowther, R. L., Chandran, C., Rumberger, J. M., Li, S., Huang, K.-S., Presky, D. H., Familletti, P. C., Wolitsky, B. A., and Burns, D. K. (1994) *Nature 267*, 532−538.
44. Yang, D. S., Hon, W. C., Bubanko, S., Xue, Y., Seetharaman, J., Hew, C. L., and Sicheri, F. (1998) *Biophys. J. 74*, 2142−2151.
45. Brisset, N. C., and Perking, S. J. (1996) *FEBS Lett. 388*, 211−216.
46. DeLuca, C. I., Davies, P. L., Ye, Q., and Jia, Z. (1998) *J. Mol. Biol. 275*, 515−525.
47. Monaco, H. L., and Zanotti, G. (1992) *Biopolymers 32*, 457−465.
48. Davies, P. L., Ewart, K. V., and Fletcher, G. L. (1993) in *Fish Biochemistry and Molecular Biology* (Mommsen, T. P., and Hochachka, P. W., Eds.) Vol. 2, pp 279−291, Elsevier, Amsterdam.
49. Tanaka, T., Har-El, R., and Tanyer, M. L. (1988) *J. Biol. Chem. 263*, 15831−15835.
50. Bezouska, K., Crichlow, G. V., Rose, J. M., Taylor, M. E., and Drickamer, K. (1991) *J. Biol. Chem. 266*, 11604−11609.
51. Suter, U., Bastos, R., and Hofstetter, H. (1987) *Nucleic Acids Res. 15*, 7295−7308.
52. Drickamer, K., and McCreary, V. (1987) *J. Biol. Chem. 262*, 2582−2589.
53. White, R. T., Dam, D., Miller, J., Spratt, K., Shilling, J., Hawgood, S., Benson, B., and Cordell, B. (1985) *Nature 317*, 361−363.
54. Ord, D. C., Ernst, T. J., Zhou, L. J., Ranbaldi, A., Spertini, O., Griffin, J., and Tedder, T. F. (1990) *J. Biol. Chem. 265*, 7760−7767.
55. Johnston, G. I., Bliss, G. A., Newman, P. J., and McEver, R. P. (1990) *J. Biol. Chem. 265*, 21381−21385.

BI9820513