Isolation and Molecular Cloning of Transferrin from the Tobacco Hornworm, Manduca sexta

SEQUENCE SIMILARITY TO THE VERTEBRATE TRANSFERRINS

(Received for publication, May 21, 1990)

Neil S. Bartfeld and John H. Law

From the Department of Biochemistry and Center for Insect Science, University of Arizona, Tucson, Arizona 85721

An iron-binding glycoprotein of $M_r = 77,000$ has been isolated from hemolymph of the adult sphinx moth Manduca sexta. Since this protein binds ferric ion both in vivo and in vitro and has a secondary structure similar to that of human serum transferrin and human lactoferrin as judged by CD spectra, we decided to clone its cDNA in order to determine its relationship to the vertebrate transferrins. Antiserum generated against this protein was used to screen a larval fat body cDNA library. A 2.0-kilobase clone was isolated in vivo and in vitro and has a secondary structure judged by CD spectra, we decided to prepare a 2.0-kilobase clone as a probe, three full-length clones were isolated, and the complete nucleotide sequence of one 2,183-base pair insert was determined. The deduced protein sequence contains an 18-amino acid signal sequence and a mature protein sequence of 663 amino acids with a calculated $M_r$ of 73,436. The sequence was used to search the National Biomedical Research Foundation (NBRF) protein database, revealing significant similarity to the vertebrate transferrins, a family of 80-kDa glycoproteins which transport and sequester iron in the blood and other body fluids. A multiple sequence alignment shows the greatest areas of similarity to be around the two iron binding sites, although the insect protein seems to contain only one such functional site. Moreover, 23 of the 24 cysteine residues in the insect protein occupy identical positions as compared with the other transferrins, indicating a similar overall tertiary structure. Comparison of the two halves of the insect sequence indicates that the protein may have arisen as a result of gene duplication. The similarity of the $M. sexta$ sequence to the vertebrate transferrins may provide important clues to transferrin evolution.

Transferrins are transport glycoproteins which permit the effective passage of the relatively toxic and readily hydrated ferric ion through the vertebrate vascular system (Huebers and Finch, 1987). The iron bound by the transferrins is thus maintained in a bioavailable form for use in the synthesis of iron-containing proteins such as hemoglobin and the cytochromes. A number of transferrins have been characterized and sequenced including human serum transferrin (HUTF) (MacGillivray et al., 1983; Yang et al., 1984), chicken ovotransferrin (CHTF) (Jeltch and Chambon, 1982), human lactoferrin (HLTF) (Metz-Boutigue et al., 1984; Rado as quoted in Anderson et al., 1989), and, most recently, a melanotransferrin (HMTF) found in human melanoma cells (Rose et al., 1986). The transferrins contain two domains as determined by x-ray crystallography of HLTF and rabbit serum transferrin (Anderson et al., 1989; Bailey et al., 1988) and exhibit extensive internal sequence homology between these domains. Each domain binds a single iron atom as well as a bidentate anion (Slabach and Bates, 1975). The distribution of disulfide bonds in the transferrins is very well conserved; six common disulfides are found in essentially similar positions in the NH$_2$-terminal domain as are the nine common disulfides in the C-terminal domain (Metz-Boutigue et al., 1984).

Although the structure and function of vertebrate transferrins have been well studied, iron binding proteins from invertebrates have yet to be closely scrutinized. Iron binding proteins from a crab (150 kDa) and a tarantula (80–100 kDa) have been described previously (Huebers et al., 1982; Lee et al., 1978), as was a transferrin-like protein isolated from a tunicate, Pyura stolonifera (Martin et al., 1984). Although all vertebrate transferrins are 80 kDa and bind two ferric ions, the $Pyura$ protein is 40 kDa and binds a single ferric ion. The two-domain vertebrate transferrins are thought to have arisen by duplication of an ancestral gene encoding a single domain (Greene and Feeney, 1968); this duplication has been confirmed for human serum transferrin by Park et al. (1996). Although the existence of the $Pyura$ and invertebrate proteins may be evolutionarily significant, it has not been prudent to speculate on their relationship to each other or to vertebrate transferrins because of the lack of sequence information available.

Although transport of fats and carbohydrates in insects has been well studied (Kanost et al., 1990), the transport of micronutrients such as iron has yet to be carefully examined. Insects undoubtedly require a large supply of iron to be used in the cytochrome heme structure of their highly aerobic muscle system. Since insect tissues are known to contain ferritin (Huebers et al., 1988; Nichol and Locke, 1989), what would be needed to complete the iron storage system in insects is a hemolymph transport protein that could receive iron at the gut and transport it to the tissues for storage in the form of ferritin.

The abbreviations used are: HUTF, human serum transferrin; PBS, phosphate-buffered saline; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; HLTF, human lactoferrin; CHTF, chicken ovotransferrin; HMTF, human melanotransferrin; MSTF, $M. sexta$ transferrin.

Received for publication, May 21, 1990

* This work was supported by Grant GM 29238 from the National Institutes of Health. The costs of publication of this article were therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) M96296.
Virtually all of the major hemolymph proteins of the tobacco hornworm, Manduca sexta, have been isolated and characterized (Kanost et al., 1980). An 80-kDa glycoprotein was tentatively identified as a transferrin because it had a brownish color and bound iron (Kanost et al., 1980; Bartfeld and Law, 1990). While we were initiating the isolation and characterization of this protein, Huebers et al. (1998) reported some of the properties of the same protein from larvae including its involvement in iron transport. In this paper we will present evidence that this protein is an insect transferrin based upon sequence comparisons and structural and functional data. This is the first reported sequence of such a protein and it bears a clear structural relationship to the vertebrate transferrins.

**Materials and Methods**

**Animals—**Adult M. sexta were raised from eggs supplied by Drs. J. P. Reinecke and J. S. Buckner (U. S. Department of Agriculture, Fargo, ND). After hatching, the larvae were raised as described previously (Prasad et al., 1986a).

**Isolation of Transferrin—**100 adult animals were bled by the flushing out method (Chino et al., 1987) using phosphate buffered saline (PBS, 0.1 M sodium phosphate, 0.15 M NaCl, pH 7.0) containing 0.01 M diisopropylfluorophosphate, 0.05 M glutathione as bleeding solution. The hemolymph was subjected to KBr density gradient ultracentrifugation at 50,000 rpm for 4 h at 5°C to remove lipoproteins (Shapiro et al., 1984). The lipoprotein-free subnatant was collected, concentrated to 10 ml by ultrafiltration through an Amicon YM-10 membrane, and applied to a Bio-Gel A 1.5 (Bio-Rad) gel filtration column (120 × 2.6 cm, flow rate 16 ml/h) equilibrated with PBS. Fractions (3.5 ml) were collected and the eluted protein monitored by absorbance at 280 nanometers. Aliquots of selected fractions were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970) on 4-15% gradient slab gels in order to identify transferrin-containing fractions as judged by the presence of the 77-kDa band. These fractions were pooled, concentrated to 5 ml, and applied to a Sephadex G-100 (Pharmacia LKB Biotechnology Inc.) gel filtration column (165 × 1.2 cm, flow rate 15 ml/h) equilibrated with PBS, and 4 ml fractions were collected. The appropriate fractions were pooled, concentrated, and dialyzed against 0.02 M Tris-HCl, pH 8.6, 0.1 M NaCl. The sample was applied to a Cibacron blue 3GA-agarose (Sigma) column (12 × 1.4 cm, flow rate 25 ml/h) equilibrated with the same buffer. The column was washed with equilibration buffer to elute unbound protein followed by elution of bound protein with a 0.1-0.5 M NaCl linear gradient (100 ml).

**Transferrin-containing fractions were pooled, concentrated, and di-**

**alized against 0.02 M Tris-HCl, pH 7.4, 0.5 M NaCl, 1 mM MgCl2, 1 M NaI.**

**The sample was applied to a concanavalin A-Sepharose (Pharmacia) column (0.8 × 7 cm, flow rate 30 ml/h) equilibrated with the same buffer. Bound protein was eluted with a 0.5 M methyl-D-mannopyranoside (Sigma) linear gradient (50 ml). The appropriate fractions were pooled, concentrated, and dialyzed against 0.01 M sodium phosphate, pH 6.6, 0.1 M NaCl. The sample was applied to a Bio-Gel-HT (Bio-Rad) hydroxyapatite column (1 × 17 cm, flow rate 10 ml/h) equilibrated with the same buffer. Bound protein was eluted with a 0.01-0.08 M sodium phosphate linear gradient (100 ml) which resulted in a homogeneous preparation of transferrin.

**Iron Binding Studies—**(a) *In vivo* experiment; 10 µCi of $^{59}$FeSO4 (10 µCi/µg) was injected into three adult animals, and they were bled 10 min later. Ten µl of hemolymph (125,000 cpm) were subjected to nondenaturing gel electrophoresis on 4-20% gradient slab gels at 4°C for 1,700 V-h and stained with Coomassie Blue. The gel was dried and subjected to autoradiography using Kodak X-Omat AR film and an intensifying screen at -80°C. (b) *In vitro* experiment; 0.2 mg of transferrin was dissolved in 1.5 ml of 0.01 M EDTA, 0.1 M NaCl, pH 5.5 (Coehran et al., 1984) in order to remove bound iron and incubated for 30 min at room temperature in a 0.5% Con Can cell (Amicon Corp.) followed by centrifugation through a YM-10 membrane. A 50-µl sample (Amicon Corp.) containing the membrane three times with 0.05 mM sodium citrate, pH 5.5, $^{59}$FeSO4 (2.5 µCi in 0.01 M HCl) was added and the pH slowly raised to 7.5 with 0.01 M NaHCO3 followed by a 30-min incubation at room temperature. The sample was centrifuged and the membrane washed three times with 0.02 M Tris-HCl, pH 8.0. A 50-µl sample (120,000 cpm) was analyzed by nondenaturing gel electrophoresis as described above.

**Amino Acid Analysis—**Amino acid data were obtained by analysis of duplicate 50-µg samples hydrolyzed in vacuo at 110°C in 6 N HCl for 24, 48, and 72 h. Samples were analyzed using a Beckman 7300 amino acid analyzer.

**Carbohydrate Analysis—**The colorimetric phenol-sulfuric acid assay of (Dubois et al., 1956) was used to estimate the weight percentage of carbohydrate in transferrin. The assay was performed on duplicate samples (232 and 255 µg of protein). The carbohydrate composition was determined from 2.6 mg of protein by hydrolysis of the oisiosaccharide chains in 2 N trifluoroacetic acid. The hydrolysate was subsequently converted to alditol acetates (Grimes and Gregor, 1976) and analyzed using a Hewlett-Packard model 5770 A gas chromatograph.

**CD Spectrum—**CD measurements of a 0.04 mg/ml protein solution in double-distilled water were performed in a 1-cm path length quartz cell at 25°C under a stream of nitrogen using a Cary 60 DS spectropolarimeter with an Aviv modification. Five repetitive scans of the protein were performed and corrected for background measurements.

The data were analyzed by the method of Chang et al. (1978).

**Production of Antiserum—**Purified transferrin (0.5 mg) was subject to SDS-PAGE to remove any traces of impurities which may have been present following excision of the transferrin band and homogenization in 1 ml of Ribi adjuvant (Ribi Immunocchemicals). The homogenate was administered subcutaneously into the hind and forelimbs of a female New Zealand White rabbit. Booster injections (0.2 mg/l of transferrin) were given subcutaneously in Ribi homogenate given every two 2-week intervals. The animal was bled through the main ear vein 6 weeks after the initial injection. Antibody specificity was determined by immunoblotting (Burnette, 1981) using the Vectastain ABC peroxidase immunodetection kit (Vector Laboratories).

**Screening and Isolation of Recombinant Phage—**The anti-transferrin antibody was used to screen a 4th day fifth instar larval fat body cDNA library in vector λZAP II (Stratagene) kindly provided by Dr. Eric Henneman (University of Arizona). Positive plaques were detected with goat anti-rabbit IgG conjugated to alkaline phosphatase (Bio-Rad). Approximately 120,000 plaques were screened, and putative positive ones were re-screened until homogeneous. The phages were excised in vivo according to the Stratagene protocol, digested with EcoRI, electrophoresed on a 1% agarose gel, and stained with ethidium bromide in order to determine insert size. The 1.4-kilobase restriction fragment was labeled by nick translation (Bethesda Research Laboratories) with [*-32P]dCTP (ICN Radiochemicals) and used as a probe to rescreen the library.

**RNA Isolation—**Five individual fourth larval fat bodies were dissected out with ice cold PBS and stored at -80°C until needed. Total RNA was extracted from fat body by the guanidinium isothiocyanate method (Chirgwin et al., 1979). Poly(A)+ RNA was prepared from total RNA by oligo(dT)-cellulose chromatography (Aviv and Leder, 1972).

**Hybrid-select Translation—**Hybrid-select translation was performed according to Miller et al. (1983) as modified by Kanost et al. (1989) using 20 µg of larval fat body poly(A)+ RNA. Total and hybrid-selected poly(A)+ RNA were added to a cell-free rabbit reticulocyte lysate in vitro translation system (Fromage) using [*-35S]methionine (Du Pont-New England Nuclear). Samples were then either subjected to SDS-PAGE or immunoprecipitated. Immunoprecipitations were performed using Pansorbin Staphylococcus aureus cells (Calbiochem) according to Cole et al. (1987).

**DNA Preparation—**Small-scale plasmid DNA isolation was performed using the alkaline lysis method described by Birnboim et al. (1983). Large-scale plasmid DNA isolation was performed by alkaline lysis followed by ultracentrifugation in a cesium chloride density gradient in the presence of ethidium bromide (Ausubel et al., 1987).

**Hybridization of the cDNA—**The 1.4-kilobase DNA insert was determined by the chain-terminating dideoxy method (Sanger et al., 1977) using a modified T7 DNA polymerase (Tabor and Richardson, 1987) (United States Biochemical Corp.). Overlapping exonuclease III deletions were generated on both strands of the insert according to the procedure of Henikoff (1986). Single-stranded DNA was prepared by a modification of the procedure of Denve et al. (1983). Double-stranded DNA for sequencing was prepared by digestion of plasmid minipreps with ribonuclease A followed by phenol extraction and ethanol precipitation. DNA (3-5 µg) was denatured with 2 M NaOH, 0.002 M EDTA and subsequent ethanol precipitation.
RESULTS AND DISCUSSION

Purification of Transferrin—The first step in the purification, KBr density gradient ultracentrifugation, separated the main lipoprotein, lipophorin, from the remainder of the hemolymph proteins (Shapiro et al., 1984). Gel filtration chromatography on Bio-Gel A 1.5 (Fig. 1A) removed the yolk precursor protein, vitellogenin (Osir et al., 1986a), as well as residual lipophorin. Gel filtration on Sephadex G-100 (Fig. 1B) allowed separation of several proteins in the 21-45-kDa range as well as apolipophorin III, an 18-kDa dissociable component of lipophorin (Kawooya et al., 1984). At this stage in the purification, the main protein remaining was the 84-kDa blue-colored biliprotein tetramer, insecticyanin (Riley et al., 1984), and a post-larval protein (Ryan et al., 1988). Insecticyanin was useful in monitoring transferrin-containing fractions from gel filtration columns since the two proteins co-eluted. Cibacron blue dye affinity chromatography removed these two proteins by selectively binding transferrin (Fig. 1C). Chromatography on concanavalin A-Sepharose (Fig. 1D) removed all of the minor bands with the exception of a 50-kDa contaminant which was removed by chromatography on hydroxylapatite (Fig. 1E). The final preparation is shown in Fig. 2.

Characterization of Transferrin—The sequence of the N-terminal 34 amino acids has been determined previously by Edman degradation of the intact protein (Bartfeld and Law, 1990) and is underlined in Fig. 8. The chemically determined amino acid composition, shown in Table I, is similar to that of human and chicken transferrins, although it is somewhat lower in glycine and cysteine.

Transferrin contained 2% carbohydrate by weight which is at the lower end of the 2–12% range reported for mammalian transferrins. Compositional analysis indicated the presence of mannose and N-acetylglucosamine in a ratio of 5:1, similar to the 9:2 ratio reported for other M. sexta glycoproteins such as arylphorin (Ryan et al., 1985) and vitellogenin (Osir et al., 1986b), but quite different from those found in the mammalian transferrins which are not as high in mannose and often contain sialic acid (Spik et al., 1979). Calculations indicated the presence of a single oligosaccharide chain per transferrin molecule. The number of chains per molecule in mammalian transferrins ranges from one to four.

The CD spectrum of transferrin (Fig. 3) indicated a structure low in a-helix (13%) and high in b-sheet (55%). The CD spectra of human serum transferrin and human lactoferrin also indicate low a-helical content (20 and 27%, respectively) and high b-sheet content (65 and 60%, respectively) (Mazurier et al., 1976) implying similar secondary structures as compared with the insect protein.

The iron binding studies verified that the protein did indeed bind iron (Fig. 4). Huebers et al. (1988) demonstrated that the in vitro labeled protein donated its iron to ferritin when incubated with larval fat body in vitro, thus serving a transport function. In contrast to the vertebrate transferrins, the insect transferrin bound only one ferric ion/protein molecule as determined by radiometric titration of the apoprotein.

Fig. 1. Purification of transferrin. The sequential column profiles used in the purification are shown. A, Bio-Gel A 1.5; B, Sephadex G-100; C, Cibacron blue SGA-agarose; D, concanavalin A-Sepharose; E, Bio-Gel-HT hydroxylapatite. Transferrin-containing peaks, as judged by SDS-PAGE, are indicated by an asterisk. The arrows indicate where the gradients were initiated.
Transferrin from M. sexta

**FIG. 2. SDS-PAGE of purified transferrin.** Lane 1, purified transferrin; lane 2, molecular weight standards: phosphorylase b (92,500), bovine serum albumin (66,200), ovalbumin (45,000), carbonic anhydrase (29,000), soybean trypsin inhibitor (21,500), and lysozyme (14,400).

**FIG. 3. CD spectrum of transferrin.** A 0.04 mg/ml protein solution in double-distilled water was scanned between 190 and 240 nm under a stream of nitrogen.

**FIG. 4. In vivo and in vitro iron binding by M. sexta transferrin.** A, autoradiogram of in vivo labeled transferrin (lane 1) and in vitro labeled transferrin (lane 2). B, nondenaturing gel of purified M. sexta transferrin stained with Coomassie Brilliant Blue. The molecular weight of the radiolabeled proteins is similar to that of native transferrin.

**FIG. 5. Identification of transferrin cDNA insert by hybrid-select translation.** Translation products were analyzed on 4–16% SDS-PAGE gels either directly or following immunoprecipitation. Lane 1, translation products of larval poly(A)+ RNA; lane 2, precipitation of products in lane 1 with immune serum; lane 3, precipitation of products in lane 1 with preimmune serum; lane 4, translation products of hybrid-selected mRNA; lane 5, immunoprecipitation of products in lane 4; lane 6, translation products of mRNA selected by control plasmid; lane 7, 14C-labeled molecular weight markers.

**TABLE I**

| Amino acid composition of M. sexta transferrin |
|-----------------------------------------------|
| Amino acid | Chemical analysis | From cDNA |
|------------|------------------|-----------|
| Aspartic acid and asparagine | 12.6 | 12.2 |
| Threonine | 3.7 | 3.6 |
| Serine | 6.8 | 8.1 |
| Glutamic acid and glutamine | 8.9 | 7.8 |
| Proline | 5.5 | 5.6 |
| Glycine | 5.5 | 5.1 |
| Alanine | 9.4 | 9.0 |
| Valine | 7.9 | 7.8 |
| Methionine | 1.5 | 1.7 |
| Isoleucine | 3.8 | 3.6 |
| Leucine | 9.1 | 8.3 |
| Tyrosine | 3.3 | 3.2 |
| Phenylalanine | 4.7 | 4.5 |
| Histidine | 2.8 | 2.7 |
| Lysine | 7.8 | 7.8 |
| Arginine | 5.4 | 4.1 |
| Cysteine | ND* | 3.6 |
| Tryptophan | ND* | 1.1 |

*ND = not determined.

with $^{59}$Fe$^{3+}$-nitritotriacetic acid (data not shown). This chelate complex has been used extensively to label vertebrate transferrins (Bates and Schlabach, 1973). Huebers et al. (1988) determined that the larval transferrin also bound a single iron atom.

**Site of Synthesis**—The fat body is the site of synthesis of most hemolymph proteins (Kanost et al., 1990). In vitro larval fat body incubation with $[^{58}$S]methionine (Prasad et al., 1986b) indicated that transferrin was synthesized in the fat body (data not shown). In addition, in vitro translation of poly(A)$^+$ RNA from larval fat body produced a 77-kDa protein that was precipitated with transferrin antibody (Fig. 5, lane 2). The presence of a lower band, a doublet consisting of the two subunits of the larval storage protein arylphorin (Kramer et al., 1980), is due to nonspecific precipitation as it was also precipitated by preimmune serum (Fig. 5, lane 3). This is probably an artifact of the immunoprecipitation since the antiserum was specific as judged by immunoblots (Fig. 6). Additionally, no reaction was observed with preimmune serum. We have no information as to whether other insect tissues may also synthesize transferrin.

**Screening of the cDNA Library**—A larval fat body cDNA library was screened with the antibody since the larval protein appeared to be identical to the adult protein based on amino acid composition (Huebers et al., 1988), molecular weight, and immunoreactivity (Fig. 6, lanes 1 and 2). Screening with antiserum resulted in the isolation of six positive clones, five of which appeared to have identical 1.7-kilobase inserts as judged by restriction analysis. The sixth clone had a 2.0-
Transferrin from M. sexta

Fig. 6. Specificity of anti-transferrin antiseraum. Either whole hemolymph or subnatant from KBr density gradient ultracentrifugation was subjected to 4-15% SDS-PAGE followed by transfer to nitrocellulose and immunoblotting with 0.1% antiserum. Lane 1, adult hemolymph; lane 2, larval hemolymph; lane 3, adult subnatant; lane 4, purified transferrin. No reaction was observed with preimmune serum.

Fig. 7. Partial restriction map and sequencing strategy for the transferrin cDNA insert. The open box indicates the sequence encoding the mature protein. The solid box indicates the sequence encoding the signal peptide. The arrows indicate the direction and length of the sequenced fragments. The scale at the bottom indicates the nucleotide position from the 5' end of the insert. The EcoRV and KpnI restriction sites present only in XTFl are indicated by asterisks.

kilobase insert. Since the message size was determined to be 2.3 kb as judged by northern blot hybridization (Bartfeld and Law, 1990), the 2.0-kilobase clone was used as a probe to rescreen the library, resulting in the isolation of three full-length clones. The three clones (AF1, AF2, AF3) were identical as judged by restriction analysis with the exception of closely spaced EcoRV and KpnI sites which were present in XTFl but absent in XTF2 and XTF3 (Fig. 7).

Characterization of the Transferrin cDNA—In order to confirm the identity of the cDNA, it was used to hybrid-select its corresponding mRNA from larval fat body. The plasmid vector lacking an insert was used as a control. When this selected message was translated in vitro, a 77-kDa protein was synthesized which migrated on SDS-PAGE gels at the same position as the transferrin immunoprecipitated from the total translation mixture (Fig. 5, compare lanes 2 and 4). Although there were background bands due to nonspecific binding of other mRNA species to the filters, transferrin was the main protein synthesized. Furthermore, immunoprecipitation of the products in lane 4, in which much less arylphorin was present than in lane 1, resulted in a single band corresponding to transferrin. The control plasmid did not select an mRNA (Fig. 5, lane 6).

Nucleotide Sequence of the Transferrin cDNA and Its Deduced Amino Acid Sequence—The partial restriction map of the 2,183-base pair transferrin cDNA insert is shown in Fig. 7 along with the employed sequencing strategy, and the complete nucleotide sequence is shown in Fig. 8. There is a 5'-untranslated region of 24 nucleotides followed by an initiating ATG codon. The sequence preceding the ATG (TGACC) matches the consensus translation initiation sequence (CCACC) at three out of five positions (Kozak, 1984). The 34 NH₂-terminal residues of the protein predicted by Edman degradation begin at position 19 of the amino acid sequence, indicating the presence of an 18 amino acid signal sequence.
FIG. 9. Multiple sequence alignment. Residue numbers of HUTF are indicated above the alignment. Identical residues at the same position in all sequences are indicated by an asterisk above the alignment. Cysteine residues are printed in lower case and are underlined. Potentiav iron binding residues are indicated by arrows above the alignment.

Table II

Percent identities of the transferrins (according to the alignment shown in Fig. 9)

|       | HUTF | HLTF | CHTF | HMTF | MSTF |
|-------|------|------|------|------|------|
| HUTF  | 61.9 | 52.3 | 43.9 | 28.2 |
| HLTF  | 61.9 | 53.2 | 42.6 | 27.6 |
| CHTF  | 52.3 | 53.2 | 40.6 | 25.6 |
| HMTF  | 43.9 | 42.6 | 40.6 | 26.3 |
| MSTF  | 28.2 | 27.6 | 25.6 | 26.3 |

This is a classic signal sequence in terms of length, the presence of a basic residue near the NH₂ terminus, and a hydrophobic core. The putative signal sequence cleavage site between Ala-18 and Ala-19 conforms to the predicted cleavage site (von Heijne, 1983). There are four consensus N-linked glycosylation sites (Asn-X-Thr/Ser) within the sequence. The TAA termination codon occurs at nucleotide 2,068, thus translation of the sequence from nucleotides 25 to 2,070 produces a 681-residue protein with a calculated Mᵦ of 75,221. After cleavage of the signal peptide the mature protein contains 663 amino acids with a Mᵦ of 73,436. Three AATAAA consensus polyadenylation signals are present 31, 77, and 81 nucleotides downstream from the stop codon. The poly(A) tail is 63, 21, and 17 nucleotides downstream from these recognition sequences, respectively. The most likely signals for polyadenylation are the ones 21 and 17 nucleotides upstream of the poly(A) tail, since the signals are most often present 11-30 nucleotides upstream from the poly(A) tail (Fitzgerald and Shenk, 1981).

The sequence of the PvuII-PvuII restriction fragment was determined to confirm any sequence differences which would explain the absence of EcoRV and KpnI restriction sites in XTF2 and XTF3. XTF2 contained five nucleotide substitutions (Fig. 8) which did not alter any of the encoded amino acids except for residue 316 which was changed from an aspartate to a glutamate. Two of these substitutions altered the EcoRV and KpnI restriction sites. XTF3 contained the same substitutions with the exception of nucleotide 921 which was not changed. Since the cDNA library was constructed from multiple fat bodies, these clones may represent allelic variants of transferrin present in our M. sexta population.

Similarity of the Insect Protein to Vertebrate Transferrins—Comparison of the deduced amino acid sequence with those in the NBRF protein database indicated significant similarity to HUTF, CHTF, and HLTF. The M. sexta sequence (MSTF) was aligned with these sequences as well as the HMTF sequence using the multiple alignment program of Feng and Doolittle (1987) (Fig. 9) in which the sequences are aligned progressively, beginning with the most similar pair. This method produces an alignment which reflects the evolutionary history of the sequences. There were 99 amino acids (15%) conserved among all the transferrins. MSTF was 26-28% identical to the other transferrins (Table II).
of these similarities was assessed by comparing the MSTF sequence with 50 randomly shuffled sequences having the same amino acid composition as either HUTF, CHTF, HLTF, or HMTF using the RDF computer program (K-tuple = 2) (Lipman and Pearson, 1985). The results of this analysis were z values of 33.0 for HUTF, 34.9 for CHTF, 17.1 for HLTF, and 35.5 for HMTF, where z > 10 is considered significant.

The other transferrins exhibit identities ranging from 41 to 62% (Table II), thus MSTF is more distantly related to these proteins than they are to each other. The greatest similarity was seen around the iron binding sites. Another feature of the transferrins is the internal homology exhibited between the two domains. HMTF, HUTF, HLTF, and CHTF exhibit 46, 41, 37, and 33% similarity between domains, respectively. The FASTp alignment (Lipman and Pearson, 1985) of the two putative domains of MSTF is shown in Fig. 10. The sequence exhibited some internal homology (19%), which indicates that the insect protein may also have arisen by gene duplication.

The residues involved in iron binding may be tentatively assigned based on the crystal structures of rabbit serum transferrin and HLTF (Bailey et al., 1988; Anderson et al., 1989) and by comparing the conserved residues at the iron binding sites of transferrins whose complete sequences are known. These residues are thought to be 2 tyrosines, a histidine, and an aspartic acid (Fig. 9). In the NH2-terminal domain, Asp-58 and Tyr-188 are conserved among all five proteins. Another conserved tyrosine is replaced by a phenylalanine (residue 95) in the insect protein, although there is another histidine which may be the actual liganding residue. A conserved histidine is replaced by a glutamate at position 249, although there is another histidine at position 254 which could possibly take its place even though it is not conserved in the mammalian proteins. In addition, Arg-124, which most likely plays a role in binding the bicarbonate ion, is also conserved. In the C-terminal domain, only Asp-392 is conserved (except in HMTF). The remaining putative iron and bicarbonate binding residues have been replaced by other residues (Fig. 9). The definitive assignment of residues involved in iron binding will have to await the determination of a higher resolution x-ray crystal structure.

Another feature of the transferrins is the conservation of disulfide bonds in each domain (Metz-Boutigue et al., 1984). There are six disulfides common to HUTF, CHTF, and HLTF in similar positions in the NH2-terminal domain and nine common disulfides in the C-terminal domain. MSTF contains 24 cysteine residues, a somewhat smaller number than the mammalian transferrins (Fig. 9). However, 23 of these 24 residues are conserved among all five transferrins. This implies a similar overall tertiary structure between the insect protein and the other transferrins. Although the locations of the disulfide bonds in MSTF are unknown, it migrates significantly faster on SDS-PAGE gels in the absence of dithiothreitol than in its presence (data not shown), indicating that it is extensively disulfide-bonded. Cysteine residues corresponding to the six common disulfides in the NH2-terminal domain of the transferrins are found in identical positions in MSTF (Fig. 9). Of the nine C-terminal disulfides common to mammalian transferrins, cysteine corresponding to five of these are present in MSTF.

Analysis of the M. sexta sequence may explain the observation that the protein contains a single iron binding site. This site is most likely in the NH2-terminal half of the molecule for several reasons. 1) There are more conserved putative iron binding ligands in the NH2-terminal half of the molecule. 2) There is a greater sequence similarity to the other transferrins in the NH2-terminal half than in the C-terminal half. 3) The likely conserved disulfide bond distribution in the N-terminal half indicates a more similar folded structure to the mammalian transferrins than the C-terminal half, which potentially lacks four of the conserved disulfides present in the other transferrins, thus the NH2-terminal half may be in a better conformation for binding iron. The relationship of MSTF to the mammalian transferrins, along with the presence of transferrin-like proteins in other arthropods, may provide important clues to the evolution of the transferrins. In addition, isolation of the cDNA will facilitate cloning of the transferrin gene and subsequent studies of its regulation.

Acknowledgments—We would like to thank Dr. Eric Hanneman for providing the cDNA library, Dr. Terry Meyer for help with the CD spectrum, Drs. Michael Wells and GloriaYepiz for assistance with the sequence alignments, Dr. Michael Kanost for all his valuable advice, Dr. Michael Wells for providing laboratory space for the cloning and sequencing experiments, and Mary Medina for animal care.

REFERENCES
Anderson, R. F., Baker, H. M., Norris, G. E., Rice, D. W., and Baker, E. N. (1989) J. Mol. Biol. 209, 711-734
Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., and Struhl, K. (eds) (1987) in Current Protocols in Molecular Biology, John Wiley & Sons, New York, pp. 2.4.3
Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 1412-1414
Bailey, S., Evans, R. W., Garett, R. C., Gorinsky, E., Hannain, S., Horsburgh, C., Jhiti, H., Lindley, P. F., Mydin, A., Sarr, R., and Watson, J. L. (1988) Biochemistry 27, 5804-5812
Bartfeld, N. S., and Law, J. H. (1990) in Molecular Insect Science (Hagedorn, H. N., Hildebrand, J. G., Kidwell, M. G., and Law, J. H., eds) Plenum Publishing Corp., New York, in press.
Baies, G. W., and Schlabaach, M. R. (1975) J. Biol. Chem. 248, 3228-3232
Birnboim, H. C. (1983) Methods Enzymol. 100, 243-255
Burnette, W. N. (1981) Anal. Biochem. 112, 195-203
Chang, C. T., Wu, C.-S. C., and Yang, J. T. (1978) Anal. Biochem. 91, 13-31
Chino, H., Hirayama, Y., Kiyomoto, Y., Downer, R. G. H., and Takahashi, K. (1987) Insect Biochem. 17, 89-97
Chirgin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutner, W. J. (1979) Biochemistry 18, 5294-5299
Cochran, M., Coates, J., and Neoh, S. (1984) FERS Lett. 176, 192-196
Cole, K. D., Fernando-Warnakulasiriya, G. J. P., Bogaski, M. S., Freeman, M., Gordon, J. J., Clark, W. A., Law, J. H., and Welle, M. A. (1987) J. Biol. Chem. 262, 11794-11800
Dente, L., Cesareni, G., and Cortese, R. (1983) Nucleic Acids Res. 11, 1945-1955
Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956) Anal. Chem. 28, 350-356
Feng, D. P., and Doolittle, R. F. (1987) J. Mol. Evol. 25, 351-360
Fitzgerald, M., and Shenk, V. (1981) Cell 24, 251-260
Greene, F. C., and Feeny, R. (1968) Biochemistry 7, 1366-1374
Grimes, W. J., and Gregor, S. (1970) Cancer Res. 30, 3900-3910
Henikoff, S. (1984) Gene (Amst). 28, 351-358
Huebers, H. A., and Finch, C. A. (1987) Physiol. Rev. 67, 520-582
Huebers, H. A., Huebers, E., Finch, C. A., and Martin, A. W. (1982) J. Comp. Physiol. 148B, 101-109
Huebers, H. A., Huebers, E., Finch, C. A., Webb, B. A., Truman, J. W., Riddiford, L. M., Martin, A. W., and Massover, W. H. (1988) J. Comp. Physiol. B 158, 281-300
Jeltsch, J.-M., and Chambon, P. (1982) Eur. J. Biochem. 122, 291-295
Kanost, M. R., Prasad, S. V., and Wells, M. A. (1989) J. Biol. Chem. 264, 965-979
Kanost, M. R., Kawooya, J. K., Law, J. H., Ryan, R. O., Van Heusden, M. C., and Zangerl, R. (1990) Ado. Insect Physiol. 22, 299-385
Kawooya, J. K., Keim, P. S., Ryan, R. O., Shapiro, J. P., Samareswara, P., and Law, J. H. (1984) J. Biol. Chem. 259, 10733-10737
Kozak, M. (1984) Nucleic Acids Res. 12, 857-872
Kramer, R. J., Mundall, E. C., and Law, J. H. (1980) Insect Biochem. 10, 279-288
Transferrin from *M. sexta*

Laemmli, U. K. (1970) *Nature* **227**, 680-685
Lee, M. Y., Huebers, H., Martin, A. W., and Finch, C. A. (1978) *J. Comp. Physiol.* **127**, 349-354
Lipman, D. J., and Pearson, W. R. (1985) *Science* **227**, 1435-1441
MacGillivray, R. T. A., Mender, E., Shewale, J. G., Sinha, S. K., Lineback-Zins, J., and Brew, K. (1983) *J. Biol. Chem.* **258**, 3543-3553
Martin, A. W., Huebers, E., Huebers, H., Webb, J., and Finch, C. A. (1984) *Blood* **64**, 1047-1052
Mazurier, J., Aubert, J. P., Loucheux-Lefevre, M. H., and Spik, G. (1976) *FEBS Lett.* **66**, 238-242
Metz-Boutigue, M.-H., Jollès, J., Mazurier, J., Schoentgen, F., Legrand, P., Spik, G., Montreuil, J., and Jolles, P. (1984) *Eur. J. Biochem.* **145**, 659-670
Miller, J. S., Paterson, B. M., Ricciardi, R. P., Cohen, L., and Roberts, B. E. (1983) *Methods Enzymol.* **101**, 650-674
Nichol, H., and Locke, M. (1989) *Insect Biochem.* **19**, 587-602
Ouir, E. O., Wells, M. A., and Law, J. H. (1986a) *Arch. Insect Biochem. Physiol.* **3**, 217-233
Ouir, E. O., Anderson, D. R., Grimes, W. J., and Law, J. H. (1986b) *Insect Biochem.* **16**, 471-478
Park, I., Schaeffer, E., Sidoli, A., Baralle, F. E., Cohen, G. N., and Zakin, M. M. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3149-3153
Prasad, S. V., Ryan, R. O., Law, J. H., and Wells, M. A. (1986a) *J. Biol. Chem.* **261**, 558-562
Prasad, S. V., Fernando-Warnakulasuriya, G. J. P., Sumida, M., Law, J. H., and Wells, M. A. (1986b) *J. Biol. Chem.* **261**, 17174-17176
Riley, C. T., Barbeau, B. K., Keim, P. S., Kézdy, F. J., Heinrikson, R. L., and Law, J. H. (1984) *J. Biol. Chem.* **259**, 13159-13165
Rose, T. M., Plowman, G. D., Teplow, D. B., Dreyer, W. J., Hellstrom, K. F., and Brown, J. P. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1261-1265
Ryan, R. O., Anderson, D. R., Grimes, W. J., and Law, J. H. (1985) *Arch. Biochem. Biophys.* **243**, 115-124
Ryan, R. O., Cole, K. D., Kawooya, J. K., Wells, M. A., and Law, J. H. (1988) *Arch. Insect Biochem. Physiol.* **9**, 81-90
Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463-5467
Schlachsch, M. R., and Bates, G. W. (1975) *J. Biol. Chem.* **250**, 2182-2188
Shapiro, J. P., Keim, P. S., and Law, J. H. (1984) *J. Biol. Chem.* **259**, 3680-3685
Spik, G., Fournet, B., Cheron, A., Strecker, G., Montreuil, J., Dorland, L., and Vliegenthart, J. F. G. (1979) *Fifth International Symposium on Glycoconjugates*, September, 1979, Kiel, Germany, pp. 21-22, Georg Thieme, Stuttgart
Tabor, S., and Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 4767-4771
von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17-21
Yang, F., Lum, J. B., McGill, J. R., Moore, C. M., Naylor, S. L., van Bragt, P. H., Baldwin, W. D., and Bowman, B. H. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 2752-2756
Isolation and molecular cloning of transferrin from the tobacco hornworm, Manduca sexta. Sequence similarity to the vertebrate transferrins.
N S Bartfeld and J H Law

J. Biol. Chem. 1990, 265:21684-21691.

Access the most updated version of this article at http://www.jbc.org/content/265/35/21684

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/35/21684.full.html#ref-list-1