Biochemical and Biophysical Characterization of Refolded Drosophila DPP, a Homolog of Bone Morphogenetic Proteins 2 and 4*

(Received for publication, June 8, 1998)

Jay Gropppe§§, Klaus Rumpel¶¶, Aris N. Economides**, Neil Stahl**, Walter Sebald‡‡, and Markus Affolter§§

From the §Department of Cell Biology and ¶¶Department of Biophysical Chemistry Biozentrum, University of Basel, CH-4056 Basel, Switzerland, **Regeneron Pharmaceuticals, Inc., Tarrytown, New York 10591, and ‡‡Theodor-Boveri-Institut (Biozentrum), Physiologische Chemie II, Am Hubland, D-97074 Würzburg, Germany.

The mature C-terminal signaling domain of the Drosophila Decapentaplegic proprotein (DPP) can be efficiently refolded from chaotrope-solubilized inclusion bodies with the aid of a membrane protein-solubilizing detergent, high concentrations (0.75–2 M) of NaCl, and low temperatures (5–15 °C). The disulfide-linked homodimeric product contains N-terminal heparin-binding sites that were utilized as intrinsic affinity tags to obtain a highly enriched preparation in one chromatographic step. A subsequent C4 reverse phase high pressure liquid chromatography step provides high purity, salt-free protein that is amenable to biophysical and structural studies at a yield of approximately 3 mg/liter of bacterial culture. The dimeric protein is correctly folded as determined by electrophoretic, spectroscopic, chemical, and proteolytic analyses. Refolded DPP is also active as shown by induction of chondrogenesis in embryonic chick limb bud cells and by high affinity binding to Noggin, an antagonist of bone morphogenetic protein signaling. In contrast to bone morphogenetic proteins extracted from demineralized bone or overexpressed in cell culture, the refolded Escherichia coli-expressed protein is not glycosylated at a conserved N-linked site and is therefore homogeneous. The C-terminal domain dimer is more hydrophobic and thus less soluble than its unfolded or partially folded forms, necessitating highly solubilizing conditions for recovery after folding in vitro. Hence solubilization of the mature ligand may be one of the principal roles of the large (250–400 amino acids) N-terminal prodomains of transforming growth factor-β superfamily members, shown to act as intramolecular chaperones in vivo.

The decapentaplegic locus (dpp) encodes a signaling ligand homologous to human bone morphogenetic proteins BMP-21

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Supported by the Kanton Basel. To whom correspondence should be addressed: Dept. of Cell Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland. Tel.: 41 61 267 2076; Fax: 41 61 267 2078; E-mail: gropppe@ubaclu.unibas.ch.

¶ Present address: Pfizer Central Research, Sandwich, Kent, CT13 9NJ UK.

§§ Supported by a START fellowship of the Swiss National Science Foundation and by the Kanton Basel.

1 The abbreviations used are: BMP, bone morphogenetic protein(s); TGF, transforming growth factor; DPPC, reduced, unfolded monomeric form of the mature C-terminal domain of DPP; MES, 4-morpholineethanesulfonic acid; hNNG, human Noggin fused to the Fc domain of and -4 (1, 2) and has been the object of intense study due to its importance in many fundamental processes during Drosophila development. The mature DPP ligand is secreted as a disulfide-linked homodimer of the C-terminal domain of a large proprotein precursor and has been shown to act as a morphogen required for the establishment of the dorsoventral axis of the embryo (3) and to impart positional information over long range to specify cell fate along the anterior/posterior axis of the limbs (4, 5). DPP signaling also occurs at short range between germ layers of the developing gut (6–8) and is required to direct the dorsal and ventral migration of tracheal cells during embryogenesis (9) and to act as a relay signal triggering epithelial cell shape changes during the process of dorsal closure (10). In contrast to the developmental roles of DPP signaling, which have been the focus of extensive genetic studies, little is known about the folding, modification, and processing of the DPP proprotein, its distribution and state extracellularly, or the biochemical and biophysical properties of the mature homodimer and specific aspects of its structure.

Characterization of the protein products of the dpp locus has been hindered by a lack of appreciable amounts of material for analysis. DPP is a relatively insoluble signaling ligand present in low concentrations and cannot be isolated preparatively from Drosophila tissue. Attempts to express the DPP proprotein by transient transfection in mammalian cell culture did not produce detectable levels of the mature protein (11). To date the most successful expression system has been a Drosophila embryonic cell line (S2 cells) stably transfected with an inducible metallothionein promoter-dpp cDNA fusion construct (12). Although mature dimeric protein can be recovered from the conditioned media, the bulk of the ligand is adsorbed to the surface of the tissue culture plate and can be recovered in enriched form by washing with 0.5 M NaCl, 1% Tween 20. The S2 cell expression system has provided DPP homodimer for the determination of its N terminus and isoelectric point (12) and demonstration of ectopic bone-forming activity in mammals (13). However, further studies to characterize the signaling activity of the ligand have been hampered by the unavailability of high purity preparations. For example, the affinity of DPP for the type I receptors could not be quantitated due to the inability to produce 125I-DPP for binding assays, necessitating replacement with human recombinant BMP-2 (14, 15). Similarly, human recombinant BMP-4 was substituted for the bona fide Drosophila ligand to assay for Noggin function in the DPP

human IgG1; CHAPS, 3-[3-cholamidopropyl]dimethylammonio]-2-hydroxy-1-propanesulfonic acid; CHAPSO, 3-[3-cholamidopropyl]dimethylammonio]-1-propanesulfonate; DPPC, oxidized, folded dimeric form of the mature C-terminal domain of DPP.
signaling pathway (16) and recombinant BMP-2 utilized in lieu of DPP to stimulate ligand-dependent phosphorylation of MAD proteins in Drosophila imaginal disk cell lines (17).

In this report, we show that the mature C-terminal signaling domain of Drosophila DPP can be efficiently refolded in vitro from chaotrope-solubilized inclusion bodies and purified to near-homogeneity by heparin affinity chromatography followed by reverse phase HPLC. Approximately 3 mg of high purity protein can be produced per liter of bacterial culture, enabling biochemical studies and assays of signaling activity to be performed with DPP ligand rather than human BMP homologs and opening the way for biophysical and structural analyses of the C-terminal domain. Our studies also show that, in addition to the extracellular domains of heterodimeric receptor complexes, DPP binds with high affinity to heparin, a glycosaminoglycan related to heparan sulfate of extracellular matrix and cell-surface proteoglycans, and with high specificity to Noggin, a protein antagonist of BMP activity in vertebrates. Because the E. coli-expressed, refolded protein migrated reproducibly faster in SDS-PAGE than homodimer secreted by insect cells, DPP appears to be glycosylated like all BMPs at a conserved N-linked site. Inspection of the crystal structure of BMP-7 (18) and a model of the DPP homodimer reveals that the modified asparagine side chains reside in the bottom of a large cleft on one side of the dimers, opposite the pair of highly basic, disordered N-terminal peptides that extend from the folded core of the mature domains. The surface of the DPP dimer appears to be predominantly hydrophobic, comprised of two large symmetry-related pairs of hydrophobic patches. As shown by reverse phase HPLC, the folded dimer is more hydrophobic and thus less soluble than its unfolded or partially folded forms. The highly solubilizing conditions found to be crucial for the production of TGF-β superfAMILY signal ligands in vitro, a membrane protein-solubilizing detergent, high concentrations of NaCl, and low temperatures, appear to be required for recovery of the folded product rather than to promote the process of folding per se, as is more commonly observed. Thus solubilization of the folded C-terminal domain of TGF-β superfAMILY members may be one of the principal roles of the large (250–400 amino acids) N-terminal prodomain, which are required as intramolecular chaperones in vivo to provide for secretion of the mature signal ligands (19).

**EXPERIMENTAL PROCEDURES**

**Expression Vector Construct**—The mature C-terminal domain of DPP (DPPC) was overexpressed in Escherichia coli with the bacteriophage TT RNA polymerase/610 promoter system (20). A PCR product encoding the mature protein was synthesized with Taq DNA polymerase (Stehelein) from a cDNA template and introduced into pJC10 (21), a high copy number derivative of pET3c, by cleavages with Ndel and BamHI. These sites were incorporated into the DNA product by oligonucleotide primers encoding N- and C-terminal sequences, respectively. The Ndel site provided a methionine initiation codon downstream of the 610 promoter and ribosome-binding site of pJC10. The BamHI site is preceded by a stop codon and flanked by a 6 terminator. Analysis of the reverse phase HPLC-purified folded dimer prepared from the E. coli-expressed protein concentration (approximately 18 mg/ml) was estimated by the method of Bradford (23) by solubilizing microtiter volumes of a 10-fold dilution of resuspended inclusion bodies directly in the phosphoric acid/methanol reagent which was then diluted with water. HPLC-purified DPPC protein and bovine serum albumin gave similar responses in the dye-binding assay, indicated that bovine serum albumin standard curves can be used to estimate DPPC protein concentration, as has been shown previously to be feasible for DPP dimer from S2 cells and for the BMP family of proteins in general (13). Insufficient sonication resulted in loose, jelly-like pellets of inclusion bodies due to contamination with high molecular weight DNA which could be eliminated by additional cycles. As an alternative to sonication, a French press could also be employed and may be more practical with larger preparations.

**In Vitro Folding**—Inclusion bodies suspended in TEB buffer were centrifuged briefly (−2 min) in a microcentrifuge and solubilized at an estimated concentration of 2.5 mg/ml in buffered chaotrope. The final molarity of chaotrope is not significantly lowered by addition of the inclusion bodies, which were diluted more than 70-fold from an estimated initial concentration of 180 mg of protein/ml pellet. 8 μl x 6 M GdmCl, buffered with 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 1 mM DTT, gave equivalent results. After incubation for 1–2 h at ambient temperature, insoluble debris was removed by a second brief centrifugation, and the solubilized protein was diluted 50-fold (final concentration 50 μg/ml), rapidly with stirring, into pre-chilled folding buffer (50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 33 mM CHAPS detergent (1.8% v/v), 1.25 mM NaCl, 2 mM reduced glutathione, and 1 mM oxidized glutathione) and incubated for at least 72 h at 4 or 15 °C.

**Purification**—The folding solution was concentrated 10-fold by ultrafiltration at 4 °C with a stirred cell device fitted with a YM10 membrane (Amicon), diluted 12.5-fold with 6 μl urea, 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, to yield a final NaCl concentration of 0.1 mM and reconstituted to approximately 50 mM. The turbid solution was cleared of precipitate by centrifugation at 5000 g for 5 min. Residual unfolded monomeric folded dimeric product, and higher order multimers were separated by heparin affinity chromatography on a 5-m M Heparin HiTrap column by fast protein liquid chromatography (Amersham Pharmacia Biotech) at ambient temperature through combined step and gradient elution with NaCl. The column was equilibrated with 6 μl urea, 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 0.1 mM NaCl at a flow rate of 1 ml/min, and the...
Refolded Drosophila DPP

Concentrated, buffer-exchanged folding reaction was pumped onto the column with a 50-ml Superloop. After washing with 5 column volumes, the bulk of the unfolded monomer was eluted with a step gradient of 0.25 M NaCl as a single peak. The folded dimeric product eluted as a broad peak with a 25-ml gradient from 0.25 to 0.4 M NaCl and the major peak was further purified to homogeneity by reverse phase HPLC on a semipreparative, 10 × 250 mm C18 column (Vydac 214TP510). Aliquots of the protein were adjusted to 60% solvent A (0.1% trifluoroacetic acid), 40% solvent B (80% acetonitrile, 0.08% trifluoroacetic acid) in a total volume of 1.0 ml, injected onto the column at a flow rate of 2.5 ml/min, and eluted with a gradient from 40 to 60% B over 60 min. Injections of approximately 1.3, 1.7, 2.3, and 3.4 mg provided similar recovery efficiencies with no apparent decrease in resolution (larger loadings were not tested). The peak of folded dimer, which reproducibly eluted with a retention time of 43.5 min, was collected manually. Peak fractions from multiple injections were pooled, the UV absorption spectrum measured, and aliquots of 15–20 ml frozen in 50 ml propylene tubes, lyophilized, and stored at −20 °C.

**Trypsin Cleavage**—Activated thiol-Sepharose chromatography was performed to demonstrate dose-dependent binding of hNG-Fc to xBMP-Fc and to allow determination of the number of hNG-Fc bound. Competitive binding of DPPC2, human BMP-4, and human Noggin against hNG-Fc was determined by automated cycles (9 and 5, respectively) of Edman degradation performed with an Applied Biosystems 473A Protein Sequencer and analyzed on line with a phenylthiohydantoin-amino acid analyzer.

**UV Absorption Spectroscopy**—To assess the yield, purity, and solubility of the protein, UV absorption spectra (240–340 nm) of the dimer preparations were determined routinely following enrichment by heparin affinity chromatography, dialysis, and concentration and purification by reverse phase HPLC. For the determination of difference spectra, 70 μg of protein in 0.1% trifluoroacetic acid was diluted 10-fold into 6.67 M GdmCl (AA-Grade, NIGU Chemie GMBH, Waldkrasburg, Germany) buffered with 50 mM Tris-HCl, pH 8.0. Final protein concentration was 100 μg/ml or approximately 7 μM. After establishment of a base line, the UV absorption spectrum from 240–310 nm was recorded. A second base line was then established with the monomer sample, the spectrum of the dimer measured, and the difference spectrum obtained from the derivative. Concentrations of the reduced monomer and the folded, oxidized dimer were estimated from absorbance at 280 nm with molar absorption coefficients of 19,940 and 40,755 M−1 cm−1, respectively. Coefficients were calculated by the method of Pato et al. (37), which includes contributions from disulfide bonds (€ 280 = 25,000 M−1 cm−1 = εTrp500 + εTrp450 + εCyS). The calculated molecular weights of the reduced monomer and the oxidized dimer, including the N-terminal cysteine residues, are 14,819 Da and 29,625 Da. Monomer and dimer at 1 mg/ml therefore correspond to absorbances at 280 nm of 1.35 and 1.38, respectively. All spectra were determined with a single-beam, microprocessor-controlled UV/VIS spectrophotometer (Hewlett-Packard 8452A Diode Array) with data storage capacity.

**Noggin Binding Assay**—Anti-myc monoclonal antibody (9E10) was coated on ELISA plates (Corning) in phosphate-buffered saline at 2 μg/ml by passive binding. Unbound 9E10 was removed by washing with phosphate-buffered saline, and nonspecific binding was blocked with 1% bovine serum albumin in phosphate-buffered saline. Xenopus laevis oocytes were injected with a master mix containing 0.3% bovine serum albumin (BSA) and 2 μg of 9E10-coated ELISA plates by incubation for 1 h and unbound BSA was removed by washing. Increasing amounts of hNG-Fc were added to the plate coated with BSA and incubated for 1 h. Unbound hNG-Fc was removed by washing, and 0.5 μg/ml alkaline phosphatase-conjugated anti-human IgG (anti-Fc AP) was added to each well. After 1 h, unbound anti-Fc AP was removed by washing, and alkaline phosphatase substrate was added, and the reactions were allowed to proceed for approximately 15 min. The extent of the reactions was determined by measuring the absorbance at 405 nm with an automatic plate reader, and the data in tabular form were converted to a graphic format with the program Cricket Graph (Computer Associates). A standard curve of hNG-Fc (human Noggin fused to the Fc domain of human IgG1) was determined to demonstrate dose-dependent binding of hNG-Fc to xBMP-Fc and to allow correlation between A405 units and the fraction of hNG-Fc bound. Competitive binding of DPPPC2, human BMP-4, and human Noggin against hNG-Fc was determined as described above, except that the concentration of hNG-Fc was held constant (110 ng/ml) and the concentration of the competitive ligands varied (1–1000 ng/ml). The buffer components of either the DPPC2 or the hNG preparations used in this assay have no effect on the binding of hNG-Fc to xBMP-Fc at the levels employed. Similar results are obtained when blocking antibodies to hNG are used and no binding of hNG-Fc to the plates is observed if xBMP-Fc is omitted.

**Chick Limb Bud Cell Assay**—BMP-induced synthesis of proteoglycan in chicken embryonic limb bud cells was assayed as described previously (29). Briefly, limb buds were isolated from day 7 chicken embryos, cultured in culture dishes pretreated with poly-L-lysine, and incubated on cellulose filters. After incubation in BMP-containing medium for 4 days, 20 μl of 30 μCi/ml Na35SO4 was added to each well, and the cultures were incubated for an additional 6 h. Cells were lysed with GdmCl and proteoglycans precipitated with Alcian blue and collected on glass paper with a Skatron cell harvester. 35SO4 incorporation was determined with a radio-TLC analyzer, and the data in tabular form were converted to a
FIG. 1. Structural features of the C-terminal domain of DPP.
Maturation of the DPP proprotein in vitro produces a homodimeric signaling ligand composed of two disulfide-linked, 132-amino acid monomers of the C-terminal domain (12). The N-formylmethionine residue, required for bacterial overexpression, was not cleaved posttranslationally by the E. coli host. Like other members of the BMP family, each monomer of the secreted form of the DPP homodimer is expected to be glycosylated at a conserved N-linked site (*) contained by the consensus motif, NX(S/T). The cystine-knot which is the structural hallmark of the mature domain of all TGF-β superfamily members is composed of a pair of linkages (Cys<sup>65</sup>-Cys<sup>132</sup>; Cys<sup>61</sup>-Cys<sup>130</sup>) through which a third (Cys<sup>22</sup>-Cys<sup>98</sup>) is formed. The homodimer is covalently linked via an intermolecular disulfide bond between residue 97 of each monomer. The N-terminal polypeptide arm, which is not conserved among TGF-β superfamily members, is a flexible extension from the folded, disulfide-linked core (shaded) and contains clusters of basic residues (bold) which have been implicated in mediating high affinity binding to heparin (28). Arginine and lysine residues within the core which are potential substrates for tryptic cleavage are also indicated. The boundary between the disordered N-terminal arm and the folded core is predicted from the electron density map of BMP-7 (18).

The amino acid sequence encoded by the expression construct is not in agreement with the cDNA-derived sequence (1), and its database entry is without providing initial estimates for the iterative calculations.

Model of the DPP Homodimer—The three-dimensional structure of the DPP homodimer was modeled with ProMod (29, 30), an automated, homology-based method. The polypeptide backbone of the DPPC monomer was produced by sequence alignment and three-dimensional superposition with the four representatives of the TGF-β superfamily (TGF-β1, TGF-β2, TGF-β3, and BMP-7) currently in the Brookhaven Protein Data Bank. As the C-terminal domains of BMP-7 and DPP were aligned without the inclusion of gaps, no additional loops were constructed de novo and only non-conserved side chains had to be introduced and existing side chains corrected by referencing a library of allowed side chain rotamers. The model was subjected to several cycles of energy minimization with CHARMM and refined by optimizing intramolecular contacts and relieving steric strain. A symmetry transformation was employed to generate a dimeric structure from the final model of the monomer. Template structures and models were visualized with RASMOL (31). The model of the monomer, when transformed to a dimer, does not result in any apparent large gaps or steric conflicts within the structure. Because the disordered N-terminal arm is unresolved in the electron density map and therefore lacking from the crystal structure of BMP-7, the structure of this polypeptide could not be incorporated into the model of the DPP homodimer.

RESULTS

In Vitro Folding Reaction

Ruppert et al. (28) have shown that the human homolog of DPP, BMP-2, can be refolded in vitro in the presence of a zwiterionic detergent and high salt, conditions determined empirically for the refolding of TGF-β2 (32). We found that the C-terminal domain of DPP (DPPC; Fig. 1) also can be efficiently refolded in vitro under these conditions. As is common for many secreted disulfide-containing proteins, the DPP homodimer contains no unpaired cysteines and thus the oxidized dimeric folding product can be denatured under non-reducing condi-

[2] Accessible through the Swiss model server on the World Wide Web at http://www.expasy.ch/swissmod/SWISS-MODEL.html.

[3] J. Groppe and M. Affolter, unpublished observations.
A variety of chromatography matrices (cation exchange, anion exchange, and chromatofocusing) and buffer combinations (sodium acetate, Tris-HCl, HEPES, MBS, 30% isopropanol alcohol, and 6 M urea) were screened for their ability to separate residual unfolded and misfolded forms from the folded dimeric product. However, all forms of the protein failed to bind these matrices and co-eluted in the flow-through. Because the DPP dimer was expected to contain two N-terminal heparin-binding sites (28), heparin-Sepharose was tested and found to provide efficient separation of the folded dimer from non-native forms (Fig. 3, A and B). All forms of the protein were bound to heparin quantitatively in low salt (6 mM urea, 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 0.1 M NaCl). Dimers, both folded and misfolded, bound with affinity intermediate between the low affinity of the monomer and the high affinity of the larger multimeric misfolded forms. Gradient elution alone resulted in partially overlapping peaks of monomer and dimer which could not be resolved by dimerers.35 How an initial step gradient resulted in the quantitative elution of the bound monomer in a single well formed peak. A subsequent linear gradient of NaCl yielded a broad peak of dimeric protein containing the folded product in highly enriched form.

After enrichment, the dimeric product could be bound quantitatively to a cation exchange matrix (SP-Sepharose; 6 M urea, 50 mM HEPES, pH 7.5, 0.15 M NaCl), but the folded protein eluted from a linear gradient of NaCl in a broad peak contaminated by the trace non-native forms, perhaps due to shared ionic properties of the proteins.35 In contrast, the folded dimer was purified to near-homogeneity by reverse phase HPLC with a linear gradient of acetonitrile (Fig. 3C), which separates proteins primarily due to differences in surface hydrophobicity. Residual unfolded monomer and misfolded forms eluted in a broad shoulder preceding a single uniform peak of dimeric product, which was retained last by the C4 matrix.

Approximately 3 mg of HPLC-purified folded dimer was obtained per liter of bacterial culture (Table I). Differences in yield of the purified product arose predominantly from variability in the yield of inclusion body protein, which in turn depended on the efficiency of induction (cf. “Experimental Procedures”). The induced protein was found exclusively in the monomer. B, effect of organic co-solvent and folding trials with other Drosophila BMP homologs. Folding reactions containing the C-terminal domains of DPP, 60A, and SCREW were performed in standard reaction buffer (std rxn) or in standard reactions supplemented with 20% Me2SO (+DMSO) and analyzed as described above. Products in the folding reactions with the C-terminal domain of DPP are indicated with arrows as follows: upper arrow, oxidized dimer and lower arrow, monomeric intermediate; C, NaCl dependence. Reactions containing the C-terminal domain of DPP were performed under standard conditions except that the concentration of NaCl was varied from 0 to 2.0 M NaCl in increments of 0.25 M. The reactions were quenched by the addition of trichloroacetic acid, adjusted to an equivalent concentration of NaCl (2.0 M) without altering the composition of CHAPS (1.8%), and then incubated on ice and collected by centrifugation. D, temperature dependence. Standard reactions containing DPPC were incubated between 5 and 35 °C in increments of 10 °C for 3 days. An aliquot of the solubilized inclusion body protein was stored in 6 M GdmCl at 5 °C during the course of the folding reaction and then diluted into trichloroacetic acid-quenched folding buffer for precipitation and analysis by non-reducing SDS-PAGE (−). Inclusion body protein analyzed in C and D was prepared under reducing conditions as described under “Experimental Procedures,” whereas the preparations shown in A and B were washed in TE buffer without DTT and subjected to a final wash with 10% acetic acid, which allowed for removal of residual lysozyme (migrates slightly faster than DPPC monomer) and other more high molecular weight contaminants. DPPC red, reduced monomer; DPPC ox, oxidized dimer; 2 DPPC reduced, non-covalently linked, dimeric aggregate.

Fig. 2. Characterization of the in vitro folding reaction. A, parameters of the reaction. Chaotrope-solubilized inclusion bodies of the C-terminal domain of DPP (DPPC) were rapidly diluted 50-fold in standard reaction buffer (2 M NaCl, 1.8% w/v CHAPS, 2 mM GSH, 50 mM Tris-HCl, pH 8.0) to a final concentration of approximately 50 μg/ml and incubated at 4 °C for 3 days (std rxn). The effects of supplementation with 5 mM EDTA, incubation at ambient temperature, reduced protein concentration, varied pH, an equimolar (1 mM) ratio of redox pair, and elimination of either NaCl or CHAPS from the reaction were analyzed. The reactions were terminated and the protein precipitated by the addition of trichloroacetic acid to 20% (see “Experimental Procedures”) and analyzed by non-reducing SDS-PAGE. An aliquot of the unfolded inclusion body protein was subjected to boiling in the presence of 2% SDS and 0.2 mM DTT and co-electrophoresed (reduced monomer).
Refolded Drosophila DPP 29057

Evidence for a Folded Product

The inclusion body-purified protein appears to be correctly refolded and oxidized to a disulfide-linked homodimer, as evidenced by several independent observations.

Chromatographic—Following lyophilization and resolubilization in 50% acetonitrile, 0.1% trifluoroacetic acid, the purified dimer was reanalyzed by C4 reverse phase HPLC and found to elute in a single sharp peak (Fig. 3D). Little or no unfolded or misfolded forms were observed, indicative of a stable disulfide-linked structure resistant to rearrangement or denaturation.

Electrophoretic—Comparison of the mobilities of oxidized and reduced forms of denatured proteins in polyacrylamide gels provides a simple assay for disulfide bonds, which restrict the flexibility of the polypeptide and decrease its hydrodynamic volume (36). Analysis of HPLC-purified dimeric and monomeric forms of DPPC clearly shows the greater mobility of the folded dimeric product compared with that of aggregated monomer, indicative of a more compact, disulfide-linked structure (Fig. 4A). SDS-PAGE analysis of the purified proteins also clearly demonstrates that the dimeric product migrates as an apparent single species or conformer, in contrast to the aggregated monomer which migrates in a disperse zone indicative of a heterogeneous population of unfolded or misfolded polypeptides.

Chemical—Activated thiol-Sepharose 4B (Amersham Pharmacia Biotech) provides a means of separating thiol-containing from nonthiol-containing proteins. A mixed disulfide between Sepharose-linked glutathione and 2-pyridyl reacts through 1-ml HiTrap heparin affinity column with a 10-ml Superloop mounted on an fast protein liquid chromatography system (Amersham Pharmacia Biotech). The bound protein was washed extensively with 6 M urea, 50 mM Tris-HCl, pH 8.0, 0.1 M NaCl until the absorbance at 280 nm of the elution returned to base line. Residual monomeric protein was eluted with a step gradient from 0.1 to 0.25 M, the folded dimeric product with a linear gradient from 0.25 to 0.4 M, and misfolded multimeric forms with a linear gradient from 0.4 to 1.0 M NaCl. Non-reducing SDS-PAGE analysis of fractionation depicted in A. 1-μl aliquots from the indicated column fractions (1.0 ml each) were boiled in SDS-PAGE loading buffer containing 2% SDS (without reducing agent) and analyzed on a denaturing 15% polyacrylamide gel. Protein was visualized by staining with Coomassie. An aliquot of the flow-through (FT) peak, not depicted in A, was also analyzed. C and D, C4 reverse phase HPLC purification and analysis of folded DPPC2. C, purification of folded dimer from residual monomer and misfolded forms. The folded dimer eluted late in the acetonitrile gradient (retention time; 40 min) as a single, uniform peak compared with unfolded or misfolded species that appeared earlier as a broad shoulder (18–38 min). D, analysis of purified DPPC2. An aliquot (~30%) of the peak material collected from the preparative chromatographic run shown in A was lyophilized, redissolved in 50% acetonitrile, 0.1% trifluoroacetic acid, and rechromatographed. The purified, re-solubilized protein eluted as a single sharp peak, equilibrating with or contaminated by little or no detectable denatured forms. The small peaks at approximately 16 and 28 min, which were not observed in the preparative profile, appear to be solvent artifacts or ghost peaks emanating from a less stable base line. The purified protein reproducibly eluted slightly later (~42 min) in the gradient when rechromatographed. HPLC depicted in C and D were performed with an analytical C4 column (Vydac 214TP54, 4.6 × 250 mm). Milligram quantities were purified with a semi-preparative C4 column (Vydac 214TP510, 10 × 250 mm) from which the dimeric product eluted in a peak from 41 to 46 min.

FIG. 3. Purification of the dimeric folding product. A and B, separation of DPPC, DPPC2, and multimeric forms by heparin affinity chromatography. A, profile of HiTrap heparin affinity column. Approximately 40 mg of chaotrope-solubilized inclusion body protein was rapidly diluted into standard folding buffer (200 ml) and incubated at 4 °C for 3.5 days. The reaction was concentrated and the buffer exchanged to 6 M urea, 50 mM Tris-HCl, pH 8.0, by ultrafiltration and pumped onto a inclusion body fraction and quantitatively recovered in a highly purified form (cf. Fig. 2A). Around 15% of the chaotrope-solubilized protein could be refolded routinely under the standard conditions and obtained in highly enriched form in one step by heparin affinity chromatography. Approximately one-half of the enriched pool of dimeric product was recovered at near-homogeneity by C4 reverse phase HPLC.

The inclusion body-purified protein appears to be correctly refolded and oxidized to a disulfide-linked homodimer, as evidenced by several independent observations.

Chromatographic—Following lyophilization and resolubilization in 50% acetonitrile, 0.1% trifluoroacetic acid, the purified dimer was reanalyzed by C4 reverse phase HPLC and found to elute in a single sharp peak (Fig. 3D). Little or no unfolded or misfolded forms were observed, indicative of a stable disulfide-linked structure resistant to rearrangement or denaturation.

Electrophoretic—Comparison of the mobilities of oxidized and reduced forms of denatured proteins in polyacrylamide gels provides a simple assay for disulfide bonds, which restrict the flexibility of the polypeptide and decrease its hydrodynamic volume (36). Analysis of HPLC-purified dimeric and monomeric forms of DPPC clearly shows the greater mobility of the folded dimeric product compared with that of aggregated monomer, indicative of a more compact, disulfide-linked structure (Fig. 4A). SDS-PAGE analysis of the purified proteins also clearly demonstrates that the dimeric product migrates as an apparent single species or conformer, in contrast to the aggregated monomer which migrates in a disperse zone indicative of a heterogeneous population of unfolded or misfolded polypeptides.

Chemical—Activated thiol-Sepharose 4B (Amersham Pharmacia Biotech) provides a means of separating thiol-containing from nonthiol-containing proteins. A mixed disulfide between Sepharose-linked glutathione and 2-pyridyl reacts through 1-ml HiTrap heparin affinity column with a 10-ml Superloop mounted on an fast protein liquid chromatography system (Amersham Pharmacia Biotech). The bound protein was washed extensively with 6 M urea, 50 mM Tris-HCl, pH 8.0, 0.1 M NaCl until the absorbance at 280 nm of the elution returned to base line. Residual monomeric protein was eluted with a step gradient from 0.1 to 0.25 M, the folded dimeric product with a linear gradient from 0.25 to 0.4 M, and misfolded multimeric forms with a linear gradient from 0.4 to 1.0 M NaCl. Non-reducing SDS-PAGE analysis of fractionation depicted in A. 1-μl aliquots from the indicated column fractions (1.0 ml each) were boiled in SDS-PAGE loading buffer containing 2% SDS (without reducing agent) and analyzed on a denaturing 15% polyacrylamide gel. Protein was visualized by staining with Coomassie. An aliquot of the flow-through (FT) peak, not depicted in A, was also analyzed. C and D, C4 reverse phase HPLC purification and analysis of folded DPPC2. C, purification of folded dimer from residual monomer and misfolded forms. The folded dimer eluted late in the acetonitrile gradient (retention time; 40 min) as a single, uniform peak compared with unfolded or misfolded species that appeared earlier as a broad shoulder (18–38 min). D, analysis of purified DPPC2. An aliquot (~30%) of the peak material collected from the preparative chromatographic run shown in A was lyophilized, redissolved in 50% acetonitrile, 0.1% trifluoroacetic acid, and rechromatographed. The purified, re-solubilized protein eluted as a single sharp peak, equilibrating with or contaminated by little or no detectable denatured forms. The small peaks at approximately 16 and 28 min, which were not observed in the preparative profile, appear to be solvent artifacts or ghost peaks emanating from a less stable base line. The purified protein reproducibly eluted slightly later (~42 min) in the gradient when rechromatographed. HPLC depicted in C and D were performed with an analytical C4 column (Vydac 214TP54, 4.6 × 250 mm). Milligram quantities were purified with a semi-preparative C4 column (Vydac 214TP510, 10 × 250 mm) from which the dimeric product eluted in a peak from 41 to 46 min.
thiol-disulfide exchange with proteins containing free thiols, whereas nonthiol-containing proteins are unreactive and therefore not covalently bound. The extent of oxidation of the DPPC protein forms was examined by passing a standard folding mixture equilibrated in 6 M GdmCl repeatedly over a column of activated thiol-Sepharose 4B and analyzing the bound and unbound fractions by non-reducing SDS-PAGE (Fig. 4B). The dimeric product appeared to be quantitatively recovered in the unbound fraction, indicating that no reaction had occurred and consistent with an oxidized form containing no free or accessible thiols. The monomeric intermediate (DPPC int) was also found to be unreactive, indicating that the majority of the cysteines had formed disulfide linkages or were inaccessible to the mixed disulfide linked to the column matrix. Most of the unfolded monomer and misfolded forms reacted covalently, as little of these forms were detected in the flow-through fraction and could be decoupled with DTT-containing buffer. Also consistent with a completely oxidized native structure containing multiple disulfides and no unpaired cysteines is the observation that the dimeric product could be denatured prior to electrophoresis in the presence of SDS by boiling at neutral pH, conditions which would allow for the scrambling of disulfide bonds by free thiols, without any evidence of rearrangement to forms with altered migration rates (cf. Fig. 4A).

Proteolytic—Proteolytic cleavage of proteins in the native conformation is usually observed only in disordered regions of the polypeptide backbone, such as segments linking separate domains, large flexible loops, or extended termini. In contrast, unfolded, misfolded, or partially folded proteins are much more sensitive to proteolysis at multiple sites along the polypeptide chain. Koenig et al. (26) have observed that the compact, disulfide-linked core of human BMP-2 is resistant to proteolysis, whereas the N-terminal polypeptide extending from the folded C-terminal domain is protease-sensitive. Incubation with trypsin yielded truncated products collectively designated DR-BMP-2, a “digit-removed” form. Similarly, incubation of DPPC int with trypsin also yielded a truncated, digit-removed form (Fig. 4C). Solubilization of the dimer in aqueous buffer with either zwitterionic detergent or 4 M urea allowed for a quantitative conversion from full length to the truncated form, which consisted of two major products as revealed by amino-terminal sequence analysis. Approximately 35 and 65% of the DR-DPPC int was found to have N-terminal extensions of four and five residues, respectively, from the domain core (cf. Fig. 1). A similar gradient of protease sensitivity with respect to the folded core was observed for BMP-2. Approximately 30 and 70% of the DR-BMP-2 had extensions of two and four residues, respectively (26). In comparison, treatment of the unfolded monomer under these conditions led to an apparent total proteolysis of the polypeptide at additional lysine and arginine residues within the polypeptide (cf. Fig. 1). The resistance of the disulfide-linked core of the folded dimer to proteolysis was significant. Digestion of the folded dimer under even more rigorous conditions (stoichiometric ratio of dimer and protease, 4 M urea, 24 h incubation with chemically modified, high stability trypsin) failed to yield additional detectable products, demonstrating that the core of the folded product did not contain any protease-accessible disordered regions expected of a partially folded or misfolded state.3

Spectroscopic—Differences in conformation and solubility of the unfolded monomer and the folded dimer were analyzed by ultraviolet absorption spectroscopy (Fig. 5). Each DPPC monomer contains 3 phenylalanine, 6 tyrosine, and 2 tryptophan residues that serve as chromophores, allowing not only for estimation of protein concentration, but also for comparison of the molecular environment of these side chains and the relative solubility of the two forms. Because aromatic amino acids do not absorb at wavelengths greater than 310 nm, a sloping base line in the 310–400 nm region is generally due to light scattering by large particles or aggregates in solution (37). The spectrum of the unfolded monomer slopes strongly in this region, indicative of aggregation (Fig. 5A). In contrast, the spectrum of the folded dimer maintains a base-line value throughout this region, demonstrating that the preparation is more soluble than the unfolded monomer at low pH (pH~2) in this mildly chaotropic buffer (0.1% trifluoroacetic acid). Differences in the curvature of the descending slope (285–295 nm) of the spectra were also observed. Such differences are due to increases in absorbance of tyrosine (285–288 nm) and tryptophan (291–294 nm) residues upon burial in the hydrophobic core of the native, folded protein and are usually accompanied by a red shift in the absorbance maximum (37). Thus, in addition to the environment-specific differences in the descending slope, the absorbance maximum of the unfolded monomer (276 nm) is red-shifted by 2 nm relative to that of the folded dimer (278 nm). To minimize potential artifacts due to the aggregated state of the monomer and to compare the stability of the two forms, uv spectra were also measured under denaturing conditions (6 M GdmCl). Although the spectra were no longer shifted relative to one another, differences in the descending slope were again observed (Fig. 5B). The difference spectra of the two protein forms in 6 M GdmCl (Fig. 5C) reveals peaks at 285 and 293 nm resulting from differences in the solvent exposure of the tyrosine and tryptophan residues, respectively. Thus even under strongly denaturing conditions the folded dimer retains some native-like structure, presumably due to the constraints imposed by the extensive and intertwined disulfide linkages of the cystine knot (cf. Fig. 1).

### Table I

**Summary of protein isolation, in vitro folding, and purification**

| Preparation step | Protein | Step-wise recovery | Buffer, chaotrope, or solvent |
|------------------|---------|--------------------|-----------------------------|
| Inclusion body isolation | mg | % | 50 mM Tris-HCl, 40 mM EDTA, pH 8.0 |
| Chaotrope-mediated solubilization, denaturation | ND | | 8 M urea or 6 M GdmCl, 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 1 mM DTT |
| In vitro folding by rapid dilution | | | 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 33 mM CHAPS detergent (1.5% w/v), 1.25 mM NaCl, 2 mM GSH, 1 mM GSSG |
| Concentration and buffer exchange by ultrafiltration (stirred-cell device) | 5.2 | 13 | 6 M urea, 50 mM Tris-HCl, 2 mM EDTA, pH 8.0, 0.1 M NaCl |
| FPLC heparin-affinity chromatography | ND | | 0.1% trifluoroacetic acid (~pH 2) |
| Dialysis, concentration by ultrafiltration (centrifugal) | 2.7 | 51 | 0.1% trifluoroacetic acid, 35% acetonitrile |
| C4 reverse phase HPLC Lyophilization | | | |

Protein yields are per liter of induced bacterial culture; ND indicates not determined. Overall yield is approximately 7%.
Noggin Binding

Holley et al. (16) have shown by injection of mRNAs in Drosophila embryos that noggin, an antagonist of BMP signaling in vertebrates, also blocks signaling by DPP. Because Noggin exerts its antagonistic effect by binding to BMP-2, -4, and -7 and blocking their ability to interact with their receptors (38), the potent biological activity of noggin may be mediated by direct binding to the dpp-encoded signal ligand. Thus we determined whether refolded DPPC2 could bind Noggin in a competition binding assay (Fig. 6). After determination of the dose dependence of epitope-tagged human Noggin (hNG) binding to antibody-immobilized Xenopus BMP-4, the ability of the refolded DPPC2, human BMP-4, and refolded hNG to compete against the binding of hNG to Xenopus BMP-4 was determined. hNG at fixed concentration was challenged with each ligand at a series of concentrations, and the fraction of epitope-tagged hNG remaining bound was determined by ELISA. The human BMP-4 standard was found to displace half of the bound hNG at a concentration of approximately 0.3 nM. Refolded DPPC2 competed slightly less efficiently, displacing half of the bound Noggin at about 3-fold higher concentration or 1 nM. The refolded hNG control was only slightly more effective than refolded DPPC2, competing half of the bound Noggin at around 0.6 nM. Thus the refolded DPPC2 bound Noggin with an affinity comparable to that of vertebrate BMP-4 and was therefore bioactive with respect to regulation of activity by antagonist binding.

Induction of Chondrogenesis

Recombinant DPP and 60A proteins produced in tissue culture have been shown to induce the formation of endochondral bone after subcutaneous implantation in rats (13), the principal and most definitive means of determining BMP activity (39, 40). However, for optimal results, insect BMP orthologs must be reconstituted with rat collagen carrier prepared from demineralized bone by GdmCl extraction, rather than with commercially available basement membrane-extracted matrices. In addition, analysis of extensive dose curves or direct comparison with other BMP activities is impractical due to the large number of implantations required. An alternative assay, induction of chondrogenesis in micromass cultures of mesenchymal cells from embryonic chicken limb buds (41–44), provides quantita-
tive dose-dependent determination of activity as well as direct comparison between BMPs. In this assay, the EC_{50} (effector concentration at half-maximal response) for BMP-2 was 33 ± 13 nM and the maximum incorporation at saturation 20,600 cpm ± 4,200 cpm (Fig. 7). The DPP ligand exhibited approximately 5-fold reduced activity relative to BMP-2, as estimated by the calculated EC_{50} value (170 ± 60 nM). The maximum incorporation induced by DPPC_{2} at saturation was estimated at 14,900 ± 2,600 cpm, approximately 70% that achieved by BMP-2. However, due to the limited solubility of these ligands above 1 μM, saturation values could not be directly measured and required estimation from the fitted curves, the shapes of which were highly sensitive to small deviations in the measured response at the highest concentrations of ligand. Thus the difference in maximum response achieved by the two ligands did not appear to be significant due to the limitations imposed at higher concentrations. The activity of DPPC_{2} as a function of concentration, on the other hand, is indeed reduced relative to the mammalian protein but may reflect the heterologous nature of the assay. Mammalian fibroblasts such as C3H10T1/2 cells are markedly less responsive to BMP-2 than the avian mesenchyme cells (28). In keeping with this reduction in responsiveness, DPP has no measurable inductive effect on this cell line. In fact, none of the mammalian fibroblast cell lines commonly employed to assay BMP activity has been found to respond to insect BMPs. Thus the reduced activity of the insect ligand relative to the human ortholog at lower concentrations in the chick mesenchyme cell assay most likely reflects reduced complementarity between the insect ligand and the avian receptors rather than a non-native ligand structure. Despite the inherent limitations imposed by the insolubility of the ligands and the heterologous nature of the assay, the refolded DPP dimer induced a strong chondrogenic response in a dose-dependent manner and was therefore clearly bioactive as a signal ligand.

**DISCUSSION**

Although the genetic control of many key developmental processes by *dpp* has been intensively studied, little has been known about the biochemical and biophysical properties of the encoded signaling ligand and specific aspects of its structure due to the lack of appreciable amounts of highly purified protein for analysis. We have shown that the mature C-terminal signaling domain of DPP can be efficiently refolded from chaotrope-solubilized inclusion bodies, yielding a properly folded, bioactive ligand that can be purified to near-homogeneity in two chromatographic steps. Approximately 3 mg of high purity protein can be produced routinely per liter of bacterial culture, enabling not only biochemical but also biophysical and structural studies to be performed.

DPP C-terminal domain dimer was produced in high yield under conditions determined empirically for TGF-β2 (32) that also provided for the production of the human homolog of DPP, BMP-2 (28). The solubilized, denatured proteins are refolded in a zwitterionic detergent and high concentrations of NaCl at low temperature in the presence of a redox system. Glutathione, in its reduced and oxidized forms at a ratio and in concentrations similar to those observed in *vivo* (45), provides for thiol-disulfide exchange leading to a fully oxidized dimeric product and is employed in many, if not most, *in vitro* reactions for the production of disulfide-containing proteins.

In contrast to the requirements for the glutathione pair, the roles of the zwitterionic detergent, high concentrations of salt,
and low temperature are not as readily apparent. Cerletti (33) determined that the optimal detergents for production of TGF-β superfamily members are the non-ionic detergent digitonin and the zwitterionic detergent CHAPS, the latter and its derivative (CHAPSO) being most preferable. These are mild detergents employed frequently and almost exclusively for solubilization of membrane proteins without denaturation. An absolute requirement for detergent to solubilize the monomer was observed (cf. Fig. 2A); however, it remains to be determined whether other detergents would suffice in this capacity, and if so, whether the specific effect of the membrane protein-solubilizing detergents is to promote folding, enhance recovery, or both.

In the absence of NaCl the monomer was only marginally soluble (cf. Fig. 2, A and C); however, at 0.25 M and above, residual unfolded monomer and the folding intermediate could be quantitatively recovered. Thus despite the limited amount of dimer observed in reactions containing low to moderate concentrations of NaCl (0.25–0.5 M), the potential for folding is present under these conditions, as opposed to reactions that lacked salt or detergent. The high concentrations of NaCl (0.75–2.0 M) found to allow for efficient production of the dimer are therefore required to promote folding and/or to provide for recovery of the native form.

Low to moderate concentrations of all salts have a salting-in effect by screening the net charge of a protein in aqueous solution, which leads to a reduction in electrostatic free energy and an increase in its solubility (46). Thus addition of NaCl at low ionic strength to the folding reaction appears to produce such a salting-in effect, allowing for the recovery of the monomeric and intermediate forms. However, at higher concentrations, NaCl found to be the preferred salt for production of TGF-β2 (32) and TGF-β3 (33) in vitro, can have more complex effects on protein solubility and structure. Consistent with their central positions in the Hofmeister series, Na+ and Cl− ions are relatively neutral with respect to their effects on the properties of bulk water but nonetheless can increase the solubility of proteins dramatically at concentrations up to 2 M (46, 47). Several observations suggest that in the case of the production of TGF-β superfamily members, high salt enhances the solubility and hence the recovery of the dimeric product but is not required to promote its folding. First, TGF-β2 can be efficiently refolded in vitro in low salt (250 mM NaCl) as a fusion protein containing a tripartite N-terminal peptide composed of a hexahistidyl tag, a thrombin cleavage site, and a collagen-binding site (48). The role of high salt can thus be mimicked by the N-terminal peptide which most likely has a solubilizing effect, as the efficiency of renaturation of other inclusion body proteins can be significantly enhanced through fusion with hydrophilic partners (49). For example, fusion of insulin-like growth factor I with the IgG binding domain of staphylococcal protein A allowed for refolding at relatively high concentrations without the need for solubilizing agents (50), and a clear correlation between the efficiency of folding and the hydrophilicity of short N-terminal fusion peptides could be demonstrated for the renaturation of human granulocyte-stimulating factor proteins (51). Thus high concentrations of salt are not strictly required to promote folding, since the TGF-β2 fusion protein could be renatured in 250 mM NaCl, but rather appear to have a solubilizing effect. As mentioned above, the unfolded and intermediate folded forms of DPPC are soluble at low concentrations of salt, without the aid of a fusion partner, suggesting that high salt is required to enhance the solubility of the folded form. Second, DPPC seems competent to fold under conditions of low salt, due to the following similarities with the productive reactions in high salt: (i) the amount of residual monomer is equivalent, indicative of consumption in the reaction, (ii) the folding intermediate is formed, and to a similar extent, and (iii) the folded dimer is indeed produced, albeit at significantly reduced efficiency. Together these observations suggest that the dimer is refolded just as efficiently at low to moderate salt concentrations as at high but is insufficiently soluble under these conditions to be quantitatively recovered. Third, the reduced DPPC monomer appears to adopt a native-like, dimeric state in the absence of high salt, as determined by fluorescence spectroscopy and sedimentation equilibrium ultracentrifugation analyses in 0.1% trifluoroacetic acid, pH 2.3. Therefore, high concentrations of NaCl are not required to promote folding by collapsing the monomer or inducing structure, in keeping with the efficient folding of the TGF-β2 fusion protein observed in low salt and consistent with a singular role in solubilizing the dimeric product.

Recovery of unfolded monomer from the folding reaction was independent of temperature (cf. Fig. 2, A and D), just as recovery of monomer was independent of the concentrations of NaCl at 0.25 M and above. However, production of dimer was strongly dependent on the temperature of the 3-day incubation. Low temperature (5–15 °C) was strictly required for efficient production of the dimer, which was present in only limited amounts in reactions incubated at near ambient temperature and higher (25–35 °C). However, the dimeric form is indeed produced at these temperatures to some extent, as is the case for reactions containing low concentrations of NaCl, again suggesting that recovery is singularly affected rather than folding. In addition, the levels of residual unfolded monomer after incubation at high temperature are equivalent to those of productive reactions (low temperature and high salt), consistent with the notion that the less productive conditions allow the reactions to proceed nearly equally but have disparate effects on the solubility of the product. The positive effect of temperatures in this range (5–15 °C) on the solubility of many proteins is well documented and has a physical basis due to the increased tendency of water to form clathrate-like structures around non-polar molecules. This water-ordering effect of lower temperatures results in a decrease in the free energy of transfer of a non-polar molecule into water, ΔG_T, a measure of the magnitude of the hydrophobic interaction (47). Interestingly, the optimum temperature of folding proteins from mesophiles in vitro is 10–15 °C (52), which coincides well with the optimum observed for the C-terminal domain of DPP. However, low temperature presumably aids in the solubilization of intermediately folded forms of most mesophilic proteins, reducing aggregation in the absence of solubilizing agents, whereas low temperature appears to be required for solubilization and recovery of the product, not the intermediate forms, of TGF-β superfamily members folded in the presence of high salt and a membrane protein-solubilizing detergent. Thus two of the factors found to be crucial for production of C-terminal dimers in vitro, high concentrations of NaCl and low temperature, appear to influence the solubility of the folded form, not the unfolded or partially folded forms as is commonly the case (53, 54).

The insolvibility of non-native forms is generally attributed to intermolecular interaction between hydrophobic side chains, inaccessible to solvent in the core of the native form, but transiently exposed to solvent during the folding process, leading to aggregation at high protein concentrations or in the absence of solubilizing factors. In this regard, it is particularly interesting that the folded dimeric product DPPC2 is actually more hydrophobic than the unfolded or misfolded forms, as shown by reverse phase HPLC (cf. Fig. 3C). The non-native forms elute in a broad shoulder preceding the peak of folded dimer, which is released last from the hydrophobic matrix by the gradient of
organic solvent. In contrast, the native forms of the vast majority of proteins are more hydrophilic than their partially folded forms and elute significantly earlier. This pronounced hydrophobicity appears to be responsible for the insolubility of DPPC$_2$ and TGF-ß superfamily members in general. The DPP C-terminal dimer not only requires solubilizing conditions to be recovered from the folding reaction as described above, but as well must be maintained in chaotropic aqueous buffers or in aqueous organic mixtures throughout purification and during storage and handling (cf. Table I), a shared property of BMP (28, 55, 56) and TGF-ß (32, 33) ligands. At relatively high concentrations (2.5 mg/ml) in a non-chaotropic, aqueous buffer (10 mM sodium acetate, pH 3.5), the DPP dimer was soluble but highly aggregated, so much so that the size of the particles could not be estimated by dynamic light scattering.

Mittl et al. (57) have noted that the surface of TGF-ß3 is mainly hydrophobic, which provides the structural basis for its low solubility in aqueous solvents. A similar analysis of the structures of the other TGF-ß superfamily members in the Brookhaven Protein Data Bank, the refined crystal structures of TGF-ß2 (58, 59), the solution structure of TGF-ß1 (60), and the crystal structure of BMP-7 (18), demonstrates that this is a conserved structural feature and accounts for the general tendency for these ligands to aggregate under non-denaturing conditions. The hydrophobic surface is composed of two symmetry-related pairs of patches that wrap around the dimeric domains in a nearly contiguous fashion, as demonstrated by the crystal structure of BMP-7 and a homology-based model of the DPPC$_2$ (Fig. 8). Thus, both with respect to surface structure and solubility properties, the mature ligands of the TGF-ß superfamily are strikingly similar to membrane proteins, which typically are surrounded by a broad band of hydrophobic surface for insertion into the non-polar lipid bilayer and extremely insoluble in aqueous buffers without the aid of detergents. The preference for membrane protein-solubilizing detergents in the folding reaction is consistent with the observed membrane protein-like properties of these ligands and a singular role for these detergents in enhancing the solubility of the folded product, as hypothesized for the roles of high concentrations of NaCl and low temperatures.

The conditions that appear to be required to solubilize the folded C-terminal domain in vitro are clearly far from physiological, thus a means for solubilizing the mature domain in vivo must have evolved. Gray and Mason (19) have shown that in vivo the prodomains of two TGF-ß superfamily members, TGF-ß1 and activin A, are required as intramolecular chaperones to provide for secretion of the mature domains, as well as to promote folding and disulfide bond formation. Subsequently, Mason and co-workers (61) showed that BMP-2 and BMP-4 also cannot be produced in cell culture in the absence of the prodomain, which appears to serve a similar role as a chaperone for this family of signal ligands. Because aggregates of the folded C-terminal dimer would be degraded before exiting the secretory organelles, the prodomain may allow for secretion of the mature form by enhancing its solubility as a complex of the two domains. After proteolytic processing in the trans-Golgi network by members of the family of subtilisin-like prohormone convertases (furins; see review by Nakayama (62)), two representatives of the TGF-ß superfamily, TGF-ß1 (63, 64) and Mullerian inhibiting substance (65), are secreted from cultured cells and embryonic testes, respectively, as non-covalent complexes. The prodomains of TGF-ß1 (66, 67) and Mullerian inhibiting substance (68) have also been shown to reassociate with their respective signaling ligands in vitro, demonstrating that stable, non-covalent complexes between the N- and C-terminal domains form both in vitro and in vivo. The interface between the two domains almost certainly includes, and thus renders solvent-inaccessible, the hydrophobic surface patches.
of the folded dimer, since the largest or the second largest patch forms the intersubunit interface in 90% of all multimeric proteins (69). Moreover, the burial of exposed hydrophobic surface area due to hydrophobic interaction is suggested to be the main driving force for the formation of these multimeric complexes (70). Similarly, formation of complexes of the N- and C-terminal domains of TGF-β superfamily members may be mediated by, and mask from solvent, hydrophobic surface patches at domain interfaces. Hence, one of the principal roles of the prodomains as intramolecular chaperones may be to enhance the solubility of the folded dimers by masking the large hydrophobic patches from the aqueous solvent, reducing kinetic partitioning into off-pathway aggregates (53, 54) as has been shown for the molecular chaperones of the hsp70 and chaperonin families, and thereby allowing for secretion.

In addition to a requirement for secretion, the prodomain was also shown to be required to promote the folding and the formation of disulphide bonds of the C-terminal domain (19). In vitro folding of the TGF-β and BMP C-terminal domains is sluggish, requiring 3–7 days to reach completion, and thus again far from the physiological situation, since disulphide-containing proteins typically fold and exit the endoplasmic reticulum within 0.5–1.5 h (71). The slow kinetics of TGF-β superfamily members is reminiscent of the in vitro folding behavior of proteases that have also been shown to require their prodomains as intramolecular chaperones, however, not to allow for secretion but to enhance the rate of folding (reviewed by Eder and Fersht (72)). Because they promote folding in a direct and positive fashion by providing essential steric information to the C-terminal domain, Ellis (73) has categorized prodomains that act as intramolecular chaperones as “steric chaperones,” in contrast to hsp70 and chaperonin proteins that are considered “non-steric” because they assist folding indirectly by preventing aggregation of newly synthesized polypeptides. In this context, the prodomains of TGF-β superfamily members appear to function non-sterically, but to prevent aggregation of the folded rather than intermediate forms, and possibly also sterically, to reduce kinetic barriers preventing conversion of reduced native-like states to the oxidized native structure.

In contrast to BMP-2, -4, and DPP signal ligands, which can be efficiently produced in vitro under highly solubilizing conditions, the C-terminal domains of SCREW, 60A, and the vertebrate homolog of 60A, BMP-7, cannot. Thus structural features that are common to the BMP-2 and -4 class of proteins, yet not shared by the BMP-7 class and other groups, appear to engender competency to refold in vitro. One striking difference between these two classes is the nature of the N-terminal polypeptides that extend from the folded core of the C-terminal dimers. The polypeptides of BMP-2, -4 ligands are composed primarily of charged or polar residues and thus hydrophilic, whereas the polypeptides extending from the folded C-terminal domains of BMP-7, 60A, and SCREW are composed of a larger fraction of non-polar residues and thus appear to be less water-soluble in comparison. This observation suggests that perhaps by increasing their solubility, e.g. by fusion to hydrophilic N-terminal peptide partners, the core domains of BMP-7, 60A, and SCREW may be rendered competent to spontaneously refold in vitro.

Although the heparin-binding property of DPP has not been reported previously, high affinity for heparin has been observed for BMP-2, -3, -4, and -7 (28, 55, 56, 74) and is a functional consequence of the primary structure of the N-terminal polypeptide (cf. Fig. 1). Ruppert et al. (28) have shown that the N-terminal arm of human BMP-2 comprises a high affinity heparin-binding site. Clusters of basic amino acids are a common feature of these polypeptides of many BMPs (28), suggesting that high affinity binding is mediated through specific ionic interactions between positively charged protein side chains and anionic carboxylate and sulfate groups of heparin and heparan sulfate glycosaminoglycan chains (reviewed by Hileman et al. (75)). This is in contrast to complexes of low affinity which are presumed to arise from relatively nonspecific ionic interactions. Isolation of both the unfolded monomer and the folded dimer by heparin affinity chromatography in 6 M urea, which would diminish the contributions to binding primarily of hydrogen bonds and hydrophobic interactions, provides evidence that ionic interactions indeed play a significant role in formation of the protein-carbohydrate complex. The thermodynamic parameters of DPPC, binding to low molecular weight heparin have been determined by isothermal titration calorimetry and have shown that the binding affinity is comparable to that of bovine fibroblast growth factor. 4 Because heparan sulfate proteoglycan co-receptors play a key role in mediating the interaction between bovine fibroblast growth factor and its receptor, the in vitro binding studies have significant implications concerning the distribution and activity of DPP in vivo during embryonic development and imaginal disk morphogenesis. Daily, encoding a Drosophila glypican modified with heparan sulfate, has recently been shown to interact genetically with dpp and thus may serve such a co-receptor role (76).

The competition binding assays provide direct biochemical evidence for interaction between the mature DPP signal ligand and Noggin, which had been previously demonstrated indirectly by injection experiments with Drosophila embryos (16). As no Drosophila Noggin homolog has been identified to date, it remains to be determined whether this interaction is only a reflection of the apparently highly conserved structure of BMPs or whether in addition to the Drosophila Chordin homolog encoded by the short gastrulation gene, SOG (77), a second antagonist related to vertebrate Noggin participates in the modulation of DPP activity.

The E. coli-expressed, refolded protein migrated at a reproducibly faster rate through denaturing polyacrylamide gels than the homodimer secreted by the Drosophila S2 embryonic cell line (12, 13), consistent with the observation that all BMPs are post-translationally modified through N-linked glycosylation by insect as well as mammalian cell lines (55, 78). The measured difference in molecular mass of approximately 1.75 kDa per monomer is in good agreement with the calculated molecular mass of a mature, high mannose oligosaccharide (1.4 kDa), indicating that the secreted DPP dimer is also modified by N-linked glycosylation by the Drosophila embryonic cell line. The C-terminal domains of a family of BMPs, which includes DPP, share a consensus N-linked glycosylation site (Fig. 9), shown to be the single site of modification of mature BMP-7 (56). Inspection of the model of DPPC indicates that the modified asparagine side chains reside in the bottom of a large hydrophilic cleft (Fig. 8). The two N-linked oligosaccharides, branched polyols of 8–10 carbohydrate residues, would be expected to fill this cleft and render the secreted dimer somewhat more soluble in non-denaturing, aqueous buffers. However, S2-secreted DPP dimer behaves qualitatively like the non-glycosylated, refolded protein from E. coli, requiring chaotropic aqueous buffers or aqueous organic solvent mixtures for solubilization (13). These modifications also impart heterogeneity onto the population of protein molecules, unlike the bacterially expressed protein which is homogeneous. For many applications, especially biophysical, a homogeneous preparration...
Fig. 9. Conserved N-linked glycosylation site of BMPs. Amino acid sequences encompassing the conserved N-linked glycosylation site of BMPs (shaded box) are aligned with homologous sequences from representatives of the four major protein subfamilies within the DVP (DPP-Vg-Related) group.

Acknowledgments—The following members of the Biozentrum der Universität Basel are acknowledged for their support and assistance: Prof. Kasper Kirschner, Dept. of Biophysical Chemistry, for unlimited access to the HPLC apparatus and UVVIS spectrophotometer and to all of the members of the Kirschner laboratory for their courtesy and helpfulness; Dr. Thomas Kiehfarb, Dept. of Biophysical Chemistry, for critical comments on the manuscript and valuable discussion about protein folding theory; Dr. Paul Jenó and Thierry Mini, Protein Chemistry Facility, for mass spectrometry measurements and N-terminal sequence determinations, respectively; Guido Capitani, Dept. of Structural Biology, for generating a dimeric model by symmetry transformation; Georges Martin, Dept. of Cell Biology, for providing the PSEC10 expression vector and technical expertise about bacterial overexpression and fast protein liquid chromatography; Thomas Marty, Dept. of Cell Biology, for automated fluorescent-deoxyribonuclease sequence analysis; Remo Amherd, Dept. of Pharmacology, for curve-fitting analysis of the chondrogenesis data; Dr. Stefania DiMarco performed the dynamic light scattering analysis while at Ciba-Geigy, Basel, and her effort and expertise are greatly appreciated. Special thanks go to Dr. John McCartney (Creative Biomolecules, Hopkinton, MA) for stimulating discussion about in vitro folding of BMPs and Ray Rischka (University of Wisconsin) for generously providing details about DVP expression in S2 cells.

REFERENCES

1. Padgett, R. W., St. Johnston, R. D., and Gelbart, W. M. (1987) Nature 325, 81–84
2. Newfeld, S. J., Padgett, R. W., Findley, S. D., Richter, B. G., Sanicola, M., de Cuevas, M., and Gelbart, W. M. (1997) Genetics 145, 297–309
3. Ferguson, E. L., and Anderson, K. V. (1992) Cell 71, 451–461
4. Nellen, D., Burck, R., Rühland, K., and Basler, K. (1996) Cell 85, 357–368
5. Leucile, P., Brok, W. J., Mg, Calleja, M., Sun, H., and Cohen, S. M. (1996) Nature 381, 387–393
6. Immerguth, K., Lawrence, P. A., and Brien, M. (1990) Cell 62, 261–268
7. Panganiban, G. E., Reuter, R., Scott, M. P., and Hoffmann, F. M. (1990) Development 110, 1041–1050
8. Reuter, R., Panganiban, G. E., Hoffmann, F. M., and Scott, M. P. (1990) Development 110, 1031–1040
9. Vincent, S., Ruberte, E., Grieder, N. C., Chen, C. K., Haerry, T., Schuh, R., and Affolter, M. (1997) Development 224, 2741–2750
10. Ringe-Dossou, R. J., and Hafen, E. (1997) Biochemistry 36, 267–669
11. Basler, K., Edlund, T., Jessell, T. M., and Yamada, T. (1993) Cell 73, 678–702
12. Panganiban, G. E., Rashka, K. E., Neitel, M. D., and Hoffmann, F. M. (1990) Mol. Cell. Biol. 10, 2669–2680
13. Sampath, T. K., Rashka, K. E., Doctor, J. S., Tucker, R. F., and Hoffmann, F. M. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 6004–6008
14. Brummel, T. J., Twombly, V., Marques, G., Wranza, J. L., Newfeld, S. J., Attisano, L., Massague, J., O'Connor, M. B., and Gelbart, W. M. (1994) Cell 78, 251–261
15. Penton, A., Chen, Y., Staeling-Hampton, K., Wranza, J. L., Newfeld, S. J., Attisano, L., Massague, J., O'Connor, M. B., and Gelbart, W. M. (1994) Cell 80, 239–250
16. Holley, S. A., Neul, J. L., Attisano, L., Wranza, J. L., Sasi, Y., O'Connor, M. B., De Robertis, E. M., and Ferguson, E. L. (1996) Cell 86, 607–617
17. Newfeld, S. J., Mehta, A., Creque, M. A., Wranza, J. L., Attisano, L., and Gelbart, W. M. (1997) Development 124, 3167–3176
18. Griffith, D. L., Keck, P. C., Sampath, T. K., Ruerger, D. C., and Carlsson, W. D. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 878–883
19. Gray, A. M., and Mason, A. J. (1990) Science 247, 1328–1330
20. Studier, F. W., and Moffatt, B. A. (1986) J. Mol. Biol. 189, 113–130
21. Clos, J., Westwood, J. T., Becker, P. B., Wilson, S., Lambert, K., and Wu, C. (1990) Cell 63, 1038–1047
22. Dekker, N., Kinner, J., Tommassen, J., and Verheij, H. M. (1995) Eur. J. Biochem. 232, 214–219
23. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
24. Laemmli, U. K. (1970) Nature 227, 680–685
25. Schagger, H., and von Jagow, G. (1987) Anal. Biochem. 166, 368–379
26. Koenig, B. B., Cook, J. S., Wolsing, D. H., Ting, J., Tiesman, J. P., Correa, P. E., Olson, C. A., Pecquet, A. L., Ventura, F., Grant, R. A., Wranza, J. L., Massague, J., and Rosenbaum, J. S. (1994) Mol. Cell. Biol. 14, 5961–5974
27. Pace, C. N., Vasgos, F., Fee, L., Grimley, G., and Gray, T. (1995) Protein Sci. 4, 2411–2423
28. Ruppert, R., Hoffmann, E., and Sebald, W. (1996) Eur. J. Biochem. 237, 295–302
29. Peitsch, M. C. (1995) Bio/Technology 13, 658–660
30. Peitsch, M. C. (1996) Biochem. Soc. Trans. 24, 274–279
31. Sayle, R. A., and Milner-White, E. J. (1995) Protein Expression Purif. 4, 406–415
32. Cerletti, N., Mcmaster, G. K., Cox, D., Schmitz, A., and Meyhak, B. (June 19, 1997) Development 124, 4461–4470
Refolded Drosophila DPP

57. Mittl, P. R., Priestle, J. P., Cox, D. A., McMaster, G., Cerletti, N., and Grutter, M. G. (1996) Protein Sci. 5, 1261–1271
58. Daspin, S., Li, M., and Davies, D. R. (1993) Proteins 17, 176–192
59. Schlunegger, M. P., and Grutter, M. G. (1993) J. Mol. Biol. 231, 445–458
60. Hinck, A. P., Archer, S. J., Qian, S. W., Roberts, A. B., Sporn, M. B., Weatherbee, J. A., Tsang, M. L., Lucas, R., Zhang, B. L., Wenker, J., and Torchia, D. A. (1996) Biochemistry 35, 8517–8534
61. Hammonds, R. G., Jr., Schwall, R., Dudley, A., Berkemeier, L., Lai, C., Lee, J., Cunningham, N., Reddi, A. H., Wood, W. I., and Mason, A. J. (1991) Mol. Endocrinol. 5, 149–155
62. Nakayama, K. (1997) Biochem. J. 327, 625–635
63. Gentry, L. E., Loubin, M. N., Purchio, A. F., and Marquardt, H. (1988) Mol. Cell. Biol. 8, 4162–4168
64. Miyazono, K., Hellman, U., Wernstedt, C., and Heldin, C. H. (1988) J. Biol. Chem. 263, 6407–6415
65. Nachtigal, M. W., and Ingraham, H. A. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 7711–7716
66. Gentry, L. E., and Nash, B. W. (1990) Biochemistry 29, 6851–6857
67. Böttinger, E. P., Factor, V. M., Tsang, M. L., Weatherbee, J. A., Kopp, J. B., Qian, S. W., Wakefield, L. M., Roberts, A. B., Thorgerisson, S. S., and Sporn, M. B. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 5877–5882
68. Wilson, C. A., di Clemente, N., Ehrenfels, C., Pepinsky, R. B., Josso, N., Vigier, B., and Cate, R. L. (1993) Mol. Endocrinol. 7, 247–257
69. Linaza, J., and Argos, P. (1997) Proteins 26, 335–343
70. Janin, J., Miller, S., and Chothia, C. (1988) J. Mol. Biol. 204, 155–164
71. Lodish, H. F., Kong, N., Snider, M., and Strous, G. J. (1983) Nature 304, 80–83
72. Eder, J., and Fersht, A. R. (1995) Mol. Microbiol. 16, 609–614
73. Ellis, R. J. (1996) Trends Biochem. Sci. 21, 45–45
74. Sampath, T. K., Muthukumar, N., and Reddi, A. H. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 7109–7113
75. Hileman, R. E., Fromm, J. R., Weiler, J. M., and Linhardt, R. J. (1998) BioEssays 20, 156–167
76. Jackson, S. M., Nakato, H., Sugiura, M., Jannuzi, A., Oakes, R., Kaluza, V., Golden, C., and Selleck, S. B. (1997) Development 124, 4113–4120
77. Marques, G., Musacchio, M., Shimell, M. J., Winnenberg-Stapleton, K., Chs, K. W., and O’Connor, M. B. (1997) Cell 91, 417–426
78. Aono, A., Hazama, M., Notoya, K., Takeboshi, S., Yamasaki, H., Tsukuda, R., Sasaki, S., and Fujisawa, Y. (1995) Biochem. Biophys. Res. Commun. 210, 670–677
79. Kingsley, D. M. (1994) Genes Dev. 8, 133–146