Changes in the Plasma Proteome of *Manduca sexta* Larvae in Relation to the Transcriptome Variations after an Immune Challenge: Evidence for High Molecular Weight Immune Complex Formation*[S]

Yan He‡¶, Xiaolong Cao‡§¶, Shuguang Zhang‡¶, Janet Rogers§, Steve Hartson§, and Haobo Jiang‡‖

Manduca sexta **is a lepidopteran model widely used to study insect physiological processes, including innate immunity. In this study, we explored the proteomes of cell-free hemolymph from larvae injected with a sterile buffer (C for control) or a mixture of bacteria (I for induced). Of the 654 proteins identified, 70 showed 1.67 to >200-fold abundance increases after the immune challenge; 51 decreased to 0–60% of the control levels. While there was no strong parallel between plasma protein levels and their transcript levels in hemocytes or fat body, the mRNA level changes (*i.e.* I/C ratios of normalized read numbers) in the tissues concurred with their protein level changes (*i.e.* I/C ratios of normalized spectral counts) with correlation coefficients of 0.44 and 0.57, respectively. Better correlations support that fat body contributes a more significant portion of the plasma proteins involved in various aspects of innate immunity. Consistently, ratios of mRNA and protein levels were better correlated for immunity-related proteins than unrelated ones. There is a set of proteins whose apparent molecular masses differ considerably from the calculated M_r's, suggestive of posttranslational modifications. In addition, some low M_r proteins were detected in the range of 80 to >300 kDa on a reducing SDS-polyacrylamide gel, indicating the existence of high M_r covalent complexes. We identified 30 serine proteases and their homologs, 11 of which are known members of an extracellular immune signaling network. Along with our quantitative transcriptome data, the protein identification, inducibility, and association provide leads toward a focused exploration of humoral immunity in *M. sexta*.**

Molecular & Cellular Proteomics 15: 10.1074/mcp.M115.054296, 1176–1187, 2016.

Cell-free hemolymph (*i.e.* plasma) of insects serves as a medium that bathes tissues and cells, stores and transfers metabolites, and allows occurrence of physiological processes. Plasma protein concentrations in various insects range from 10 to 100 mg/ml, which fluctuate during development (1). Different groups of hemolymph proteins include hexamerins acting as amino acid sources for metamorphosis, lipophorins for lipid transport, vitellogenins for embryo development, enzymes (*e.g.* esterases, lipases) for lipid hydrolysis, cytokines for intercellular communications, peptide hormones for endocrine regulation, and carriers of lipid hormones. Additionally, a substantial body of literature is available on proteins involved in immune responses (2, 3). Hemolymph is also a battleground wherein plasma proteins and hemocytes attack invading organisms such as viruses, bacteria, fungi, and parasites (4, 5). Some proteins recognize pathogens and propagate the signals of wounding and microbe invasion, others either act as effectors to stop bleeding, kill the pathogens, or modulate potency and duration of the defense reaction (6–10). Fat body, analogous to vertebrate adipose tissues and liver, is the major source of insect plasma proteins. Several studies have described compositions of insect hemolymph proteomes of the fruit fly, honeybee, mosquito, and silkworm (11–15), but little is known about proteome–transcriptome correlations, posttranslational modifications, or protein complex formation during immune responses.

We have been studying the innate immune system of a biochemical model insect *Manduca sexta*, particularly serine proteases (SPs), serine protease homologs (SPHs), and serpins in larval hemolymph (2). Analogous to blood coagulation and complement activation in mammals, rapid responses around wounded tissues or invading pathogens are mediated by a plasma SP cascade in insects. Accumulating evidence indicates that pattern recognition receptors (PRRs), SPHs,

From the ‡Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078; §Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078

Received August 3, 2015, and in revised form, December 4, 2015
Published, MCP Papers in Press, January 25, 2016, DOI 10.1074/mcp.M115.054296

Author contributions: H.J and S.H. designed the research; S.Z. and J.R. performed the experiment; Y.H., X.C., and S.Z. analyzed data; Y.H., S.Z., S.H., and H.J. wrote the paper.

serpins, phenoloxidases (POs), and antimicrobial peptides (AMPs) may form macromolecular complexes during an immune reaction (16–19). Gene sequences and expression profiles of these proteins are now available as a result of the *M. sexta* genome project and RNA-Seq analyses (6–8, 10). Immuno-transcriptome analyses (20–22) have elucidated sequences and levels of mRNAs that encode defense proteins in *M. sexta* fat body and hemocytes that are analogous to certain human leukocytes. Furusawa *et al.* (23) identified 58 nonredundant proteins in *M. sexta* larval plasma using one- and two-dimensional electrophoresis. We recently published an analysis of the plasma peptidome and identified 138 peptides (arbitrarily defined as M_r <25 kDa to include all AMPs) as well as 130 proteins that remained soluble after 50% acetonitrile precipitation (24). Here, we describe the identification of larger hemolymph proteins, their differential expression after a bacteria challenge, and correlations of induced changes between plasma proteins and the corresponding mRNAs in fat body and hemocytes. We also present evidence for immune-response-related posttranslational modifications and protein–protein interactions based on discrepancies between theoretical protein M_r's and mobility on an SDS-polyacrylamide gel.

EXPERIMENTAL PROCEDURES

Experimental Design and Statistical Rationale—**C**ontrol (C) and **i**nduced (I) plasma samples were collected from *M. sexta* larvae injected with buffer or mixture of bacteria, respectively (24). The samples including biological replicates ($n = 3$, three larvae per replicate) were studied using a gel-LC approach (25) (Fig. 1). Proteins in the gel slices were identified by searching MS/MS data against a sequence database of *M. sexta* proteins. Their levels were quantified using normalized **s**pectral **c**ounts (NSCs) and compared by Student's *t* test to reveal significant differences ($p < 0.05$) between the I and C. Correlations of plasma protein levels and corresponding mRNA levels in fat body (F) and hemocytes (H) were analyzed by ordinary least squares regression using NSCs and normalized RNA-Seq read numbers (NRNs) (22). Possible correlations between mRNA and protein level changes (*i.e.* I/C) were examined using $\log_2(NRN_I/NRN_C)$ and corresponding $\log_2(NSC_I/NSC_C)$ values. NSC distributions of proteins in different gel slices and their theoretical M_r values were examined to identify major posttranslational modifications in certain hemolymph proteins.

Preparation of Cell-Free Hemolymph from Buffer- and Bacteria-Injected M. sexta Larvae—The same plasma samples used in our previous work (24) were analyzed using the gel-LC approach. Briefly, each of day 1, fifth instar larvae was injected with a mixture of *Escherichia coli*, *Micrococcus luteus*, and insoluble β-1,3-glucan from *Alcaligenes faecalis* or with sterile phosphate buffered saline as a negative control. At 24 h after the injection, prolegs of the insects were cut and hemolymph was collected using tubes containing a crystal of 1-phenyl-2-thiourea and 1 mM *p*-aminobenzamidine. After centrifugation, equal volumes of the plasma samples from three immune challenged insects were pooled as induced plasma-1 (IP1). Similarly, control plasma-1 (CP1) was prepared from three larvae injected with the saline. This experiment was repeated twice on different days to obtain CP2, CP3, IP2, and IP3.

SDS-PAGE Separation of Plasma Proteins, In-Gel Trypsinolysis, and Sample Preparation for MS Analysis—The pooled plasma samples (40 μl each) were separately mixed with 6×SDS sample buffer (8 μl) and heated at 95 °C for 5 min. Sixty μg of protein from each sample were loaded onto a 4–15% precast gradient polyacrylamide gel (Bio-Rad, Hercules, CA) and run at constant current of 30 mA for 45 min. After brief staining with Coomassie Brilliant Blue R-250, the gel was destained in 30% methanol and 10% acetic acid. Each of the six lanes was cut into nine gel slices (Fig. 1), and the 54 slices were individually cut into small pieces prior to extensive destaining by 50% acetonitrile (ACN)/50 mM NH_4HCO_3, pH 8.0. The gel pieces were then dehydrated with 100% ACN, dried briefly, and rehydrated with 10 mM tris (2-carboxyethyl) phosphine in 50 mM NH_4HCO_3 for 1 h at room temperature. The reducing solution was replaced with 55 mM iodoacetamide in 50 mM NH_4HCO_3 to alkylate Cys residues for 1 h at 25 °C in the dark. After further rinsing with 50 mM NH_4HCO_3, the gel pieces were dehydrated with ACN and then infiltrated with sequencing-grade trypsin (8 μg/ml) in 50 mM NH_4HCO_3. After digestion for 16 h at 37 °C, trypsinolytic peptides were extracted from the gel pieces with 1% trifluoroacetic acid for subsequent LC-MS/MS analysis.

LC-MS/MS Analysis—The 54 control and induced samples were individually analyzed on an LTQ-OrbitrapXL mass spectrometer (Thermo Fisher Scientific, Grand Island, NY) coupled to a New Objective PV-550 nanoelectrospray ion source and an Eksigent NanoLC-2D chromatography system (24). Briefly, peptides were separated using a 5–40% acetonitrile gradient in 0.1% formic acid performed over 40 min at a flow rate of 300 nl/min. During each 1 s full-range FT-MS scan, the six most intense ions were analyzed via MS/MS in the linear ion trap mode. MS/MS spectra were collected using a trigger threshold of 8,000 counts and monoisotopic precursor selection. Parental ions were rejected for MS/MS if their charge states were not assignable, if they were previously identified as contaminants on blank gradient runs, or if they had already been twice selected for MS/MS. Three to six technical replicates were performed for each biological sample in the LC-MS/MS analyses.

Protein Database, Identification, and Validation—The MS/MS spectra were searched against a high-quality, nonredundant database (denoted here as Msexta_060614.fasta, see supplemental text file) for protein identification. The database consisted of immunity-related protein sequences (6–10) and complete protein sequences deduced from the rectified open reading frames and confirmed by homology and domain search (26). Centroided ion masses were extracted by extract_msn.exe utility from Bioworks (v3.3.1) for database searching with Mascot (v2.2.04, Matrix Science, Boston, MA). The following parameters were used in the search: no charge state deconvolution or deisotoping, trypsin digestion, maximum missed cleavage of 1, parental ion mass tolerance of 5.0 ppm, fragment ion mass tolerance of 0.60 Da, formylation or acetylation of protein N termini, Met oxidization, cyclization of glutamate to pyroglutamate, iodoacetamide or acrylamide derivatization of Cys as variable modifications. Scaffold (v4.2.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide or protein identifications. Protein probabilities were assigned by Protein Prophet algorithm (27), and identifications were accepted if they could be established at greater than 98.0% probability and contained at least two identified peptides (Table S6). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principle of parsimony. Reversed protein sequences of the same database were used as decoys, demonstrating a project-wide false discovery rate of 0.5%. Average values of the unique spectral counts (USCs) in three biological replicates for individual proteins were used to analyze distributions of proteins in different gel slices.

Function Prediction, Statistical Analysis, and Ratio Calculation—The accepted protein identifications were used as queries to search the insect RefSeq collection at National Center for Biotechnology Information NCBI; by BLASTP under default conditions to identify

their homologs in other species in order to predict function. To study changes in protein levels, NSCs for individual proteins in the CP[1] or IP samples were calculated by multiplying their observed SCs by the ratio of the average SC of all matched peptides in all the six samples to the total SC of all matched peptides in the samples. As an example, since the total spectral counts for CP1–3 and IP1–3 proteins in gel were 63,886, 66,646, 64,540, 69,891, 70,899 and 70,538, respectively, the SC value for a protein identified in the $IP_{gel}2$ sample was normalized according to: $NSC_{IP2,\ gel} = SC_{IP2,\ gel} \times (63{,}886 + 66{,}646 + 64{,}540 + 69{,}891 + 70{,}899 + 70{,}538)/(6 \times 70{,}899)$. Peptidomic data from our previous study of the ACN-treated samples (24) were similarly reanalyzed using the improved database (Msexta_060614.fasta). NSCs were analyzed by Student's *t* test to reveal statistically significant differences ($p < 0.05$) between induced and control samples. Changes in protein levels (*i.e.* I/Cs) were calculated by dividing the average NSC values from the induced samples by the corresponding average NSC values from the control samples. In calculating I/C values, when NSCs in the denominator were all zero, their average (0) was replaced with 0.5 instead of 1 (24). Proteins were regarded as up-regulated if I/C $\geq$ 1.67, whereas proteins with I/C $\leq$ 0.60 as down-regulated.

Correlation of Protein and mRNA Levels as Well as Their Changes after the Immune Challenge—Amino acid sequences of the identified proteins were used to search the CIFH09 database (http://darwin.biochem.okstate.edu/blast/blast_links.html) by TBLASTN. CIFH09 is a collection of cDNA contigs assembled form RNA-Seq reads of control (C) and induced (I) fat body (F) and hemocytes (H) (21). Raw numbers of CF, CH, IF, and IH reads assembled into the cDNA contig with the highest sequence identity to a protein query were converted to normalized read numbers (NRNs) (22). To test if there is a direct correlation between mRNA and protein levels, the $\log_2$NRN and corresponding average $\log_2$NSC values were analyzed by ordinary least squares regression in the following pairs: CF-CP, CH-CP, IF-IP, and IH-IP (P for plasma). To test whether a correlation exists between mRNA and protein level changes (*i.e.* I/C), $\log_2(NRN_{IF}/NRN_{CF})$ and $\log_2(NRN_{IH}/NRN_{CH})$ and corresponding $\log_2(NSC_{IP}/NSC_{CP})$ values were examined similarly in terms of $F\text{-}P_{gel}$, $F\text{-}P_{ACN}$, $H\text{-}P_{gel}$, and $H\text{-}P_{ACN}$, respectively.

RESULTS

Proteomics Workflow and Protein Identification—To identify *M. sexta* plasma proteins, especially those involved in innate immunity, we collected hemolymph samples from the larvae injected with buffer or bacteria and analyzed them by the gel-LC-MS/MS approach (Fig. 1) (25, 28). A high-quality, nonredundant protein database (Msexta_060614.fasta) was searched with the MS/MS data for protein identification and spectral counting (Table S1). In the dataset, 217,791, or 13.7%, of the 1,594,513 spectra matched those of trypsinolytic peptides. The matching spectra corresponded to 654

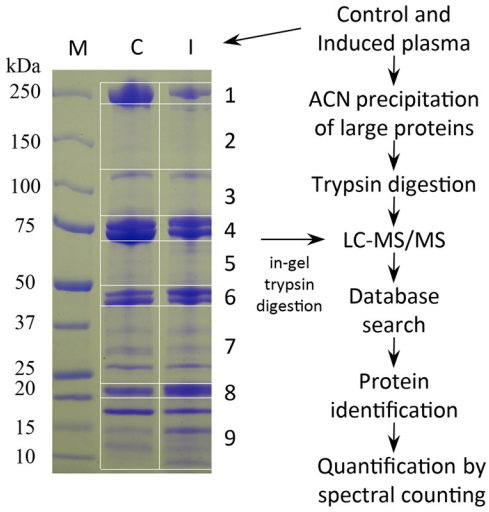

FIG. 1. **Schematic overview of sample preparation and data analysis.** The control and induced plasma samples were separated by SDS-PAGE on a gradient gel that was later cut into nine pieces. Proteins in the gel slices were treated with trypsin and analyzed by nanoLC-MS/MS. As reported previously (24), the same plasma samples were treated with 50% ACN and proteins in the supernatants were analyzed. Protein identification and quantification were performed as described in the Materials and Methods.

independent proteins in the database, and 150 were known or predicted to be defense-related based on homology (Table S2). This represents a dramatic increase in coverage of the *M. sexta* hemolymph proteome, reflecting quality of the protein database, sensitivity increase due to prefractionation, as well as our selection of day 1, fifth instar larvae for injection. Insects at this stage are large enough to provide adequate hemolymph but not too much storage proteins that overwhelm the detection of less abundant proteins.

In this study, we identified a total of 401 proteins with a putative signal peptide, 222 of which were also detected in the peptidome analysis (Table S2) (24). Among the 401, 138 may be involved in defense, including PRRs, SPs (including hemolymph proteases or HPs), SPHs, serpins, and AMPs (6–10); the other 263 may participate in lipid/ion transport and other extracellular functions, based on sequence homology. To our surprise, 253 of the 654 proteins do not contain signal peptide for secretion. Some of them may be related to immunity, including galectin-3; serpin-2; serpin-17B; prophenoloxidase-1 (proPO1); proPO2; plasmatocyte spreading peptide (PSP) binding proteins 1, 6, 7; Eiger; Smt3; and Uev1A. While Eiger possesses a transmembrane region, the other cytosolic proteins (*e.g.* PSP binding proteins and proPOs) may enter plasma via nonconventional secretion pathway or cell rupture (29–31). Consistent with the latter possibility, we detected 44 ribosomal proteins, 7 translation factors, 4 histones, 8 proteasome subunits, and others. Noting that 90% of these "contaminants" had much lower normalized spectral counts (NSCs) in induced plasma (IP) than control plasma (CP) sam-

[1] The abbreviations used are: CP, control plasma from larvae injected with buffer; IP, induced plasma from larvae injected with bacteria; CF, IF, CH, and IH, control (C) and induced (I) fat body (F) and hemocytes (H); NRN, normalized cDNA read number; AMP, antimicrobial peptide; CXE, carboxylesterase; Hp, hypothetical protein; HP, hemolymph protease; PAP, proPO-activating protease; PGRP, peptidoglycan recognition protein; PO and proPO, phenoloxidase and its precursor; PI, protease inhibitor; PRR, pattern recognition receptors; PSP, plasmatocyte spreading peptide; SP and SPH, serine protease and its homolog.

ples, decreased hemocyte lysis after immune challenge may be responsible for the bias.

Quantitative Analysis and Data Comparison of Gel-Fractionated and ACN-Treated Samples—The pairwise Pearson coefficients of the 654 proteins demonstrated excellent data consistency within the **c**ontrol (C, 0.97–0.99 for gel; 0.94–0.96 for ACN) or **i**nduced (I, 0.99 for gel; 0.94–0.96 for ACN) group (Table S3). The C-I coefficients were low (0.55–0.65) for ACN (24), indicating that levels of most ACN-stable proteins differ dramatically between the CP and IP samples. In contrast, the C-I coefficients were much higher (0.93–0.95) for gel, suggesting that most proteins identified in the gel did not change considerably in their abundances after the immune challenge.

We then examined the total number of proteins identified to compare the methods of sample preparation. A total of 654 proteins were identified in the gel-fractionated samples, over twice that (318) in the ACN-treated samples (*i.e.* supernatants of the hemolymph pools after solvent precipitation of larger proteins) (24) (Fig. 2*A*), confirming that ACN precipitated a large number of proteins. The smallest peptide was 6 kDa (cecropin-7) in ACN-treated and 7 kDa (cecropin-6) in gel-fractionated. The largest protein (Dumpy, 1,935 kDa) identified in one of the three ACN-treated samples was larger than that (apolipophorin-like, 475 kDa) in the gel-fractionated samples. In terms of numbers of proteins in different size ranges, the two methods detected 17 (gel) and 18 (ACN) below 10 kDa, 13 of which were identical (Fig. 2*B*). Most proteins fell into the M_r range of 10 to 80 kDa (239 or 75% in ACN; 553 or 85% in gel). In the range of >80 kDa, 60 and 84 proteins were identified in ACN and gel, respectively. The large proteins identified in ACN but not in gel were: C-type lectins X3, X5, and X6; Dumpy; papilin; SP56; PI3KC2A; ROS-like Tyr kinase; and seven others. Their low NSCs suggested that concentration of the ACN supernatants allowed their meagre detection (22). In terms of the ratios of NSC_{ACN} and NSC_{gel} per protein, there was a monotonic decrease from 14.4 to 0.1 (Fig. 2*C*) as M_r increased from <10 to >200 kDa. This indicates that small proteins in the ACN-treated samples are better represented than large ones.

Up- and Down-Regulated Proteins—The levels of 77 proteins (including the gel and ACN data) increased significantly after the immune challenge (Table I). As anticipated, a majority (41, or 53%) of them are related to immune responses. These include eight PRRs: hemolin, immulectin-1, immulectin-12, peptidoglycan recognition protein-3 (PGRP3), PGRP5, Hdd1, hemicentin-2, and Reeler, whose mRNA levels were also up-regulated after the challenge. We detected seven SP-related proteins (HP5, HP17a, HP20, proPO-activating protease-2 (PAP2), PAP3, scolexin A, SPH4) and six protease inhibitors (serpin-2, 5, 11, protease inhibitor-6 (PI6), PI-like protein, dipetalogastin-like Kazal-type PI), which may mediate or modulate the extracellular immune signal transduction. The most dramatic increases occurred in the category of AMPs, including diapausins, attacins, lebocins, cecropins, gloverin, gallerimycin-1, and lysozyme-1. Most of the proteins increased >10-fold, especially the four attacins, which showed >200-fold changes in the ACN and gel samples. This is consistent with the fact that these effectors of the innate immune system are highly up-regulated at the transcription level after immune challenge. Other highly induced proteins include heat shock protein (HSP) 25.4's, lipases, esterases, peptidases, lipid-binding proteins, and some hypothetical proteins. While some of them mediate metabolic changes in response to the stress

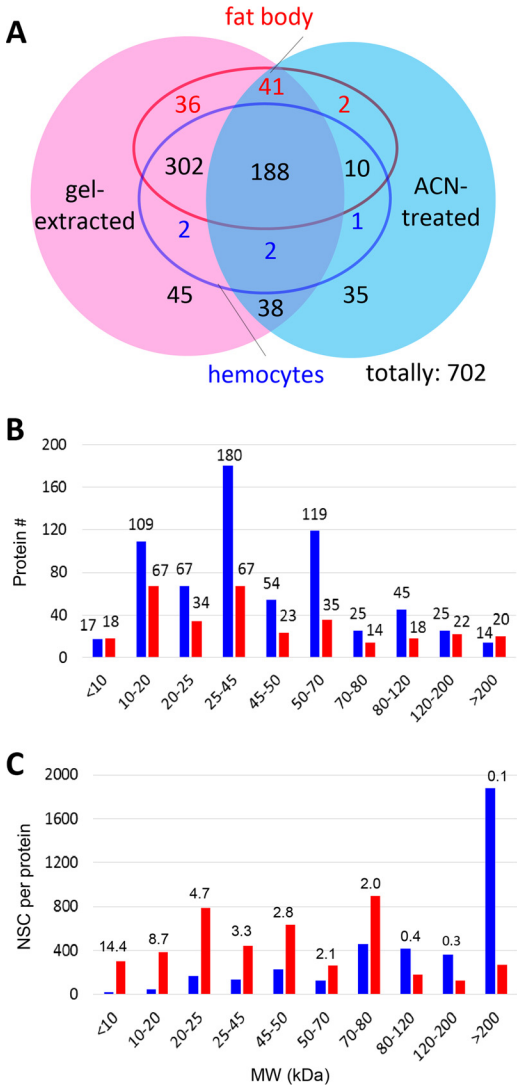

FIG. 2. Comparison of proteins identified by different methods (*A*) and in various M_r ranges (*B, C*). (*A*) Among the 654 proteins identified in the gel-fractionated samples (shaded pink), 502 are expressed in hemocytes (blue circle) and 575 in fat body (red circle). For the 319 proteins in the ACN-treated samples (shaded cyan), 203 are expressed in hemocytes and 243 in fat body. (*B, C*) Size distribution of proteins identified in the gel slices (blue bars) and ACN-treated samples (red bars) with the numbers (*B*) or average NSCs (*C*) of proteins shown as bars. Protein numbers ($\#_{gel}$ and $\#_{ACN}$) and ratios of the NSCs per protein (NSC_{ACN}/NSC_{gel}) in various size ranges are listed on top of the bars in (*B*) and (*C*), respectively.

TABLE I

A list of 77 proteins up-regulated in gel-fractionated and ACN-treated samples after immune challenge

# [a]	Group	Protein name	MW (kDa)	IH/CH gel [b]	p [c]	IF/CF RNA [d]	IH/CH RNA [e]	IH/CH ACN [f]	p [c]
117	AMPs	attacin2	22	407.01	0.00	1443.90	161.66	1983.70	0.00
28		attacin4	24	19.43	0.00	101.98	17.50	3663.78	0.00
236		attacin7	24	1202.23	0.00	106.06	6.53	3257.49	0.00
91		attacin8	24	50.19	0.00	94.33	17.50	3469.22	0.00
68		attacin9	22	471.71	0.00	5207.69	N.A.	2061.03	0.00
287		attacin10	24	500.84	0.03	86.17	9.50	1806.47	0.08
389		cecropin1	7	1.91	0.37	289.99	84.57	9.61	0.00
400		cecropin5	7	N.A.	N.A.	924.34	N.A.	477.79	0.00
413		cecropin6	7	16.65	0.00	205.41	29.16	271.28	0.00
431		diapausin1	8	20.53	0.00	18.12	8.75	227.58	0.01
481		diapausin12	7	8.99	0.03	54.37	N.A.	208.44	0.01
244		diapausin13	7	2.91	0.03	90.62	8.75	3.56	0.02
404		gallerimycin1	8	7.11	0.37	1504.31	8.31	374.18	0.00
134		gloverin	19	4.80	0.00	142.77	239.13	4.08	0.00
205		lebocin-A	38	41.67	0.00	268.97	N.A.	2.66	0.00
252		lebocin-B	16	37.21	0.02	682.68	2.92	13.14	0.00
411		lebocin-C	16	N.A.	N.A.	1685.55	17.50	40.46	0.00
427		lebocin-D	16	N.A.	N.A.	459.92	0.89	677.16	0.00
44		lysozyme_1	16	2.10	0.00	3.57	58.77	1.94	0.00
306	PRRs	IML-1	35	7.25	0.00	4.91	0.86	9.25	0.37
568		IML-12	36	101.41	0.03	332.28	N.A.	N.A.	N.A.
482		PGRP-3	26	21.14	0.00	44.02	14.58	13.88	0.37
261		PGRP-5	22	2.84	0.00	54.37	10.09	10.82	0.00
384		Hdd1	44	49.94	0.00	658.51	40.96	197.35	0.00
120		hemicentin_2	170	4.40	0.01	6.35	1.77	0.00	0.37
10		hemolin	46	10.12	0.00	748.52	3.96	5.43	0.00
237		Reeler1	18	119.79	0.00	1159.95	3.02	487.70	0.00
447	Signal transducers & modulators	HP5	48	2.92	0.04	30.21	5.78	N.A.	N.A.
453		HP17a	67	47.36	0.01	90.62	67.07	N.A.	N.A.
207		HP20	38	2.64	0.00	4.39	2.92	11.62	0.00
297		PAP2	48	125.74	0.00	592.06	5.60	301.77	0.02
104		PAP3	46	4.82	0.00	271.86	3.01	4.14	0.00
45		perlecan	456	0.16	0.05	0.64	1.01	10.70	0.01
257		scolexinA	31	76.86	0.00	5268.11	0.73	440.05	0.00
316		SPH4	43	8.19	0.00	26.85	N.A.	N.A.	N.A.
556		serpin-2	42	18.57	0.00	0.91	17.04	N.A.	N.A.
184		serpin-5	45	47.09	0.00	9.88	3.17	N.A.	N.A.
461		serpin-11	47	40.33	0.00	36.25	6.53	N.A.	N.A.
333		Kazal-type PI, dipetalogastin-like	31	47.50	0.01	120.83	2.92	205.87	0.00
293		PI-like pr.	53	53.31	0.04	197.33	N.A.	368.20	0.00
283		PI6	9	26.30	0.02	839.76	67.07	724.66	0.00
37	Others	antennal esterase	59	1.85	0.00	3.24	34.14	8.22	0.00
77		antennal esterase	59	1.89	0.01	6.63	40.63	2.03	0.04
314		apyrase_2	60	2.33	0.01	1.78	1.34	N.A.	N.A.
436		Calnuc	67	40.42	0.02	2.22	1.28	0.00	0.37
615		cuticle pr. 1	14	16.04	0.00	0.07	0.00	N.A.	N.A.
70		CXE7	63	3.49	0.01	14.07	5.94	1.66	0.00
31		CXE16	62	3.57	0.00	2042.00	8.38	4.31	0.00
337		FK506-BP2	24	3.10	0.04	1.22	1.35	47.40	0.37
100		Hp (hypothetical pr.)	22	3.92	0.02	N.A.	N.A.	2.07	0.09
428		Hp ChlncDraft_55808	10	17.30	0.01	277.90	9.03	156.28	0.07
367		Hp KGM_03936	77	3.71	0.01	0.69	0.72	N.A.	N.A.
250		Hp KGM_06199A	14	68.67	0.00	477.27	N.A.	422.42	0.00
462		Hp KGM_16225	55	13.25	0.00	N.A.	N.A.	N.A.	N.A.
349		Hp KGM_20797	13	9.24	0.00	6.40	5.34	26.29	0.00
102		Hp KGM_21511	175	1.98	0.00	N.A.	N.A.	1.01	0.99
377		Hp, no conserved domain	14	2.19	0.00	4.42	N.A.	3.34	0.32
372		HSP 25.4a	18	49.89	0.02	664.56	N.A.	226.30	0.07
473		HSP 25.4b	18	33.25	0.03	36.71	N.A.	N.A.	N.A.
187		HSP 25.4f	22	4.13	0.00	5.27	N.A.	13.88	0.37
206		HSP 25.4g	23	4.70	0.01	68.43	0.00	115.63	0.15
642		JHBP2, an-0128	26	5.12	0.00	N.A.	N.A.	N.A.	N.A.
176		JHBP, OBP fmxg18C17	26	1.47	0.01	N.A.	N.A.	2.29	0.03
391		lysosomal α-mannosidase-1	83	3.42	0.03	4.35	1.69	N.A.	N.A.
398		mitochondrial aldehyde DH	53	4.22	0.04	0.50	3.80	N.A.	N.A.
323		molting fluid carboxypeptidase A	54	13.30	0.00	0.51	1.11	10.53	0.37
549		myosin family pr.	78	11.54	0.04	3.29	0.19	N.A.	N.A.
613		neuroendocrine convertase 1-like	78	12.21	0.02	151.04	2.92	N.A.	N.A.
224		neutral lipase	38	1.74	0.00	6.03	4.75	1.85	0.04
137		nidogen	173	2.14	0.01	1.62	3.50	N.A.	N.A.
480		proliferation-ass. pr. 2G4	42	4.62	0.04	1.02	1.26	N.A.	N.A.
319		triacylglycerol lipase-1	56	110.87	0.00	0.36	1.19	N.A.	N.A.
376		ubiquitin-activator E1-like	116	16.52	0.04	1.00	1.44	N.A.	N.A.
523		uncharacterized inducible pr. B	10	26.86	0.02	6.12	N.A.	N.A.	N.A.
51		vanin-like pr. 1	64	4.34	0.00	7.94	2.79	N.A.	N.A.
222		yellow-d, MRJP domain	50	2.77	0.00	1.47	N.A.	N.A.	N.A.
368		γ-interferon-inducible lysosomal thiol reductase-like1	27	59.57	0.00	N.A.	N.A.	N.A.	N.A.

[a] Protein identification numbers with the intracellular ones shaded green.

[b] Ratios of average normalized spectral counts (NSCs) of proteins in the induced (I) and control (C) samples separated by SDS-PAGE (I/C >1.67: red; I/C <0.60: blue).

[c] Results of the Student's t-tests conducted using NSCs of proteins in the I and C samples ($p < 0.05$: green).

[d,e] Relative abundances (RAs) or adjusted read numbers (ARNs) of mRNAs in fat body (IF/CF) and hemocytes (IH/CH) (RA or ARN >5, red; RA or ARN <0.2: blue; N.A.: not available) (21).

[f] Recalculated I/C ratios of average NSCs for proteins identified in the ACN-treated samples (I/C >1.67: red; I/C <0.60: blue). Detailed information of these proteins can be found in Table S2.

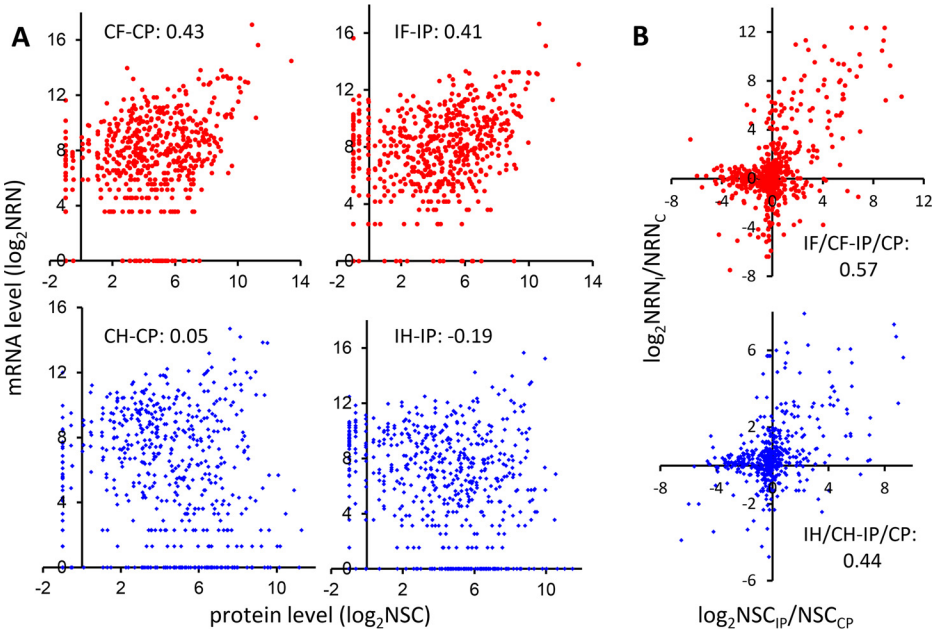

FIG. 3. Correlation of the mRNA and protein abundances (*A*) and their changes (*B*) for gel-fractionated samples. (*A*) Levels. Upper panels: normalized read numbers (NRNs, *y*-axis) of specific transcripts in the control (CF, *left*) and induced (IF, *right*) fat body (red dot) and their corresponding normalized spectral counts (NSCs, *x*-axis) in the control (CP, *left*) and induced (IP, *right*) plasma are plotted in $\log_2$ scale, representing the relative mRNA and protein levels, respectively. Lower panels are the same as *upper panels*, except that the NRNs of mRNA in the control (CH, *left*) and induced (IH, *right*) hemocytes (blue dot) are shown. (*B*) Ratios. The mRNA and protein level changes are represented by NRN_I/NRN_C of fat body (red dot, *upper panel*) or hemocytes ((blue dot, lower panel) and NSC_{IP}/NSC_{CP} of plasma in $\log_2$ scale, respectively. The correlation coefficients for all the comparisons are indicated in the panels.

condition, functions for the others in immunity remain to be determined. In other words, the quantitative proteomic analysis provided significant new leads for investigation of insect immune responses.

In contrast to the up-regulated proteins described above, the levels of 78 proteins significantly decreased after the challenge (Table S4). Eight proteins are related to defense responses, including diapausin-5, immulectin-9, extracellular leucine-rich repeat protein-8 (ELRRP8), peroxidasin, HP6, HP8, HP21, and metalloprotease inhibitor. The decreases in HP6, HP8, and HP21 (range: 0.69–0.94, average: 0.85) were small based on the gel data, while the other five reduced substantially (range: 0.00–0.56, average: 0.18). Most (35) of the 47 gel proteins were undetected in the ACN-treated samples, whereas 36 of the 37 ACN proteins were identified in the gel-fractionated samples. This anomaly may be due to their property differences and operational variances. The gel samples were obtained after electrophoresis, staining, destaining, and in-gel trypsin digestion; the simple ACN treatment had less operational variations but protein size, stability, concentration, and other properties affect their precipitation by ACN, especially those in the high M_r ranges (Fig. 2*C*) (24).

Proteins versus mRNAs—We compared the proteome results from gel and ACN to those obtained by transcriptome analyses of fat body and hemocytes. For this comparison, I/C ratios of the ACN data (24) were recalculated using the same method for the gel data. Based on RNA-Seq data, 491 of the 654 proteins in gel and 198 of the 317 in ACN were expressed in both fat body and hemocytes (Fig. 2*A* and Table S2). Adding those expressed in only one tissue, the 571 proteins in gel and 244 in ACN and their respective cDNA contigs were identified for retrieving read numbers from CF, IF, CH, and IH (24). Using normalized read numbers (NRNs) from the RNA-Seq analysis (22) and corresponding NSCs, we first compared the fat body mRNA and plasma protein levels after the immune challenge. As shown in Fig. 3*A* (*upper panel*), there was a correlation between CF and CP_{gel} (0.43) and between IF and IP_{gel} (0.41). The positive correlation remained when ACN data were analyzed, but the coefficients reduced to 0.41 (CF-CP_{ACN}) and 0.26 (IF-IP_{ACN}) (Fig. S1*A*, *upper panel*). The correlations were much lower between hemocyte transcripts and plasma proteins. The CH-CP_{gel}, IH-IP_{gel}, CH-CP_{ACN}, and IH-IP_{ACN} coefficients were 0.05, -0.19, 0.00, and -0.25, respectively (Fig. 3*A* and Fig. S1*A*, *lower panel*). The negative or low positive correlations between the transcript and protein levels seemed to result from relatively small contribution of hemocytes to plasma proteomes, as shown in the previous study (24). In contrast, due to the sheer volume of fat body and its high levels of protein synthesis, moderately positive correlations exist between fat body transcript and plasma protein abundances.

We then tested whether there was a stronger correlation between the challenge-induced changes (*i.e.* I/C ratios) in mRNA *versus* protein levels. The scatter plots (Fig. 3*B* and

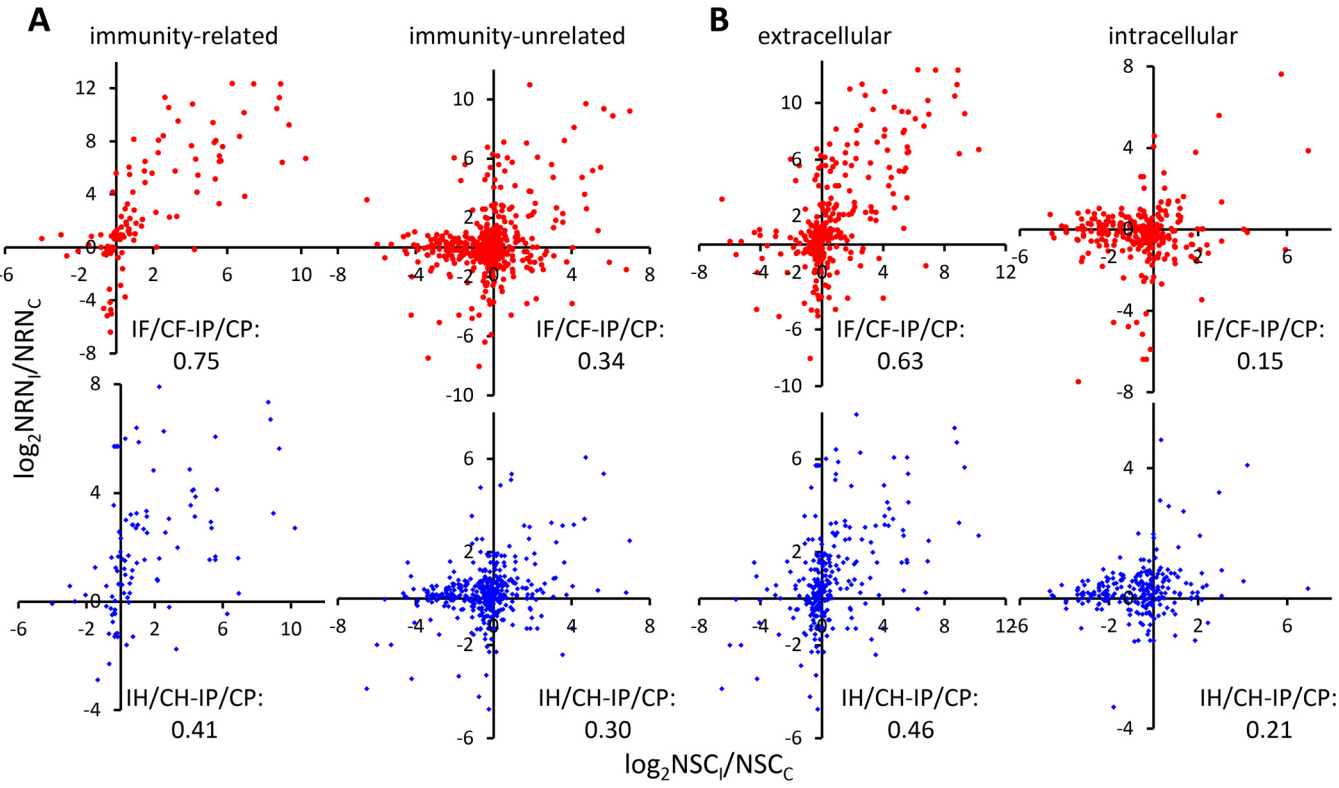

FIG. 4. **Comparison of the mRNA level changes in fat body (red dot, upper) or hemocytes (blue dot, lower) and their protein level changes in the gel samples.** (*A*) immunity-related and -unrelated; (*B*) extra- and intracellular proteins. See Fig. 3 for the definitions of *x*- and *y*-axes.

Fig. S1*B*) clearly demonstrated a positive correlation (0.44–0.69) in both samples: Most gel and ACN proteins showed the same tendency of changes as their mRNAs did, fat body mRNA in particular. Only a small number of proteins showed mRNA level increases but protein level decreases or *vice versa*. Similar to the relationships between protein and mRNA levels, the protein level changes showed a higher correlation to immune-induced mRNA level changes in fat body (gel: 0.57; ACN: 0.69) than in hemocytes (gel: 0.44; ACN: 0.52). These data support the long-standing notion that fat body makes greater contribution to the pool of plasma proteins than do hemocytes. Moreover, this is consistent with the functional disparity between fat body and hemocytes in humoral and cellular immunity (2–4).

Since the above correlations to certain degree reflect gene transcription and translation in fat body and hemocytes between the control and induced samples, we postulated that induced changes in levels of immunity-related mRNA *versus* defense proteins would be more tightly correlated than induced changes in levels of immunity-unrelated mRNA *versus* nondefense proteins. That is, the correlation of changes in immune-related mRNAs *versus* proteins would improve upon removal of nonimmune proteins and their mRNAs from the population. Similarly, we postulated that separating extracellular from intracellular proteins and their mRNAs would lead to similar segregation of correlation coefficients. The correlations for immunity-related genes did increase to 0.75 (*versus* 0.57 for all genes) for fat body, whereas a slight decrease was seen for hemocytes (from 0.44 to 0.41; Fig. 4*A*); those for immunity-unrelated genes decrease to 0.34 (fat body) and 0.30 (hemocytes). This is consistent with the first prediction. On the other hand, the correlations for extracellular proteins slightly increased from 0.57 to 0.63 (fat body) and from 0.44 to 0.46 (hemocytes) (Fig. 4*B*). Those for intracellular proteins dramatically decreased to 0.15 (fat body) and 0.21 (hemocytes). The smaller increase (0.63 *versus* 0.75) is consistent with the fact that extracellular proteins include not only immunity-related (*e.g.* PRRs, SPs, serpins, AMPs) but also immunity-unrelated (*e.g.* lipophorins, storage proteins). While some PRRs and AMPs showed large increases in mRNA and protein levels, their good correlation is partly masked by those unrelated (*i.e.* noises). In contrast, a few intracellular, immunity-related proteins were detected in the plasma samples (Table S2). Our previous studies on *M. sexta* immune signal transducers showed their gene expression changed less dramatically after the challenge (9, 22). The weak correlation is largely covered by the noises of other proteins (*e.g.* ribosomal proteins).

Predicted versus Observed Protein Sizes—In theory, proteins should migrate to positions corresponding to their calculated M_r's after SDS-PAGE and major deviations often suggest posttranslational modifications. By comparing the

calculated M_r's and size ranges of the gel slices in which the proteins were detected, we noticed marked discrepancies—some proteins are detected a few slices away from the expected locations (Table S5). This can be an artifact since protein levels have a major impact on their apparent distributions in gel and hypothetical M_r is not a perfect predictor of electrophoretic migration (Fig. S2). Ten highly abundant proteins (ΣUSC: 6000–500) were detected in nearly all slices, suggestive of diffusion among gel fractions during electrophoresis or staining. Thirty-nine abundant proteins (ΣUSC: 500–150) were mostly located in 5–7 gel slices. A majority of the 105 (ΣUSC: 150–50), 212 (ΣUSC: 50–15), and 289 (ΣUSC: 15–0) proteins were identified in 3–4, 1–2, and 1 gel slice, respectively. Therefore, we reexamined the protein distribution patterns in gel slices and found that abnormal migration of 130 proteins still cannot be explained by diffusion.

Proteolytic activation of SP, SPH, and other protein precursors and its regulation by serpins account for some of the discrepancies, which actually represent the immune signaling, modulation, and execution mechanisms in insect hemolymph (2, 32). Due to their relatively low abundances, the narrow spreads in the gel in many cases reflect proteolytic cleavage or serpin-protease complex formation rather than diffusion. We identified 22 SPs (HP1a, 1b, 2, 5, 6, 8, 9, 14a, 15, 16a, 17a, 19, 20, 21, 22, 25, GP57, GP59, SP34, PAP1–3), 8 SPHs (SPH1b, 2–4, 33, 101, Scolexins A&B), and 17 serpins (1A, 1E, 1K, 1N, 1Z, 2–6, 9, 11–13, 15b, 17b, and 30) in the gel slices (Table S5). Most intact serpins are 45–50 kDa SP inhibitors, and their N-terminal most part (40–45 kDa) form covalent complexes with the catalytic domains (30–35 kDa) of their cognate SPs and then slowly dissociate from the latter (32). Likewise, each hemolymph protease may be detected as a precursor, N- and C-terminal fragments of the active SP, and a 75 kDa complex with the serpin fragment. Among five of the *M. sexta* serpin-1 splicing variants, serpin-1E forms SDS-stable complexes with HP1 and HP8 (33). Serpin-3 through -6 are also known to form complex with HPs, including PAP1, PAP3, HP1, HP6, HP8, and HP21 (34, 35). These findings are consistent with the detection of serpin-1E, -3, -4, and -12 in the range of 25–80 kDa in induced hemolymph (Table II). HP8 and PAP3 identified at 70–80 kDa may represent their catalytic domains in complex with serpin-6 (19). Since most serpin-HP complexes resided in gel slice-4 (70–80 kDa), we were unable to distinguish them or identify their components in this analysis. Serpin-9 displayed interesting changes: its level was higher in naïve plasma and became lower after the challenge (Table II). It existed as an intact protein, with some in a cleaved form and some spectral counts in the higher M_r slices, possibly representing a complex with an unknown HP. In addition, 18 other proteins may be cleaved by proteases, including Cys protease inhibitor (36), lebocin-A (37), cathepsin 26–29 kDa like-2 (38), furin, Eiger, and lacunin (Table S5).

There is a special type of M_r discrepancies unrelated to diffusion, proteolysis, or formation of serpin-protease complexes. Low but considerable spectral counts in gel slices 1–3 (80 to >200 kDa) were detected for HP1a; HP8; HP14a; PAP1; PAP3; SPH2; SPH101; and serpin-1E, -3, -4, -9, and -12. Most of them are involved in the generation of active POs that produce quinone intermediates to crosslink nucleophiles in proteins and polysaccharides (5, 17). We have observed 54 other proteins migrate to gel slices 1–3 (80 to >400 kDa), which are a lot greater than their theoretical M_r's. Among them, β-1,3-glucan recognition protein-1; ELRRP1; hemocyte aggregation inhibitor protein; Hdd1; hemolin; immulectin-2, -4; LRR transmembrane protein-3; apolipophorins III; HSP25.4d; microbe-binding protein; Nimrod B; PGRP1; proPO1; proPO2; proPSP; ML2; attacin-4; gloverin; lysozyme-1; vanin-like-1; and PI-like protein (totally 22) are known or predicted to play roles in the humoral and cellular immune responses. Although it is unclear whether or not the other 32 proteins participate in defense, carboxylesterase (CXE)-7, CXE16, and antennal esterase-1 and -2 were up-regulated considerably. We interpret these 66 proteins as components of high M_r immune complexes crosslinked by covalent bonds, which allow them to sustain the reducing and denaturing conditions of SDS-PAGE.

DISCUSSION

A Dynamic Proteome of the Larval Hemolymph—Here, we used orthogonal fractionation by SDS-PAGE and LC-MS/MS to perform a deep analysis of the plasma proteome in *M. sexta* larvae. By combining results from the gel-free approach (*i.e.* ACN precipitation) (24), we identified a total of 702 proteins (Fig. S2), some with known functions and others with functions predicted based on sequence homology. These findings provide an overview of larval hemolymph proteins in the lepidopteran insect. Among them, 171 (24%) are related to immunity, 77 up-regulated (I/C >1.67), and 430 (61%) extracellular. Additionally, hemocyte lysis seems to contribute 272 intracellular proteins to the plasma.

Hemolymph SPs and SPHs mediate extracellular signal transduction and pathogen killing (2). Understanding their functions and inhibitory regulation in *M. sexta* is useful for investigating similar systems in insect vectors of human diseases. However, these systems are formidably large based on genome information. For instance, there are 193 SP/SPH and 32 serpin genes in the *M. sexta* genome ((8) and Kanost *et al.*, unpublished data). Identification of the 30 SPs (eight in the ACN supernatants), 8 SPHs, and 13 serpins substantially narrows the scope of our exploration to those detected in plasma. A focused research is expected to facilitate the elucidation and reconstitution of the SP-SPH system and its regulation by serpins in *M. sexta*.

Another proteome-level finding is that 272 (39%) of the 702 plasma proteins do not contain a signal peptide, and their abundances decreased considerably after the immune challenge. Since the CP and IP samples were prepared under the same conditions, this is unlikely to be a technical artifact.

TABLE II

Distribution of selected immunity-related proteins in different gel slices

Pr. *Mr* (kDa)	Name \ Gel Slice *Mr* range (kDa)	Control 500–200	200–120	120–80	80–70	70–50	50–45	45–25	25–20	20–10	Induced 500–200	200–120	120–80	80–70	70–50	50–45	45–25	25–20	20–10	Sum of USC Control	Induced
500–200	hemocytin	148	338	253	48	79	10	24	0	1	117	220	183	29	44	4	10	0	0	901	607
	perlecan	76	141	43	3	18	0	5	0	0	6	21	5	0	6	0	12	0	2	285	51
	hemicentin 1	105	69	8	0	0	0	0	0	0	111	64	19	0	0	0	0	0	0	182	193
	lacunin	1	18	35	0	0	0	0	0	0	4	21	52	0	0	0	0	0	0	54	77
	multicystatin	0	3	0	3	19	0	1	1	22	0	3	3	5	17	3	1	1	20	49	53
200–120	hemicentin 2	0	31	0	0	0	0	0	0	0	22	95	7	0	1	0	0	0	0	31	125
	peroxidasin	0	59	20	0	0	0	0	0	0	0	46	2	0	0	0	0	0	0	79	48
	thrombospondin 1	0	28	26	0	0	0	5	0	0	0	30	7	0	0	0	12	0	0	60	48
	LRR protein	16	102	68	68	114	61	95	45	74	26	102	78	59	117	48	79	39	61	641	608
120–80	hemicentin 4	0	8	147	0	11	1	0	0	0	0	19	140	3	15	1	0	0	0	167	178
	proPO2	0	18	51	154	79	34	40	0	2	0	27	43	123	63	31	36	1	2	378	325
80–70	NimB	6	16	56	29	22	12	8	13	12	7	15	53	23	20	9	8	9	4	172	147
	transferrin1	27	120	232	330	199	121	128	33	31	59	154	278	375	225	169	167	47	55	1221	1527
	proPO1	0	10	50	172	122	28	43	0	1	0	18	42	163	77	17	30	0	1	425	348
	ELRRP9	0	0	0	2	65	0	0	0	0	0	0	0	19	51	0	0	0	0	67	70
	HP14a	0	0	1	26	59	0	0	0	0	0	1	15	29	62	0	1	0	0	86	108
70–50	HP17a	0	0	0	0	0	0	0	0	0	0	0	0	5	9	0	2	0	0	0	15
	ELRRP8	0	0	0	3	14	0	0	0	0	0	0	0	0	1	0	0	0	0	17	1
	bGRP2	0	0	0	0	29	0	0	0	0	0	0	0	0	53	0	1	0	0	29	53
	HP19	0	0	0	0	18	0	0	0	0	0	0	0	10	24	0	0	0	0	18	34
	PPBP1	0	0	0	0	3	5	19	0	0	0	0	0	0	0	43	10	0	0	72	53
	serpin-3	0	0	0	8	100	75	58	14	2	0	18	42	52	97	79	57	17	17	257	379
	HP16a	0	0	0	0	36	19	2	0	0	0	0	0	0	58	30	12	0	0	57	99
	MBP	0	0	0	0	54	2	8	0	0	0	4	10	3	64	9	21	0	1	64	111
	serpin-13	0	0	0	0	80	4	13	0	0	0	0	0	0	67	1	8	0	0	96	75
	bGRP1	0	0	8	11	82	18	26	0	0	0	1	4	1	68	16	23	1	0	146	114
	Eiger	0	0	0	0	0	0	7	12	9	0	0	0	0	0	0	12	13	9	28	34
50–45	HP21	0	0	0	1	42	0	0	0	0	0	0	0	0	32	0	1	0	0	42	33
	HP2	0	0	0	0	6	0	0	0	0	0	0	0	0	5	0	0	0	0	6	5
	serpin-9	0	0	6	1	4	13	14	0	0	0	0	0	0	0	3	0	0	0	38	3
	ELRRP2	0	0	0	0	3	40	1	0	0	0	0	0	0	0	29	0	0	0	44	29
	HP25	0	0	0	0	0	20	0	0	0	0	0	0	0	1	20	0	0	0	20	20
	serpin-11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	1	0	0	0	14
	serpin-5	0	0	0	0	0	0	3	0	0	0	0	0	0	0	45	45	0	0	3	90
	serpin-30	0	0	0	0	0	10	27	0	0	0	0	0	0	0	4	24	0	0	36	28
	HAIP	1	22	38	25	115	155	83	46	55	2	24	34	19	109	153	82	52	43	539	519
	ELRRP1	0	8	16	4	57	113	80	24	21	0	9	12	1	53	128	74	25	15	322	317
	hemicentin-5	0	0	1	0	31	95	68	10	18	0	0	0	0	23	93	59	9	16	223	201
	serpin-6	0	0	0	0	11	59	16	0	0	0	0	0	0	18	69	18	0	0	85	105
	serpin-4	0	0	0	0	26	83	22	5	0	2	10	14	10	52	92	26	2	2	136	209
	PAP3	0	0	0	0	24	1	0	0	0	0	0	2	1	59	41	10	0	0	25	113
	hemolin	0	6	5	0	23	46	5	0	1	21	49	60	43	143	162	112	76	70	86	733
45–25	PI-like pr.	0	0	0	0	0	0	0	0	0	0	0	3	4	9	0	0	0	0	0	16
	HP8	0	14	14	2	16	10	44	1	3	0	12	16	10	20	8	30	0	6	105	101
	HP1b	0	0	0	0	0	8	0	0	0	0	0	0	0	0	11	1	0	0	8	11
	serpin-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	0	0	0	6
	IML-2	0	0	0	0	1	3	31	0	0	0	0	4	0	8	7	35	0	2	34	56
	PAP1	0	0	0	0	0	23	44	0	0	0	0	4	11	10	24	50	0	1	67	99
	serpin-1Z	0	0	0	0	1	14	8	0	1	0	0	0	0	0	15	5	0	1	24	21
	serpin-1K	0	0	0	0	4	15	10	0	1	0	0	0	0	0	10	10	0	0	31	20
	serpin-1A	0	0	0	0	4	13	12	0	2	0	0	0	0	3	13	12	0	2	31	30
	serpin-1E	3	22	34	24	78	149	104	54	61	8	29	32	22	7	148	93	45	44	527	492
	SPH2	0	1	11	1	61	104	46	5	1	0	0	3	0	42	92	66	1	2	229	206
	HP1a	0	4	6	0	23	75	22	1	0	0	4	3	0	20	71	19	0	0	131	117
	serpin-12	0	0	0	0	60	21	6	1	0	0	1	4	1	73	29	23	7	3	88	141
	lebocin-A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0	13
	scolexinA	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	16	0	0	0	16
	Kazal-type PI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	3	3	0	17
	IML-3	0	0	0	0	0	0	47	2	0	0	0	0	0	0	0	54	0	3	49	56
	scolexinB	0	0	0	0	0	0	34	2	2	0	0	0	0	0	0	29	1	1	38	31
	IML-7	0	0	0	0	0	0	46	2	2	0	0	0	0	0	0	40	0	0	51	40
	IML-10	0	0	0	0	1	7	70	4	9	0	0	0	0	5	0	65	0	8	91	78
	IML-4	0	0	0	0	0	0	6	0	0	0	4	7	0	5	2	22	0	2	6	42
25–20	attacin2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	60	46	0	106
	Spz7	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0	0	4	11	14	16
	lysozyme_like_1	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	1	9	8	10
	PGRP-1	0	0	0	0	0	0	1	0	56	0	3	0	0	0	0	11	3	55	57	71
	attacin10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	3	0	0	9
	attacin1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	4	0	0	9
	attacin7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	11	5	0	29
	attacin3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	21	5	0	30
	attacin9	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	14	107	63	1	184
	attacin8	0	0	0	0	0	0	1	2	1	0	0	0	0	0	0	41	49	44	5	133
	attacin4	0	2	1	0	0	0	6	11	4	0	5	2	0	5	0	95	117	95	24	318
	Spz2	0	0	0	0	0	0	14	8	0	0	0	0	0	0	0	9	9	0	21	17
	IML-9	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	1	0	0	11	1
20–10	Reeler1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5	31	0	37
	gloverin	0	0	0	0	0	0	0	0	11	0	9	3	0	1	0	1	0	41	11	56
	WAP15	0	0	0	0	0	0	0	2	3	0	0	0	0	0	0	0	1	3	5	4
	lysozyme_1	0	0	0	0	1	0	3	9	51	1	2	5	0	7	7	11	17	82	64	130

Proteins with unique spectral counts (USCs) in two or more gel slices (Table S5) are tabulated. So are the ones with USCs in one gel slice but their theoretical M_r values do not match the M_r range of that slice. Each cell, labeled with average USC of the three biological replicates, is colored with a gradient from blue to green and then to red based on its $\log_2(\text{USC}+1)$. Red rectangles enclose the slices in which theoretical sizes of the proteins correspond; USCs outside are considered abnormal. SPs/SPHs (including HPs) and protease inhibitors (including serpins) are colored blue and red, respectively.

Rather, we propose three mechanisms to explain this phenomenon: 1) control and induced hemocytes attach to and are stabilized by tissues to different extents, 2) induced hemocytes as a whole or in a subpopulation are more resilient to centrifugal force that causes cell rupture, and 3) unconventional secretory mechanisms (29) assist the release of some cytosolic proteins, which may differ between the control and induced states. Based on the number of identified proteins and their abundances, we consider cell lysis to be an important source of plasma proteins, such as proPOs.

Differential Protein Expression in Response to the Immune Challenge—We observed 155 changes in protein expression associated with immune challenge, and posit that most, if not all, of these changes represent the *Manduca* defense re-

sponse. Of the 77 up-regulated proteins (Table I), over half are known to be related to immunity. These proteins recognize microbial surface components, transduce or modulate extracellular signals (HPs, SPHs, serpins), or act as effectors (AMPs, PIs). We also detected two putative intracellular immune signal transducers (Smt3, Uev1A) (Table S2). Additionally, some of the up-regulated proteins (*e.g.* HSP25.4's, hypothetical proteins, lipid-binding proteins) may participate in defense in unknown ways. HSP25.4 isoforms a–c, f, and g were up-regulated 49.9, 33.3, 2.1, 4.1, and 4.7-fold, and their functions are worth exploring. HPs KGM_06199A and _06199B, each containing a von Willebrand factor type C domain stabilized by four disulfide bonds, may act as PIs or AMPs. Uncharacterized inducible proteins A and B, 59- and 62-residue peptides rich in certain residues (*e.g.* Gly), could also be novel AMPs. Others (*e.g.* apyrase-2, CXEs, esterases, lipases, neuroendocrine convertase-1, lysosomal carboxypeptidases) are probably responsible for changes in metabolism and protein processing after immune challenge.

Correlations of mRNA and Protein Abundances—Discrepancies between mRNA and protein levels have been well documented in mammalian systems (39, 40). Here, we demonstrated the same phenomenon in a systematic way in an insect for the first time. As shown in Fig. 3*A* and Fig. S1A, the correlation coefficients between plasma and fat body or hemocytes were $0.43_{CF\text{-}CP,\ gel}$, $0.41_{IF\text{-}IP,\ gel}$, $0.41_{CF\text{-}CP,\ ACN}$, $0.26_{IF\text{-}IP,\ ACN}$, $0.05_{CH\text{-}CP,\ gel}$, $-0.19_{IH\text{-}IP,\ gel}$, $0.00_{CH\text{-}CP,\ ACN}$, and $-0.25_{IH\text{-}IP,\ ACN}$. These results, supported by a well-established quantitative method with biological replicates, reinforce concerns regarding the overreliance of transcriptome or PCR data to interpret biological functions.

Despite the poor correlations in levels, we performed additional comparisons, such as mRNA and protein level changes (*i.e.* I/C), major *versus* minor contributors of plasma proteins (F–H), gel *versus* ACN data, immunity-related *versus* -unrelated proteins, and extra- *versus* intra-cellular proteins (Figs. 3*B* and 4 and Fig. S1B). The gel data (I/C) were more extensive than the ACN data, but the latter showed better parallel, likely due to strong induction of AMP genes ($0.69_{ACN} > 0.57_{gel}$ for fat body; $0.52_{ACN} > 0.44_{gel}$ for hemocytes). Extracellular protein level changes were better correlated than intracellular ones ($0.63 > 0.15$ for fat body and $0.46 > 0.21$ for hemocytes). Immunity-related protein level changes had the highest correlation coefficient (0.75_F) with the corresponding mRNA level changes in fat body, much higher than those in hemocytes (0.41_H) and immunity-unrelated (0.34_F and 0.30_H) (Fig. 4). Therefore, although direct correspondence of mRNA and protein levels was poor, choosing a set of process-related genes expressed in a major contributing tissue allowed a relatively good correlation uncovered based on the level changes.

Posttranslational Modification and Immune Complex Formation—While prefractionation of the protein mixtures by SDS-PAGE was performed primarily to enhance protein discoveries (41, 42), we also gained important insights into the changes in hemolymph proteome, one being posttranslational modifications and the other immune complex formation.

For 162 of the 654 proteins identified, we noted major discrepancies between their theoretical M_r values and their mobility of gel electrophoresis (Table II and Table S5). Most (130) of the inconsistencies remained after the effect of diffusion (Fig. S2) were considered. Glycosylation and proteolytic processing can be used to explain the M_r differences, which are neglected in most proteomic studies. Fortunately, proteolytic activation of the precursors of SPs (including HPs, PAPs), SPHs, POs, PSP, and spätzle-1 has been well documented in *M. sexta* (2, 16, 32, 43). It is also known that serpins and other protease inhibitors are involved in the control of immune SPs (19, 33–35, 44). While the current analysis lacks the resolution needed to reveal components of the serpin-protease complexes, it did provide consistent results and new leads (*e.g.* serpin-9 and -12) for future studies.

One important result of this study is the detection of high M_r immune complexes in the plasma proteome, which are stable in the presence of SDS and β-mercaptoethanol. By identifying 66 proteins (10–80 kDa) in gel slices 1–3 (80 to >200 kDa) (Table S5), we validated and extended the previous observations of such complexes (16, 19, 43, 45, 46). Nearly half (32) of them (*e.g.* PRRs, SPs, SPHs, proPOs, AMPs) may take part in immune responses. In contrast, only 171 or 24% of the 702 plasma proteins are related to immunity, indicating that defense proteins are more frequently captured in the high M_r complexes. Their coexistence with POs led us to propose that PO-catalyzed covalent crosslinking is responsible for their stability. Consistent with this working model, we have also seen similar components in >500 kDa complexes eluted from a gel filtration column (data not shown). It is possible that the complexes observed in the gel slices 1–3 were dissociated from higher M_r ones that are partially crosslinked since complete crosslinking may prevent them from entering the gel. Consequently, the nature and extent of these covalent modifications remain to be fully defined.

Acknowledgments—We thank Dr. Jack Dillwith for his critical comments on the manuscript.

* This work was supported by National Institutes of Health Grant GM58634 (to H. Jiang). Mass spectrometry analyses were performed in the DNA/Protein Resource Facility at Oklahoma State University, using resources supported by the NSF MRI and EPSCoR programs (DBI/0722494). This article was approved for publication by the Director of the Oklahoma Agricultural Experiment Station and supported in part under project OKL02450. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

S This article contains supplemental material.

‖ To whom correspondence should be addressed: Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, Tel: (405)-744-9400, E-mail: haobo.jiang@okstate.edu.

¶ Y.H., X.C., and S.Z. contributed equally to this work.

The mass spectrometry data (PXD003267 and 10.6019/PXD003267) have been deposited in the PRIDE repository (http://www.ebi.ac.uk/pride) using tools provided by the ProteomeXchange Consortium.

REFERENCES

1. Kanost, M. R. (2003) Hemolymph. *Encyclopedia of Insects,* V. H. Resh, and R. Carde, eds. Academic Press, New York, 505–508
2. Jiang, H., Vilcinskas, A., and Kanost, M. R. (2010) Immunity in lepidopteran insects. *Adv. Exp. Med. Biol.* **708,** 181–204
3. Lemaitre, B., and Hoffmann, J. (2007) The host defense of *Drosophila melanogaster. Annu. Rev. Immunol.* **25,** 697–743
4. Strand, M. R. (2008) The insect cellular immune response. *Insect Sci.* **15,** 1–14
5. Zhao, P., Lu, Z., Strand, M. R., and Jiang, H. (2011) Antiviral, anti-parasitic, and cytotoxic effects of 5,6-dihydroxyindole (DHI), a reactive compound generated by phenoloxidase during insect immune response. *Insect Biochem. Mol. Biol.* **41,** 645–652
6. Zhang, X., He, Y., Cao, X., Gunaratna, R. T., Chen, Y. R., Blissard, G., Kanost, M. R., and Jiang, H. (2015) Phylogenetic analysis and expression profiling of the pattern recognition receptors: Insights into molecular recognition of invading pathogens in *Manduca sexta. Insect Biochem. Mol. Biol.* **62,** 38–50
7. Rao, X. J., Cao, X., He, Y., Hu, Y., Zhang, X., Chen, Y. R., Blissard, G., Kanost, M. R., Yu, X. Q., and Jiang, H. (2015) Structural features, evolutionary relationships, and transcriptional regulation of C-type lectin-domain proteins in *Manduca sexta. Insect Biochem. Mol. Biol.* **62,** 75–85
8. Cao, X., He, Y., Hu, Y., Zhang, X., Wang, Y., Zou, Z., Chen, Y., Blissard, G. W., Kanost, M. R., and Jiang, H. (2015) Sequence conservation, phylogenetic relationships, and expression profiles of nondigestive serine proteases and serine protease homologs in *Manduca sexta. Insect Biochem. Mol. Biol.* **62,** 51–63
9. Cao, X., He, Y., Hu, Y., Wang, Y., Chen, Y. R., Bryant, B., Clem, R. J., Schwartz, L. M., Blissard, G., and Jiang, H. (2015) The immune signaling pathways of *Manduca sexta. Insect Biochem. Mol. Biol.* **62,** 64–74
10. He, Y., Cao, X., Li, K., Hu, Y., Chen, Y. R., Blissard, G., Kanost, M. R., and Jiang, H. (2015) A genome-wide analysis of antimicrobial effector genes and their transcription patterns in *Manduca sexta. Insect Biochem. Mol. Biol.* **62,** 23–37
11. Zhang, Y., Dong, Z., Wang, D., Wu, Y., Song, Q., Gu, P., Zhao, P., and Xia, Q. (2014) Proteomics of larval hemolymph in *Bombyx mori* reveals various nutrient-storage and immunity-related proteins. *Amino Acids* **46,** 1021–1031
12. Woltedji, D., Fang, Y., Han, B., Feng, M., Li, R., Lu, X., and Li, J. (2013) Proteome analysis of hemolymph changes during the larval to pupal development stages of honeybee workers (*Apis mellifera ligustica*). *J. Proteome Res.* **12,** 5189–5198
13. Karlsson, C., Korayem, A. M., Scherfer, C., Loseva, O., Dushay, M. S., and Theopold, U. (2004) Proteomic analysis of the *Drosophila* larval hemolymph clot. *J. Biol. Chem.* **279,** 52033–52041
14. Handke, B., Poernbacher, I., Goetze, S., Ahrens, C. H., Omasits, U., Marty, F., Simigdala, N., Meyer, I., Wollscheid, B., Brunner, E., Hafen, E., and Lehner, C. F. (2013) The hemolymph proteome of fed and starved *Drosophila* larvae. *PLoS ONE* **8,** e67208
15. Dwivedi, S. B., Muthusamy, B., Kumar, P., Kim, M. S., Nirujogi, R. S., Getnet, D., Ahiakonu, P., De, G., Nair, B., Gowda, H., Prasad, T. S., Kumar, N., Pandey, A., and Okulate, M. (2014) Brain proteomics of *Anopheles gambiae. OMICS* **18,** 421–437
16. Yu, X. Q., Jiang, H., Wang, Y., and Kanost, M. R. (2003) Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta. Insect Biochem. Mol. Biol.* **33,** 197–208
17. Nappi, A. J., and Christensen, B. M. (2005) Melanogenesis and associated cytotoxic reactions: Applications to insect innate immunity. *Insect Biochem. Mol. Biol.* **35,** 443–459
18. Wang, Y., and Jiang, H. (2004) Prophenoloxidase (proPO) activation in *Manduca sexta*: an analysis of molecular interactions among proPO, proPO-activating proteinase-3, and a cofactor. *Insect Biochem. Mol. Biol.* **34,** 731–742
19. Zou, Z., and Jiang, H. (2005) Manduca sexta serpin-6 regulates immune serine proteinases PAP-3 and HP8. cDNA cloning, protein expression, inhibition kinetics, and function elucidation. *J. Biol. Chem.* **280,** 14341–14348
20. Zou, Z., Najar, F., Wang, Y., Roe, B., and Jiang, H. (2008) Pyrosequence analysis of expressed sequence tags for *Manduca sexta* hemolymph proteins involved in immune responses. *Insect Biochem. Mol. Biol.* **38,** 677–682
21. Zhang, S., Gunaratna, R. T., Zhang, X., Najar, F., Wang, Y., Roe, B., and Jiang, H. (2011) Pyrosequencing-based expression profiling and identification of differentially regulated genes from *Manduca sexta*, a lepidopteran model insect. *Insect Biochem. Mol. Biol.* **41,** 733–746
22. Gunaratna, R. T., and Jiang, H. (2013) A comprehensive analysis of the *Manduca sexta* immunotranscriptome. *Dev. Comp. Immunol.* **39,** 388–398
23. Furusawa, T., Rakwal, R., Nam, H. W., Hirano, M., Shibato, J., Kim, Y. S., Ogawa, Y., Yoshida, Y., Kramer, K. J., Kouzuma, Y., Agrawal, G. K., and Yonekura, M. (2008) Systematic investigation of the hemolymph proteome of *Manduca sexta* at the fifth instar larvae stage using one- and two-dimensional proteomics platforms. *J. Proteome Res.* **7,** 938–959
24. Zhang, S., Cao, X., He, Y., Hartson, S., and Jiang, H. (2014) Semi-quantitative analysis of changes in the plasma peptidome of *Manduca sexta* larvae and their correlation with the transcriptome variations upon immune challenge. *Insect Biochem. Mol. Biol.* **47,** 46–54
25. Blagoev, B., Ong, S. E., Kratchmarova, I., and Mann, M. (2004) Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat. Biotechnol.* **22,** 1139–1145
26. Cao, X., and Jiang, H. (2015) Integrated modeling of protein-coding genes in the *Manduca sexta* genome using RNA-Seq data from the biochemical model insect. *Insect Biochem. Mol. Biol.* **62,** 2–10
27. Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75,** 4646–4658
28. Hubner, N. C., Ren, S., and Mann, M. (2008) Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics* **8,** 4862–4872
29. Nickel, W. (2010) Pathways of unconventional protein secretion. *Curr. Opin. Biotechnol.* **21,** 621–626
30. Matsumoto, Y., Oda, Y., Uryu, M., and Hayakawa, Y. (2003) Insect cytokine growth-blocking peptide triggers a termination system of cellular immunity by inducing its binding protein. *J. Biol. Chem.* **278,** 38579–38585
31. Kanost, M. R., Gorman, M. J. (2008) Phenoloxidases in insect immunity. In *Insect Immunology*, N. E. Beckage, ed. Elsevier, San Diego, 30
32. Jiang, H., and Kanost, M. R. (2000) The clip-domain family of serine proteinases in arthropods. *Insect Biochem. Mol. Biol.* **30,** 95–105
33. Ragan, E. J., An, C., Yang, C. T., and Kanost, M. R. (2010) Analysis of mutually exclusive alternatively spliced serpin-1 isoforms and identification of serpin-1 proteinase complexes in *Manduca sexta* hemolymph. *J. Biol. Chem.* **285,** 29642–29650
34. Tong, Y., Jiang, H., and Kanost, M. R. (2005) Identification of plasma proteases inhibited by *Manduca sexta* serpin-4 and serpin-5 and their association with components of the prophenol oxidase activation pathway. *J. Biol. Chem.* **280,** 14932–14942
35. Christen, J. M., Hiromasa, Y., An, C., and Kanost, M. R. (2012) Identification of plasma proteinase complexes with serpin-3 in *Manduca sexta. Insect Biochem. Mol. Biol.* **42,** 946–955
36. Miyaji, T., Murayama, S., Kouzuma, Y., Kimura, N., Kanost, M. R., Kramer, K. J., and Yonekura, M. (2010) Molecular cloning of a multidomain cysteine protease and protease inhibitor precursor gene from the tobacco hornworm (*Manduca sexta*) and functional expression of the cathepsin F-like cysteine protease domain. *Insect Biochem. Mol. Biol.* **40,** 835–846
37. Rayaprolu, S., Wang, Y., Kanost, M. R., Hartson, S., and Jiang, H. (2010) Functional analysis of four processing products from multiple precursors encoded by a lebocin-related gene from *Manduca sexta. Dev. Comp. Immunol.* **34,** 638–647
38. Nägler, D. K., and Ménard, R. (1998) Human cathepsin X: A novel cysteine protease of the papain family with a very short proregion and unique insertions. *FEBS Lett.* **434,** 135–139
39. Chen, G., Gharib, T. G., Huang, C. C., Taylor, J. M., Misek, D. E., Kardia, S. L., Giordano, T. J., Iannettoni, M. D., Orringer, M. B., Hanash, S. M., and Beer, D. G. (2002) Discordant protein and mRNA expression in lung

adenocarcinomas. *Mol. Cell. Proteomics* **1,** 304–313
40. Kendrick, N. (2014) A gene's mRNA level does not usually predict its protein level. Kendrick Laboratories, Inc., Madison, WI
41. Gautier, V., Mouton-Barbosa, E., Bouyssié, D., Delcourt, N., Beau, M., Girard, J. P., Cayrol, C., Burlet-Schiltz, O., Monsarrat, B., and Gonzalez de Peredo, A. (2012) Label-free quantification and shotgun analysis of complex proteomes by one-dimensional SDS-PAGE/NanoLC-MS: Evaluation for the large scale analysis of inflammatory human endothelial cells. *Mol. Cell. Proteomics* **11,** 527–539
42. Gao, B. B., Stuart, L., and Feener, E. P. (2008) Label-free quantitative analysis of one-dimensional PAGE LC/MS/MS proteome: Application on angiotensin II-stimulated smooth muscle cells secretome. *Mol. Cell. Proteomics* **7,** 2399–2409
43. Jiang, H., Wang, Y., Yu, X. Q., and Kanost, M. R. (2003) Prophenoloxidase-activating proteinase-2 from hemolymph of *Manduca sexta.* A bacteria-inducible serine proteinase containing two clip domains. *J. Biol. Chem.* **278,** 3552–3561
44. Suwanchaichinda, C., Ochieng, R., Zhuang, S., and Kanost, M. R. (2013) Manduca sexta serpin-7, a putative regulator of hemolymph prophenoloxidase activation. *Insect Biochem. Mol. Biol.* **43,** 555–561
45. Clark, K. D., and Strand, M. R. (2013) Hemolymph melanization in the silkmoth *Bombyx mori* involves formation of a high molecular mass complex that metabolizes tyrosine. *J. Biol. Chem.* **288,** 14476–14487
46. Clark, K. D. (2015) Altered tyrosine metabolism and melanization complex formation underlie the developmental regulation of melanization in *Manduca sexta. Insect Biochem. Mol. Biol.* **58,** 66–75