

# Highly Variable Immune-Response Proteins (185/333) from the Sea Urchin, *Strongylocentrotus purpuratus*: Proteomic Analysis Identifies Diversity within and between Individuals<sup>1</sup>

Nolwenn M. Dheilly,\* Sham V. Nair,<sup>2\*</sup> L. Courtney Smith,<sup>†</sup> and David A. Raftos\*

*185/333* genes and transcripts from the purple sea urchin, *Strongylocentrotus purpuratus*, predict high levels of amino acid diversity within the encoded proteins. Based on their expression patterns, 185/333 proteins appear to be involved in immune responses. In the present study, one- and two-dimensional Western blots show that 185/333 proteins exhibit high levels of molecular diversity within and between individual sea urchins. The molecular masses of 185/333-positive bands or spots range from 30 to 250 kDa with a broad array of isoelectric points. The observed molecular masses are higher than those predicted from mRNAs, suggesting that 185/333 proteins form strong associations with other molecules or with each other. Some sea urchins expressed >200 distinct 185/333 proteins, and each animal had a unique suite of the proteins that differed from all other individuals. When sea urchins were challenged in vivo with pathogen-associated molecular patterns (PAMPs; bacterial LPS and peptidoglycan), the expression of 185/333 proteins increased. More importantly, different suites of 185/333 proteins were expressed in response to different PAMPs. This suggests that the expression of 185/333 proteins can be tailored toward different PAMPs in a form of pathogen-specific immune response. *The Journal of Immunology*, 2009, 182: 2203–2212.

Recent studies of host defense have uncovered profound differences among animal phyla in the molecules used to mediate immune responses (1–3). It seems that the immune systems of different animals have evolved a variety of solutions to meet a basic requirement to combat pathogens. The resulting, highly diversified immune responses may reflect the specific physiology, lifespan, habitat, and associated microbial populations of particular animal groups, or they may have arisen by chance via evolutionary radiation. A number of metazoan phyla have now been studied, identifying a range of alternative immunological mechanisms that exhibit high levels of molecular diversity (2). These discoveries have driven a paradigm shift in our understanding of invertebrate immune responses, from systems that are simple and static, to those that are complex and have novel mechanisms for generating molecular hypervariability, a key requirement for keeping pace in the “arms race” against microbial pathogens.

Some invertebrate immune systems are proving to be surprisingly complex. Recent analyses of the purple sea urchin (*Strongylocentrotus purpuratus*) genome identified several large gene families, including gene models for 222 TLRs, 203 NOD-like receptors (NLR), 218 scavenger receptor cysteine-rich (SRCR) molecules, and 104 C-type lectins (4–6). The complexity and

large size of these gene families suggest that the receptors they encode may recognize individual pathogen-associated molecular patterns (PAMPs)<sup>3</sup> with a high degree of specificity. They might also act combinatorially, providing highly diverse cognitive capabilities (5). In addition to diversified receptors, the *S. purpuratus* genome contains homologs of RAG1 and RAG2, the molecules responsible for the somatic recombination of immunoglobulins in vertebrates (7). This suggests that the molecular tools required to generate molecular hypervariability might also exist among invertebrates.

*185/333* genes represent another high-diversity immune response system in sea urchins. This family was first identified during a transcriptome analysis of sea urchin immune response genes (8). Genes that were up-regulated in coelomocytes (immune cells) after the injection of LPS were identified by screening high-density arrayed, conventional cDNA libraries with probes generated by subtractive suppression hybridization. Surprisingly, ~60% of the expressed sequence tags (ESTs) characterized in this transcriptome analysis were members of a set of closely related transcripts with similarities to two uncharacterized sequences from *S. purpuratus*, called DD185 (GenBank accession AF228877 (9)) and EST333 (GenBank accession R62081 (10)), hence the designation *185/333*.

Screening an arrayed cDNA library constructed from immunologically activated coelomocytes indicated that the frequency of *185/333* mRNAs was enhanced more than 75-fold compared with a nonactivated arrayed cDNA library (8). Northern blots also showed striking increases in *185/333* expression in coelomocytes from bacterially activated sea urchins compared with injury controls (9). Based on this significant increase in gene expression, *185/333* transcripts were investigated in more detail, revealing an

\*Department of Biological Sciences, Macquarie University, North Ryde, New South Wales, Australia; and <sup>†</sup>Department of Biological Sciences, George Washington University, Washington, DC 20052

Received for publication September 18, 2007. Accepted for publication December 3, 2008.

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

<sup>1</sup> This study was funded in part by an Australian Research Council Discovery grant to D.A.R. (DP0880316) and by a United States National Science Foundation grant to L.C.S. (MCB 04-24235). N.D. is the recipient of an International Macquarie University Research Scholarship postgraduate scholarship.

<sup>2</sup> Address correspondence and reprint requests to Dr. Sham V. Nair Department of Biological Sciences, Macquarie University, North Ryde, NSW 2109, Australia. E-mail address: snair@bio.mq.edu.au

<sup>3</sup> Abbreviations used in this paper: PAMP, pathogen-associated molecular pattern; aCF, artificial coelomic fluid; CF, coelomic fluid; EST, expressed sequence tag; IPG, immobilized pH gradient; LC, liquid chromatography; MS, mass spectrometry; 1DE, one-dimensional electrophoresis; pI, isoelectric point; SNP, single nucleotide polymorphism; 2DE, two-dimensional electrophoresis.

unexpected level of sequence diversity (8, 11). To date, for 185/333 cDNAs, 689 have been characterized from 14 sea urchins. These sequences are predicted to encode 286 different proteins (11, 12).

The diversity evident among 185/333 transcripts is generated in several ways. 185/333 mRNAs are comprised of 25 different blocks or "elements" of nucleotide sequence that are present or absent in numerous combinations. This results in "element patterns" that are repeatedly identified in different mRNAs (8, 12), contributing significant diversity to the family of 185/333 messages. Single nucleotide polymorphisms (SNPs) and small insertions or deletions (indels) are also frequent in all 185/333 sequences, adding to diversity. Additionally, there are surprisingly high levels of frame shifts and the insertion of early stop codons in the mRNA sequences (11). The processes that generate this diversity among 185/333 mRNAs are not yet well defined, but they may include differences among the estimated fifty 185/333 gene loci (11, 13, 14), high levels of allelic polymorphism at each locus, RNA editing, and/or low-fidelity RNA polymerases (15), followed by posttranslational modifications. All of these processes may have been driven by positive selection associated with anti-pathogen defense (8, 11, 12, 14). Nucleotide sequence variability results in high levels of nonconservative amino acid substitutions among predicted proteins of the type consistent with intense evolutionary selection pressure (8, 12, 14, 15).

185/333 mRNAs are predicted to encode proteins with a hydrophobic leader, a glycine-rich region with multiple endoprotease cleavage sites, an RGD motif, a histidine-rich region, numerous *N*-linked and *O*-linked glycosylation sites, acidic patches, and several types of tandem and interspersed repeats (11). The deduced proteins do not contain cysteines, transmembrane regions, GPI linkage sites, identifiable domains, or any predictable folding patterns. The only regions with at least some similarity to other molecules are the RGD motif and one of the histidine-rich domains, which is comparable to histatins, a group of mammalian salivary proteins with powerful antifungal activities (16, 17). Brockton et al. (18) have shown that 185/333 proteins are expressed by two subsets of coelomocytes (small phagocytes and polygonal cells), and that the number of 185/333<sup>+</sup> cells increases in response to immunological challenge. Initial analyses of 185/333 proteins indicated that recombinant 185/333 proteins appear as multimers and that native proteins are present in a broad range of sizes in agreement with the mRNA sequences (18).

Here, we present the first comprehensive analysis of 185/333 proteins. It reveals a broad diversity of these proteins within individual sea urchins. Different sea urchins express different suites of 185/333 proteins, and expression is altered by immunological challenge. The data indicate that differential expression or posttranslational modification of 185/333 proteins might allow *S. purpuratus* to tailor immune responses toward specific pathogens.

## Materials and Methods

### Sea urchins

Adult *S. purpuratus* were purchased from Marinus Scientific after collection from the coast of southern California. They were maintained in the laboratory as described previously (19). *S. purpuratus* becomes immunorequiescent after long-term housing of >8 mo without significant disturbance (19, 20). Immunorequiescence can easily be reversed by injecting PAMPs, or in response to injury (19–21).

### Immunological challenge and sample collection

Animals were challenged by injecting 2  $\mu$ g of LPS or 4  $\mu$ g of peptidoglycan (Sigma-Aldrich) per milliliter of coelomic fluid (CF), as previously described (10, 19). Control animals were injected with an equivalent volume of artificial CF (aCF) (11). CF (100  $\mu$ l) was withdrawn from each sea

urchin immediately before injection, and then at various times after injection. A 23-gauge needle attached to a 1-ml syringe was inserted through the peristomium into the coelomic cavity and CF was withdrawn without anticoagulant. The CF was immediately expelled into a 1-ml tube and mixed with 100  $\mu$ l of urea sample buffer (2.4 M Tris-HCl (pH 6.8), 0.25% SDS, 4 M urea, 20% glycerol). Samples were stored at  $-70^{\circ}\text{C}$  until used. Proteins were precipitated using 2-D Clean-Up kits (GE Healthcare) according to the manufacturer's instructions and resuspended in urea sample buffer (8 M urea, 4% CHAPS, 60 mM DTT). The total protein content of each sample was determined with 2-D Quant kits (GE Healthcare).

### One-dimensional electrophoresis (1DE)

CF proteins in urea sample buffer ( $\sim 100$   $\mu$ g/well) were separated on 10% Tris-glycine precast polyacrylamide gels (Criterion gel system; Bio-Rad) at 130 V for 2 h, or on 7.5% bis-Tris polyacrylamide gels at 180 V for 1 h. After electrophoresis, proteins were visualized using Sypro Ruby (Sigma-Aldrich) following the manufacturer's protocol, or with Coomassie blue using standard protocols. Alternatively, the proteins were transferred to nitrocellulose membranes by Western blotting as described below.

### Two-dimensional electrophoresis (2DE)

Isoelectrofocusing was performed using an IPGphor IEF system (GE Healthcare). Immobilized pH gradient (IPG) gel strips (11 cm, pH 3–6 or pH 3–10; GE Healthcare) were rehydrated overnight with 200  $\mu$ g of CF proteins in 200  $\mu$ l of rehydration buffer (8 M urea, 2% CHAPS, 50 mM DTT, and 0.5% carrier ampholytes; Immobiline; GE Healthcare). Isoelectrofocusing was undertaken at 100 V for 3 h, 250 V for 20 min, and 8000 V for 6 h, to obtain a total of 26,000–30,000 Vh. The proteins in the IPG strips were reduced (1% DTT, 15 min) and alkylated (2.5% iodoacetamide, 15 min) before separation in the second dimension by SDS-PAGE using 10% Tris-glycine precast polyacrylamide gels (Criterion gel system; Bio-Rad) (22). After electrophoresis, protein spots on the gels were visualized by Coomassie blue staining or with Sypro Ruby. Alternatively, proteins were transferred to nitrocellulose membranes, as described see below.

### Antibodies

The polyclonal antisera against 185/333 proteins used in this study were the same as those reported by Brockton et al. (18). Antisera were generated against synthetic peptides corresponding to elements 1, 7, and 25a, which are present in most 185/333 cDNAs (see Ref. 11 and Fig. 5). The peptides were conjugated to keyhole limpet hemocyanin and injected into two rabbits per peptide on four separate occasions (Quality Controlled Biochemicals). Only those antisera for which the preimmunization bleeds did not cross-react with sea urchin CF proteins by Western blot (18) were used in this study. The three antisera used were designated anti-185-66, anti-185-68, and anti-185-71.

### Western blotting and immunostaining

Proteins separated by 1DE or 2DE were transferred from polyacrylamide gels to nitrocellulose membranes by electroblotting using a Criterion blotting system (Bio-Rad). Transfers were performed at 100 V for 1 h at room temperature using Towbin buffer (0.25 M Tris-HCl, 1.92 M glycine (pH 8.3)) with 20% methanol. Once transfer was complete, membranes were blocked by incubation in skim milk solution (7% skim milk powder in TBST; 10 mM Tris-HCl, 137 mM NaCl, 0.5% Tween 20 (pH 7.5)) overnight. Following several washes with TBST (three times for 5 min), membranes were incubated with anti-185/333 antisera (1/20,000 dilution of an equal mix of anti-185-66, -68, and -71 in TBST, or each antiserum separately) for 1 h at room temperature. The blots were washed with TBST (three times for 5 min) and incubated for 1 h at room temperature with goat anti-rabbit IgG conjugated with HRP (1/30,000 in TBST; Sigma-Aldrich). After washing in TBST (three times for 5 min), 185/333<sup>+</sup> proteins were visualized using ECL chemiluminescence (GE Healthcare) with blue light-sensitive high-performance chemiluminescence film (Hyperfilm ECL, GE Healthcare). Each blot was exposed to film for varying lengths of time (1–15 min) to optimize the exposure. In some cases, image processing (Photoshop; Adobe Systems) was used to combine autoradiographs from the various exposures into composite images. Proteome maps for 185/333<sup>+</sup> protein spots on 2DE Western blots were established using Progenesis software (PerkinElmer) to plot molecular mass and isoelectric point (pI). The relative intensities of different protein spots were calculated using Adobe Photoshop (Adobe Systems).

### Anti-185/333 ELISA

CF from LPS- and aCF-injected sea urchins was adjusted to  $5 \times 10^5$  coelomocytes/ml with calcium-magnesium-free seawater with EDTA and

imidazole (CMFSW-EI; 10 mM KCl, 7 mM Na<sub>2</sub>SO<sub>4</sub>, 2.4 mM NaHCO<sub>3</sub>, 460 mM NaCl, 70 mM EDTA, and 50 mM imidazole (pH 7.4)) and was mixed with an equal volume of CMFSW-EI containing 1% (v/v) Nonidet P-40 to lyse the cells, followed by centrifugation (12,000 × g, 10 s) to remove debris. The supernatant was diluted 1/20 with TBS and aliquoted in triplicate into 96-well ELISA plates (200 μl/well; Corning). Plates were incubated overnight at 4°C so that proteins could adhere to the wells. The plates were washed once with TBS and blocked for 1 h with TBS containing 4% (w/v) BSA. After blocking, 100 μl of anti-185 antisera (1/20,000 dilution of an equal mix of anti-185-66, -68, and -71 in TBST) was added per well for 2 h at room temperature with gentle shaking. The plates were washed three times for five minutes each with TBST before adding 100 μl of anti-rabbit IgG-alkaline phosphatase conjugate (1/20,000 in TBST; Sigma-Aldrich) per well. The plates were incubated with the secondary Ab for 2 h before being washed three times with TBST. After the final wash, 200 μl of *p*-nitrophenyl phosphate (2 mg/ml in 0.1 M glycine, 1 mM MgCl<sub>2</sub>, 1 mM ZnCl<sub>2</sub> (pH 10.4); Sigma-Aldrich) was added per well and incubated for 30 min before absorbance was read in a microplate spectrophotometer at 415 nm. Data were corrected for the absorbance in wells prepared without sea urchin proteins. Controls included wells in which the primary (anti-185) Abs or both the primary and secondary Abs were omitted, and wells in which an irrelevant Ab (rabbit anti-tunicate collectin peptide (24, 25)) was used in place of the anti-185 antisera.

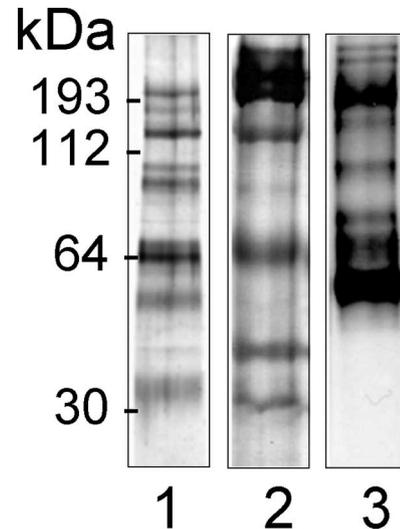
### Mass spectrometry and data analysis

Mass spectrometry (MS) was performed on CF proteins separated by either 1DE or 2DE. To extract proteins from 1DE, Coomassie blue-stained gels were washed twice in water (10 min each). Individual lanes were cut into 16 slices of equal sizes so that proteins in different molecular mass ranges could be analyzed separately by MS (see Fig. 8). To extract proteins after 2DE, spots from 2DE gels that corresponded to individual 185/333<sup>+</sup> proteins on Western blots were excised.

Gel slices or excised protein spots were washed briefly with 100 mM NH<sub>4</sub>HCO<sub>3</sub> before being destained (3 × 10 min) with 25 mM NH<sub>4</sub>HCO<sub>3</sub>/acetonitrile (1/1) and dehydrated in 100% acetonitrile for 5 min. After dehydration, gel pieces were air-dried, reduced with 10 mM DTT in 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 45–60 min at 56°C, and alkylated with 55 mM iodoacetamide in 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 45 min at room temperature. The gel pieces were washed once with 100 mM NH<sub>4</sub>HCO<sub>3</sub> for 5 min and twice with 25 mM NH<sub>4</sub>HCO<sub>3</sub>/acetonitrile (1/1) for 5 min before being dehydrated with 100% acetonitrile. The dehydrated gel slices were air-dried and rehydrated with trypsin (12.5 ng/μl in 50 mM NH<sub>4</sub>HCO<sub>3</sub>; Promega) for 30 min at 4°C. An additional aliquot of 50 mM NH<sub>4</sub>HCO<sub>3</sub> was added before the proteins were digested at 37°C overnight. The resulting tryptic peptides were extracted from the gel pieces by washing twice with 2% formic acid in 50% acetonitrile. Extracts were combined and concentrated to 10 μl by vacuum centrifugation.

Mass spectrometry was performed at the Australian Proteome Analysis Facility (Macquarie University). The tryptic digest extracts from 1DE gel slices were subjected to data-dependent nanocapillary reversed phase liquid chromatography followed by electrospray ionization using a Thermo LCQ Deca ion trap mass spectrometer (Thermo Scientific); liquid chromatography (LC)-MS/MS). For LC-MS/MS, a microbore HPLC system (TSP4000; Thermo Scientific) was modified to operate at capillary flow rates using a simple T-piece flow-splitter. Columns (8 cm × 100 μm inside diameter) were packed with 100 Å, 5-μm Zorbax C18 resin at 500 ψ. Integrated electrospray tips for the columns were made from fused silica, pulled to a 5-μm tip using a laser puller (Sutter Instrument). An electrospray voltage of 1.8 kV was applied using a gold electrode via a liquid junction upstream of the column. Samples were introduced onto the analytical column using a Surveyor autosampler (Thermo Scientific). The HPLC column eluent was eluted directly into the electrospray ionization source of the ion trap mass spectrometer. Peptides were eluted with a linear gradient of buffer A (0.1% formic acid) and buffer B (acetonitrile containing 0.1% formic acid) at a flow rate of 500 nl/min. Automated peak recognition, dynamic exclusion, and daughter ion scanning of the top three most intense ions were performed using the Xcalibur software as previously described (23).

GPM open source software (Global Proteome Machine Organization; www.thegpm.org) was used to search peptide sequences against a combined *S. purpuratus* database created with sequences downloaded from the National Center for Biotechnology Information (NCBI). This FASTA format database contained 44,037 protein sequences comprising all *S. purpuratus* sequences held by NCBI as of April 2008. This database also incorporated a list of common human and trypsin peptide contaminants. Search parameters included MS and MS/MS tolerances of ±2 Da and ±0.2 Da, tolerance of up to three missed tryptic cleavages, and K/R-P cleavages.



**FIGURE 1.** 1DE SDS-PAGE and Western blot of CF proteins from an individual sea urchin. CF proteins were loaded at 100 μg/well onto 10% SDS-PAGE gels. Lane 1, Sypro Ruby staining. Lane 2, Coomassie blue staining. Lane 3, Western blot immunostained with an equal mixture of anti-185 sera (anti-185-68, -66, and -71; 1/20,000). The positions of molecular mass markers are shown on the left.

Fixed modifications were set for carbamidomethylation of cysteine, and variable modifications were set for oxidation of methionine. Peptides isolated from 1DE gel slices were deemed to significantly match corresponding peptides in the available databases if their GPM log<sub>e</sub> score was lower than the statistically significant cutoff values assigned by the GPM algorithm.

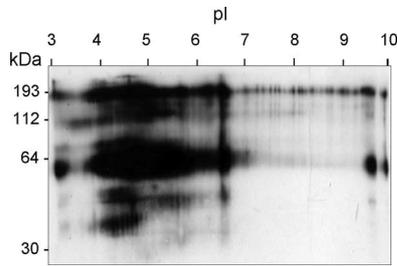
Protein spots from 2DE gels were analyzed using MALDI-TOF-TOF using an LTQ FT Ultra hybrid mass spectrometer (Thermo Scientific). MS ion searches were compared with 185/333 sequences in a custom database containing 81 translated full-length 185/333 cDNA sequences (12) and 689 EST sequences (8) using the Mascot search engine (Matrix Sciences; www.matrixscience.com/) set for carbomethylation (C) and oxidation (M) variable modifications, with peptide and fragment mass tolerances of ±50 ppm and ±0.5 Da respectively, and a maximum missed cleavage of one.

## Results

### One-dimensional SDS-PAGE analysis of CF proteins

1DE was used to provide a preliminary assessment of the diversity of proteins in CF, particularly those that were detected by anti-185 Abs (Fig. 1). The molecular masses of CF proteins from all of the sea urchins analyzed in the present study ( $n = 13$ ) ranged from 20 kDa to >193 kDa (see Fig. 6). Sypro Ruby-stained CF proteins from the individual sea urchin ranged from 30 kDa to >193 kDa (Fig. 1, lane 1). There were substantial differences in banding patterns when electrophoretically separated CF proteins from the same sea urchin were stained with Coomassie blue compared with Sypro Ruby (Fig. 1, compare lanes 1 and 2). This was probably due to different sensitivities and physicochemical properties of the stains (26).

Immunostained 1DE Western blots of CF proteins identified numerous 185/333<sup>+</sup> bands ranging from ~20 kDa to >193 kDa (Fig. 1, lane 3). Controls that omitted the primary (anti-185) and/or secondary (anti-IgG) Abs, and irrelevant controls using anti-tunicate collectin Abs instead of anti-185 antisera, were negative (data not shown). Most 185/333<sup>+</sup> bands from all of the sea urchin samples analyzed ( $n = 13$ ) had molecular masses ranging from 50 kDa to >193 kDa, although some were as low as 20 kDa (see Fig. 6). Many of the 185/333<sup>+</sup> bands on Western blots were not visible on Sypro Ruby- and Coomassie blue-stained 1DE gels, suggesting that the 185/333 proteins were present at extremely low concentrations.



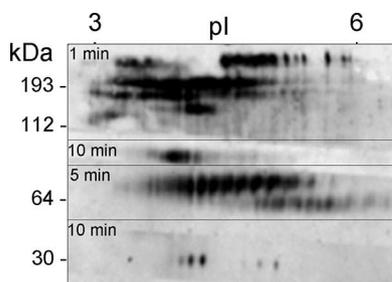
**FIGURE 2.** 2DE Western blot of 185/333<sup>+</sup> proteins. The gel was loaded with 200  $\mu$ g of CF proteins from a single sea urchin and the blot was immunostained with an equal mixture of anti-185 sera (anti-185-66, -68, and -71; 1/20,000). pIs are shown as pH units on the *top* of the blot, and molecular masses (kDa) are indicated on the *left*.

#### Two-dimensional Western blots of CF proteins

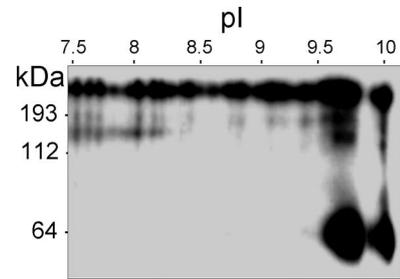
Given that cDNA analyses predicted a broad array of 185/333 proteins in the CF (8, 11, 12), we extended our analysis of 185/333 proteins using 2DE, which affords far greater resolution than does 1DE. Sea urchin CF proteins were initially separated by isoelectric focusing (pH 3–10) and then by 10% SDS-PAGE. Large numbers of 185/333<sup>+</sup> spots were evident after 2DE, to the extent that they often appeared as dense smears (Fig. 2). Most of the 185/333 proteins recognized by the antisera in all of the sea urchins analyzed ( $n = 13$ ) had pIs between 3 and 7 with apparent molecular masses of 40 kDa to >193 kDa.

To improve the resolution of individual 185/333<sup>+</sup> proteins on 2DE Western blots, additional isoelectric focusing separations were conducted on immobilized pH gradient (IPG) strips with a pH range of 3–6. Composite images obtained from film exposures of three different time intervals were used to visualize 185/333<sup>+</sup> proteins of varying abundance (Fig. 3). For the individual sea urchin shown in Fig. 3, image analysis of the different exposures revealed that fifty-one 185/333<sup>+</sup> spots were evident after a 1-min film exposure. An additional 117 spots were evident after 5 min, and a further 96 spots appeared after 10 min, making a total of 264 spots that were detected on the composite image for this individual. The number of discrete 185/333<sup>+</sup> spots varied substantially among individuals. Three of the 13 sea urchins tested did not express detectable levels of 185/333 proteins. However, the number of discrete 185/333<sup>+</sup> spots was often >200 in other individuals.

The enhanced protein separation capabilities of 2DE also showed that each discrete 185/333<sup>+</sup> band evident in 1DE contained numerous variants with similar molecular masses but different pIs. For example, at least fifteen 185/333<sup>+</sup> spots with dif-



**FIGURE 3.** Composite image of a 2DE Western blot. The gel was loaded with 200  $\mu$ g of CF proteins from sea urchin 12. The blot was immunostained with an equal mixture of the three anti-185 antisera (anti-185-66, -68, and -71; 1/20,000) and exposed to autoradiographic film for 1, 5, or 10 min. The different exposures were merged to give a final composite image. pIs are shown on the *top* as pH units, and molecular masses (kDa) are shown on the *left*.



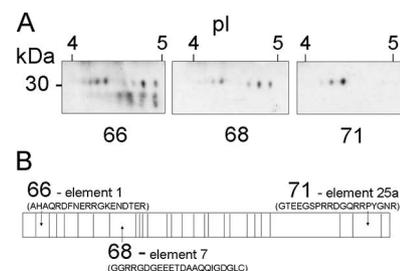
**FIGURE 4.** Enlarged region of a 2DE Western blot of CF proteins from animal 12 immunostained with an equal mixture of the three different anti-185 sera (anti-185-66, -68, and -71; 1/20,000). pIs are shown on the *top* as pH units, and molecular masses (kDa) are shown on the *left*.

ferent pIs were evident at  $\sim$ 75 kDa (Fig. 3). A further eight 185/333<sup>+</sup> spots appeared at  $\sim$ 60 kDa and at least six different 185/333<sup>+</sup> spots were present at  $\sim$ 30 kDa (Fig. 3). These pI variants were often regularly spaced from each other, differing by  $\sim$ 0.1–0.2 pH units. There were also numerous 185/333<sup>+</sup> spots with identical pI but different molecular masses (Fig. 4). For instance, three proteins in the pI 7.5–7.75 range each had three different molecular mass forms at  $\sim$ 150 kDa, 193 kDa, and >193 kDa (Fig. 4).

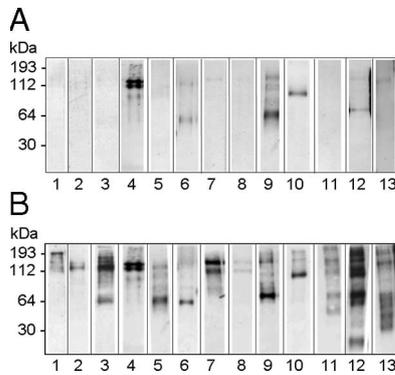
The three different anti-185 sera, which were raised against different regions of the most commonly predicted 185/333 polypeptide sequences, identified subsets of 185/333 proteins (Fig. 5). Within the small region of a 2DE Western blot (pI range of 4–5 and molecular masses of 30 kDa or lower), anti-185-66 resolved 14 distinct variants, of which 7 were 30 kDa and had a pI range of 4–5 (Fig. 5A). An additional seven anti-185-66-positive spots were smaller than 30 kDa, with a pI range of 4.5–5. In comparison, anti-185-68 recognized five of the seven proteins at 30 kDa with a pI range of 4–5 that were present on the anti-185-66 blot, but did not recognize the set of seven lower molecular mass spots. Anti-185-71 recognized only two of the 30 kDa proteins that were recognized by the other two antisera. This result is consistent with the expression of truncated 185/333 cDNAs (11).

#### Diversity of 185/333 proteins between individuals

There were major differences in the expression profiles of 185/333 proteins among different sea urchins. Some animals did not show 185/333 expression before LPS injection (Fig. 6A, *lanes 3 and 11*), although expression was evident after challenge (Fig. 6B, *lanes 3 and 11*). In other cases, 185/333 expression, which was very low



**FIGURE 5.** Different anti-185 sera recognize subsets of 185/333 proteins. *A*, Enlarged regions of three different 2DE Western blots loaded with 200  $\mu$ g of CF proteins from the same sea urchin and immunostained separately with anti-185-66, anti-185-68, or anti-185-71 (1/20,000). *B*, Schematic diagram of a 185/333 protein showing the peptides used to generate the three different anti-185 sera. The rectangle represents the protein sequence, which is composed of the leader and 25 nucleotide sequence “elements”. The positions and sequences of the synthetic peptides against which the three anti-185 sera were raised are indicated (15).



**FIGURE 6.** IDE Western blots of CF from 13 different sea urchins sampled before (A) and 96 h after (B) challenge with LPS. The equivalent lane numbers in both panels refer to the same animals from which CF samples were obtained before and after LPS challenge. The blots were immunostained with an equal mixture of the three different anti-185 sera (anti-185-66, -68, and -71; 1/20,000).

in preinjection samples, increased significantly after LPS injection. A broad distribution of 185/333 molecular masses ranging from 20 kDa to >193 kDa was apparent among the 13 animals analyzed by IDE Western blots (Fig. 6). Although many individuals expressed some 185/333 proteins with identical molecular masses (e.g., 112 kDa, Fig. 6, lanes 1–4; 100 and 120 kDa, Fig. 6, lanes 7 and 8), the suite of 185/333<sup>+</sup> bands expressed by each individual sea urchin was unique. Three animals did not express detectable 185/333 proteins, even after challenge with LPS (data not shown). In total, nineteen 185/333<sup>+</sup> bands with distinct molecular masses were identified among the 13 sea urchins analyzed by IDE. The molecular masses of these proteins seemed to be evenly spaced at 8–10 kDa apart over the molecular mass range of 20 to >193 kDa (Fig. 6).

*MS analysis of 185/333<sup>+</sup> proteins*

Proteins on Coomassie blue-stained 2DE gels that corresponded to 185/333<sup>+</sup> spots on Western blots were analyzed by MS. This analysis failed to identify 185/333 proteins unambiguously using the default criteria on the Mascot search engine. Although MS of each

Table I. Mass spectrometric (LC-MS/MS) data for peptides isolated from IDE gels of *S. purpuratus* CF that match known 185/333 sequences<sup>a</sup>

Animal	Gel Slice	Log <sub>e</sub>	m + h	z	Sequence	GenBank Accession No.		
6	6	-1.30	2594.5	3	MAVLTLATMAATTSII IATTQKVTK <sup>b</sup>	gb ABK88425.1		
17	2	-3.20	1735.8	3	GQGGFVGGRPGGMQMGGPR	gb ABR22418.1		
		-3.20	1735.8	3	GQGGFVGGRPGGMQTGSPR	gb ABZ10666.1		
		-3.30	1493.6	2	FDGPESGAPQMEGR	gb ABZ10664.1		
		-9.80	2886.3	3	RGDGEETDAAQQIGDGLGGPGQFDGPGR	gb ABZ10668.1		
		-2.40	1735.8	3	GQGGFVGGRPGGMQMGLR	gb ABK88417.1		
		-3.30	2131.9	3	RGDGEETDAAQQIGDGLGGR	gb ABZ10666.1		
		-1.70	1506.7	2	FDGPFVGGAPQMGGR	gb ABR22467.1		
		-4.60	1157.6	2	KPFGDHPFGR	gb ABR22410.1		
		-2.00	2819.2	3	GDGEEETDAAQQIGDGLGGSGQFDGPRR	gb ABR22477.1		
		3	3	-2.40	1735.8	3	GQGGFVGGRPGGMQMGGPR	gb ABR22418.1
				-2.40	1735.8	3	GQGGFVGGRPGGMQTGSPR	gb ABZ10666.1
				-4.20	1649.7	3	RFDGPESGAPQMEGR	gb ABZ10664.1
				-5.20	1493.6	2	FDGPESGAPQMEGR	gb ABZ10664.1
				-3.00	2131.9	3	RGDGEETDAAQQIGDGLGGR	gb ABZ10666.1
				-1.50	2320.2	2	PQTDQRNRLVSAATKAAMRM <sup>b</sup>	gb ABK88425.1
				-5.80	2887.3	3	RGDGEETDAAQQIGDGLGGPGQFDGPGR	gb ABZ10668.1
				-1.40	1735.8	3	GQGGFVGGRPGGMQMGLR	gb ABK88417.1
				-2.60	1979.9	3	MGGRNSTNPEFGGSRPDGAG <sup>b</sup>	gb ABR22330.1
				-1.70	2744.3	3	RNSTNPEFGGSRPDGAGGRPLFGQGR <sup>b</sup>	gb ABR22330.1
		4	4	-4.60	2927.3	3	RGDGEETDAAQQIGDGLGGPGQFDGHGR	gb ABZ10669.1
-2.40	1479.6			2	FDGPESGAPQMDGR	gb ABZ10666.1		
-4.00	1980.0			3	FGGSRPDGAGGRPFFGQGR	gb ABZ10669.1		
-2.40	2886.3			3	RGDGEETDAAQQIGDGLGGPGQFDGPGR	gb ABZ10668.1		
-2.00	1157.6			2	KPFGDHPFGR	gb ABR22410.1		
-5.10	1633.7			2	RFDGPESGAPQMEGR	gb ABZ10664.1		
-0.17	2564.1			3	FDGPESGAPQMEGRRQNGVPMGGR	gb ABK88329.1		
-3.10	2132.0			3	RGDGKEETDAAQQIGDGLGGR	gb ABR22436.1		
-1.40	1926.3			3	RGDGEETDAAQQIGDGLGGPGQFDGHGR	gb ABZ10669.1		
-1.7	3398.5			3	DFNERREKENDTERGQGGFVGGRPGGMQMGGP	gb ABK88476.1		
5	5	-3.5	1980.0	3	FGGSRPDGAGGRPFFGQGR	gb ABZ10669.1		
		-1.5	1831.9	3	RFDGPEPAPQMEGR	gb ABK88803.1		
		-2.7	2131.0	3	RGDGEETDAAQQIGDGLGGR	gb ABZ10666.1		
		-1.8	2927.3	3	RGDGEETDAAQQIGDGLGGPGQFDGHGR	gb ABZ10669.1		
		7	7	-1.6	1515.7	2	ADVVEIIVNEEDVN <sup>b</sup>	gb ABA19607.1
				-3.3	1649.7	3	RFDGPESGAPQMEGR	gb ABZ10664.1
		22	2	-4	1493.6	2	FDGPESGAPQMEGR	gb ABZ10664.1
				-3.5	1980.0	3	FGGSRPDGAGGRPFFGQGR	gb ABZ10664.1
				-2.50	1033.5	2	FGAPQMGGR	gb ABR22467.1
				-1.40	2814.3	3	RGRGQGRFVGGRPGGMQMGGPRQDGGPMG	gb ABK88373.1
-3.70	1649.7	3	RFDGPESGAPQMEGR	gb ABZ10664.1				

<sup>a</sup> The data are from IDE gel slices of CF from three different sea urchins. The GenBank accession numbers of the 185/333 sequences that matched peptides from IDE gels are also shown. The amino acid sequences shown are the matching 185/333 peptides from the NCBI database. Parameter definitions are: log<sub>e</sub>, values indicate the probability that a putative peptide sequence corresponding to a mass spectrum arises stochastically. The lower the log<sub>e</sub> value, the more significant the assignment of the peptide sequence to the mass spectrum. The log<sub>e</sub> values listed in this table are all lower than the significance cutoff values assigned by the GPM algorithm, and so they are deemed to be statistically significant (43); m + h, peptide mass in Da + 1; z, peptide charge.

<sup>b</sup> Peptides that matched sequences for previous analysis of nucleotide sequences that might represent frameshift mutations (12, 13).

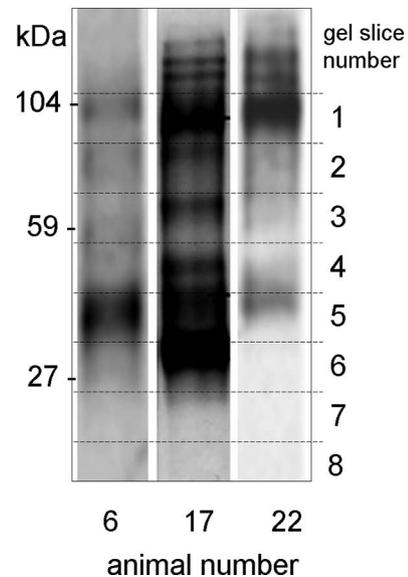
protein successfully identified ion fragments with mass/charge properties that were similar to known 185/333 sequences, these matches did not yield sufficiently high statistical probabilities to confirm the identity of any of the proteins (data not shown).

In contrast, MS analysis of CF proteins separated by IDE unequivocally identified 185/333 proteins. A total of 41 peptides isolated from IDE gel slices matched with known 185/333 sequences from the NCBI database (Table I). The same 185/333 peptide was often identified in more than one gel slice. For example, the peptide FDGPESGAPQMEGR appeared in gel slices 2–4 from animal 17. It is unlikely that the occurrence of the same peptide in multiple fractions was due to contamination of the different fractions with exactly the same 185/333 protein. If the same protein was present in more than one fraction, we would expect to find exactly the same combination of peptides from that protein in more than one fraction. However, this was not the case. Only single peptides were found in each fraction. Given this repetitive identification of the same peptide in different gel slices, a total of 23 unique 185/333 peptides were identified among all of the gel slices. In some cases, the GPM algorithm allocated slightly different predicted amino acid sequences to the same MS spectrum. For instance, GPM identified two peptides (GQGGFGGRPGGM QMGGPR and GQGGFGGRPGGMQTGSPR; underlining identifies amino acids that differ) from the same MS spectrum. This occurred because the amino acid differences between these two peptides were predicted to yield peptides with indistinguishable mass/charge ratios. Matches to both of these peptides were present in the NCBI database. In some cases, the subtle differences between peptides involved the inclusion of an arginine resulting in closely related peptides of different lengths, such as RGDGEEETDAAQIQIGDGLGGR and RGDGEEETDAAQIQIGDGLGGPGQFDGHR. It is noteworthy that of the 41 peptides identified, 36 were located in the glycine-rich (N-terminal) and central regions of the predicted proteins. This is consistent with frameshifts and early stop codons that have been identified in about half of the cDNAs (11). In many cases, the introduction of these frameshifts or early stop codons would have resulted in the loss of epitopes for the antiserum anti-185-68, which was raised against the C-terminal region of full-length 185/333 proteins, and to a lesser extent, the epitope for anti-185-71, which is in the central region of predicted full-length 185/333 proteins. As a result, these antisera would not have identified many of the 185/333 proteins encoded by truncated mRNAs in the Western blot shown in Fig. 5A. Some of the peptides shown in Table I matched sequences that previous analyses have suggested are derived from frameshift mutations (12, 13). All of the peptides that matched 185/333 sequences came from gel slices 2–7, which were also shown by Western blot to contain the vast majority of 185/333 proteins (Figs. 6 and 7). Of the three sea urchins analyzed by IDE and LC-MS/MS, animal 17 had the highest overall expression levels for 185/333 proteins (Fig. 7) and yielded 37 of the 41 peptides that matched 185/333 sequences.

#### 185/333 protein expression increases after LPS challenge

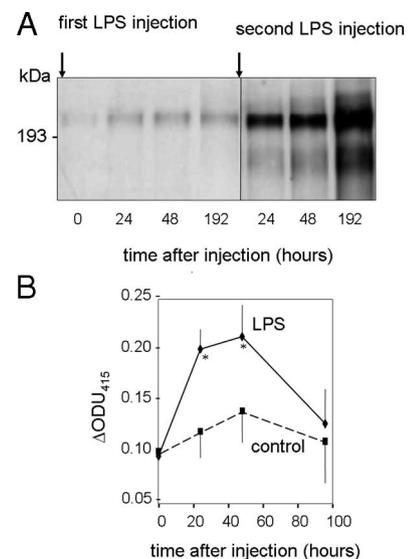
IDE Western blots of CF from five of the six animals injected with LPS showed increases in 185/333<sup>+</sup> protein expression between 24 and 192 h (representative results are shown in Fig. 8A). Repeated injections with LPS performed 14 days after the primary challenge increased 185/333 protein expression to levels that were higher than those evident after the first injection (Fig. 8A). Sea urchins injected with aCF either did not respond to the injection or showed weak responses in terms of 185/333 protein expression (data not shown).

These Western blot data were confirmed by ELISA (Fig. 8B). Injecting LPS into sea urchins significantly ( $p < 0.05$ ) increased

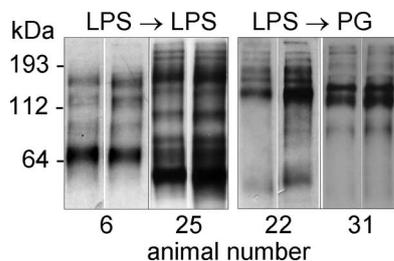


**FIGURE 7.** IDE Western blots of CF proteins from three different sea urchins (animals 6, 17, and 22) immunostained with an equal mixture of anti-185 sera (anti-185-66, -68, and -71; 1/20,000). The numbers on the *left* show the approximate positions of slices cut from IDE gels for analysis by mass spectrometry (see Table I). In total, gels were cut into 16 slices. The positions of only eight of those slices containing the highest molecular mass proteins are shown here.

the titer of 185/333 proteins detected in CF by ELISA. By 48 h after LPS injection, anti-185 reactivity had increased 2.4-fold compared with the 0 h time point ( $p < 0.05$ ). It returned to levels that were indistinguishable ( $p > 0.05$ ) from those before LPS challenge by 96 h. In response to aCF (control) injections, the titer of 185/333 proteins in CF also increased up to 48 h postinjection, but only to a level that was 1.4-fold greater than the preinjection time point. This increase was significantly less than the response to LPS



**FIGURE 8.** The titer of 185/333 proteins in CF increases after immune challenge. **A**, IDE Western blots of CF from a single sea urchin collected at various times after the animal was injected twice with LPS. The second LPS injection was administered 360 h after the first. **B**, 185/333 protein expression levels determined by ELISA ( $\Delta\text{ODU}_{415}$ ) at various times after naive sea urchins ( $n = 5$ ) were injected with LPS or aCF (controls). Bars indicate SEM. Asterisks denote time points that differed significantly ( $p < 0.05$ ) between LPS-injected sea urchins and controls.



**FIGURE 9.** 1DE Western blots of CF collected after sea urchins had been injected with LPS or PG. Animals 6 and 25 were injected with LPS twice with a 360-h interval, while animals 22 and 31 were injected first with LPS and then with PG 360 h later. CF was collected 96 h after each injection, and the Western blots were immunostained with an equal mixture of the three different anti-185 sera (anti-185-66, -68, and -71; 1/20,000).

( $p < 0.05$ ). Negative controls confirmed the specificity of the ELISA. Wells in which the primary (anti-185 sera) or secondary (anti-rabbit IgG) Abs were omitted, or when the primary Ab was replaced by an irrelevant control (anti-tunicate collectin), did not yield absorbance readings that were significantly greater than background levels ( $p > 0.05$ , data not shown).

#### Diversity of 185/333 protein expression after immunological challenge

1DE analysis did not detect changes in the types of 185/333 proteins expressed by individual sea urchins in response to LPS or PG injections, even though densitometry (data not shown) indicated there was an increase in the relative quantity of 185/333 proteins after repeated challenge (Fig. 9). The molecular masses of the predominant 185/333 proteins in CF from animals 6 and 25 did not change in response to an initial injection of LPS, or after a second LPS injection 2 wk later. A similar result was apparent for animals

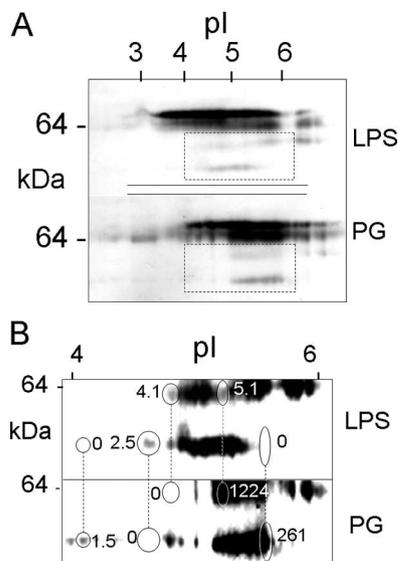
22 and 31, which were challenged with LPS followed by PG 2 wk later. The 185/333<sup>+</sup> bands of these two animals did not show any major changes in molecular mass when the challenge was switched from LPS to PG.

In contrast, 2DE revealed clearly discernable changes in the pIs of 185/333 proteins when sea urchins were challenged with PG after initially responding to LPS, even though the molecular masses of the predominant forms remained the same (Fig. 10). This is best exemplified in the small region of the 2DE Western blots of CF from animal 31 (pIs of 4–6, molecular masses of 40–65 kDa) shown in Fig. 10B. CF from this sea urchin was collected after an initial challenge with LPS (Fig. 10B, upper panel), and then again after a subsequent injection of PG (Fig. 10B, lower panel). Of the thirty 185/333<sup>+</sup> spots in this region, seven were found only after LPS challenge (e.g., 63 kDa, pI 4.8), and six were found only after challenge with PG (e.g., 42 kDa, pI 5.5). The remaining 17 proteins were present after both the LPS and PG challenges. The relative expression levels of some of these shared proteins also differed between challenges. For instance, one of the proteins (63 kDa, pI 5.2) had a densitometry value of 5.1 after LPS injection and 1224 after PG challenge, representing a 240-fold increase in expression intensity in response to PG.

#### Discussion

This study has identified substantial diversity among 185/333 proteins, as reflected by their broad range of molecular masses and pIs. Most importantly, there are obvious differences in the suites of 185/333 proteins expressed by different sea urchins, and these suites of proteins undergo subtle but extensive changes in response to different types of immune challenge. The diversity of 185/333 proteins and the changes in their expression detected by the current proteomic analysis correspond with previous studies of mRNAs from sea urchins responding to different PAMPs (11). All of these data suggest that sea urchins may be capable of altering the expression of 185/333 proteins to tailor specific responses against different PAMPs.

Despite the overall agreement between our current observations of 185/333 proteins and predictions based on the prior analyses of 185/333 nucleotide sequences (11, 12), there are some discrepancies. The predicted sizes of 185/333 proteins based on cDNA sequences range from 4 to 55.3 kDa (11, 12), with predicted pIs ranging from 5.42 to 11.54. However, the present study shows that native 185/333 proteins have a far wider range of molecular masses (20 to >193 kDa using 1DE analysis), with predominant bands often being at the high end of this range. They also have far more acidic pIs than expected, mainly between 3 and 7. It is unlikely that the discrepancy between predicted and observed molecular masses is due to disulfide-bonded oligomerization because none of the predicted 185/333 amino acid sequences identified to date contains cysteines (8, 11). Even if cysteines were present in missense sequences of frame-shifted proteins, the strong reducing conditions used to prepare samples would have disrupted any oligomers held together by disulfide bonds. Another explanation for the discrepancy in molecular masses is that 185/333 proteins are glycosylated and form large complexes covalently linked to carbohydrates. There are numerous conserved sites for *N*-linked glycosylation within the histidine-rich region of the 185/333 proteins (elements 11–25), and there are conserved sites for *O*-linked glycosylation in the carboxyl-terminal region (8, 11). However, deglycosylation of *N*-linked oligosaccharides failed to decrease the molecular masses of 185/333 proteins to the size of predicted monomers (L. C. Smith, unpublished data). Given these results, it seems likely that the disparity between predicted and observed molecular masses reflects oligomerization based on mechanisms



**FIGURE 10.** 2DE Western blots of CF collected from a single sea urchin (animal 31) that had been injected first with LPS and then 360 h later with PG. CF was withdrawn at 96 h after each injection. Western blots were immunostained with an equal mixture of the three different anti-185 sera (anti-185-66, -68, and -71; 1/20,000). *A*, Full proteome maps (pI of 3–6; 40–150 kDa; 30-s film exposure) of the CF samples after LPS and PG challenges. *B*, Enlargements (pI of 4.0–6.3; 40–85 kDa; 5-min film exposure) of the boxed areas in *A*. The numbers shown in each panel are the relative expression intensities of five 185/333<sup>+</sup> spots that differed substantially in expression after LPS compared with PG injections.

other than disulfide bond formation that are resistant to the reducing treatments used in 1DE and 2DE. This conclusion is supported by studies of recombinant 185/333 proteins, in which the expression of a single form of 185/333 protein yields a range of expressed proteins with molecular masses corresponding to monomers, homodimers, and higher order oligomers (18).

Other data also suggest that the diversity of molecular masses evident among 185/333 proteins is increased by the expression of truncated molecules. Many of the SNPs and indels found previously in *185/333* transcripts are predicted to result in frameshifts, and the encoded proteins may be either truncated and/or have missense sequences (11, 13). No such frameshifts have been identified in the *185/333* genes. However, comparisons of gene and mRNA sequences from individual animals suggest that posttranscriptional modifications may be responsible for these missense proteins (15). If these messages are translated, it would explain why the three different anti-185 antisera used in the present study detected different subsets of 185/333 proteins. The antisera were raised against amino acid sequences in N terminus (element 1), the middle (element 7), and the C terminus (element 25a) of the predicted proteins. The antiserum targeted to the N terminus (anti-185-66) identified the largest number of 185/333 proteins on 2DE Western blots, presumably because its epitope is most likely to be present in every protein, including the shortest truncated forms. Of the 689 translated cDNA sequences identified to date, 676 contain the epitope within element 1 that is recognized by the anti-66 antiserum (11). Antisera directed toward more C-terminal regions recognize decreasing numbers of 185/333<sup>+</sup> spots, probably because these epitopes are not present on truncated proteins or those with missense sequence at the C terminus. Current cDNA sequence data from Terwilliger et al. (11) show that 660 of 689 cDNA sequences contain the anti-68 epitope in element 7, while only 375 cDNA sequences contain the anti-71 epitope in element 25a.

The variability evident in the molecular masses of 185/333 proteins was matched by substantial diversity in their pIs. One source for this diversity is found in mRNAs where sequence variability results in changes in charged amino acids at particular positions leading to predicted proteins with very similar molecular masses but different pIs (8, 11, 12). Terwilliger et al. (see supplemental figures S1 and S2 in Refs. 11, 12) identified numerous sequence positions in *185/333* mRNAs that encode two to four different amino acids, many with different charges. In the present study, similar subtle differences were detected by MS, which identified a number of 185/333 peptides that differ by just a single amino acid. 2DE Western blots also showed a variety of pIs for 185/333 proteins that had very similar molecular masses. In many cases, the charge variants making up a single molecular mass class of 185/333 proteins had pIs spanning the full pH range from 3 to 10, although most variants had a pI within the range of 3–6. The different pI forms were often equally spaced, suggesting that adjacent protein spots represent variants that differ by just a single charged residue.

Although the variability in pI detected in the present study agrees with the diversity seen among cDNA sequences, our proteomic data also provide evidence for additional posttranslational modifications. Some 185/333 proteins have identical isoelectric points but significantly different molecular masses. This suggests that individual 185/333 proteins may be conjugated with some other molecule(s) that alters their molecular mass but not their isoelectric point. Such conjugation would provide another explanation for why many 185/333<sup>+</sup> bands are much larger in molecular mass than predicted. However, the cDNA data suggest that it is also feasible for 185/333 proteins with significantly different se-

quences and different pIs to have the same molecular mass, resulting in a vertical ladder of spots on 2DE Western blots.

The diversity of 185/333 proteins explains our inability to unequivocally identify 185/333 proteins by MS of individual proteins isolated by 2DE. MS analysis of the 185/333<sup>+</sup> spots from 2DE gels did not identify any statistically significant matches to proteins in a custom database of 185/333 sequences. However, all of the 185/333 proteins isolated by 2DE yielded ion fragments that possessed mass/charge characteristics similar to those of known 185/333 protein sequences. The available evidence suggests that the existing database of 185/333 sequences may represent only a small fraction of the 185/333 variants present in sea urchin populations. Consequently, the chance of finding a statistically significant match in this restricted data set to a single 185/333 protein isolated from the CF of an individual sea urchin may be extremely low, even though the purity of the proteins isolated by 2DE would have been high. In other words, we may not have identified matches because the variability of 185/333 sequence means that many 185/333 proteins will not yet be in our database of 185/333 sequences and so cannot be identified by MS. Our difficulties in matching isolated proteins to the existing 185/333 database highlight the problems of employing MS techniques, which search for precise similarities between peptides to characterize hypervariable proteins. It is interesting that even though members of the Down syndrome cell adhesion molecule (DSCAM) family are important proteins in the immunological responses of insects, proteomic analyses of hemolymph extracts from immunologically challenged *Drosophila* have not yet been able to identify DSCAMs by MS.

To circumvent the problem of matching individual 185/333<sup>+</sup> spots to our restricted sequence data set, we employed "shotgun" MS (LC-MS/MS), in which all of the CF proteins within a particular molecular mass range from 1DE gels were analyzed simultaneously by MS. Unlike 2DE, in which the individual proteins spots analyzed by MS presumably contain just a single 185/333 isotype, the 1DE gel slices subjected to shotgun MS were likely to contain numerous different 185/333 isotypes. This greatly increased the chance that at least one of these isotypes would match a previously identified sequence in our 185/333 database. 1DE also has the advantage that it is not limited by the complex separation characteristics of 2DE, which, for instance, is relatively inefficient in separating hydrophobic proteins. Additionally, we used an alternative search engine (GPM as opposed to Mascot) with slightly altered search stringencies (three missed tryptic cleavages) to increase our chances of identifying matching 185/333 peptides in the available databases. This process successfully identified numerous peptides that matched precisely to sequences in the NCBI database without affecting the robustness of the sequence matches. The gel slices in which matching peptides were identified corresponded to regions on 1DE and 2DE Western blots that contained high concentrations of 185/333 proteins, which confirms the specificities of the anti-185 sera used in this study. However, shotgun MS probably still identified just a small fraction of the 185/333 peptides present within the peptide mixture, because of the limited numbers of potentially matching peptides in the currently available dataset. The limited sensitivity of MS in identifying highly variable proteins in the absence of comprehensive sequence databases also explains why our MS analysis only identified multiple forms of 185/333 proteins in one of the sea urchins analyzed, even though both 1DE and 2DE Western blotting revealed substantial intraindividual diversity in all sea urchins.

The difficulties associated with analyzing the diversity of 185/333 proteins were compounded by the substantial variation evident in the suites of 185/333 proteins expressed by different sea urchins.

None of the 13 sea urchins analyzed by 1DE Western blotting expressed the same pattern of 185/333<sup>+</sup> bands, even though there were a number of molecular mass forms that were shared by some individuals. On average, two to three major bands were evident in each sea urchin, along with numerous minor bands. The differences between individuals implies that there may be a variety of mechanisms acting in concert to produce novel arrays of 185/333 proteins in individual sea urchins. These factors could include the presence or absence of different 185/333 genes in different sea urchins, the presence of different alleles at a given 185/333 gene locus, differential gene expression of 185/333 family members, posttranscriptional processing or editing of the transcripts that insert frameshifts and SNPs (15), and posttranslational processing of the proteins (13).

Our data suggest that these potential mechanisms for molecular diversification may allow sea urchins to vary the suites of 185/333 proteins that they express in response to different types of immunological challenge. ELISA provided direct evidence that 185/333 expression increases after the injection of PAMPs, while 2DE analysis of 185/333 proteins from individual sea urchins identified many subtle changes in the patterns of 185/333 proteins responding to different PAMPs. There were clear differences in the pIs and molecular masses of the proteins expressed in response to LPS compared with those synthesized by the same sea urchin in response PG. These results agree with changes that are apparent in the cDNAs from animals responding to different PAMPs, including LPS,  $\beta$ -1,3-glucan, or dsRNA (11). Terwilliger et al. (11) showed that a diverse size range of 185/333 messages was present in the CF of immunoreactive sea urchins, but this broad expression profile changed to one single, major mRNA size after immune challenge. Sequence analysis of mRNAs from animals responding to immune challenge also indicated that the predominant 185/333 variants in CF are different before and after challenge. Such alterations in 185/333 transcription explain the changes detected during the present study in the suites of 185/333 proteins found in the CF of the same sea urchin responding to different PAMPs.

The identification of PAMP-specific responses suggests that 185/333 protein expression can be tailored to meet different forms of immunological challenge. The obvious implication is that 185/333 proteins are involved in some form of pathogen-specific immunological response. There is growing evidence for this level of immunological specificity among a variety of invertebrates (1, 2, 5, 27–29). For instance, bacterial immunization in bumble bees can induce acquired immune responses that are highly specific to the original, inoculating species of bacteria (29), although the mechanisms underlying this response are not known. In addition to 185/333 proteins, other large gene families may be capable of fine-scale discrimination between pathogens. These gene families encode variable lymphocyte receptors from agnathans (30–34), V region-containing chitin-binding proteins (VCBPs) from cephalochordates and tunicates (3, 35, 36), DSCAMs in insects (37, 38), and fibrinogen-related proteins (FREPs) from molluscs (39–42). A repertoire of >19,000 different DSCAMs is expressed in hemocytes of *Drosophila* and mosquitoes (37, 38), and thousands of FREPs are present in freshwater snails (39–42). None of these gene families is closely related to one another, suggesting that mechanisms for immune diversification have evolved many times and that additional unique systems are likely to be identified among other organisms in the future. In this context, our continuing investigations of 185/333 proteins will provide insights into how animals other than mammals adapt in the “arms race” against the microbes.

## Acknowledgments

We thank Dr. Virginia Brockton for assistance in processing and shipping protein samples, and Dr. Katherine Buckley for suggesting improvements to the manuscript. The research has been facilitated by access to the Australian Proteome Analysis Facility established under the Australian Government's Major National Research Facilities Program.

## Disclosures

The authors have no financial conflicts of interest.

## References

- Flajnik, M. F. 2004. Immunology: another manifestation of GOD. *Nature* 430: 157–158.
- Flajnik, M. F., and L. Du Pasquier. 2004. Evolution of innate and adaptive immunity: can we draw a line? *Trends Immunol.* 25: 640–644.
- Litman, G. W., J. P. Cannon, and L. J. Dishaw. 2005. Reconstructing immune phylogeny: new perspectives. *Nat. Rev. Immunol.* 5: 866–879.
- Hibino, T., M. Loza-Coll, C. Messier, A. J. Majeske, A. H. Cohen, D. P. Terwilliger, K. M. Buckley, V. Brockton, S. V. Nair, K. Berney, et al. 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev. Biol.* 300: 349–365.
- Rast, J. P., L. C. Smith, M. Loza-Coll, T. Hibino, and G. W. Litman. 2006. Genomic insights into the immune system of the sea urchin. *Science* 314: 952–956.
- Sodergren, E., G. M. Weinstock, E. H. Davidson, R. A. Cameron, R. A. Gibbs, R. C. Angerer, L. M. Angerer, M. I. Arnone, D. R. Burgess, R. D. Burke, et al. 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* 314: 941–952.
- Fugmann, S. D., C. Messier, L. A. Novack, R. A. Cameron, and J. P. Rast. 2006. An ancient evolutionary origin of the RAG1/2 gene locus. *Proc. Natl. Acad. Sci. USA* 103: 3728–3733.
- Nair, S. V., H. Del Valle, P. S. Gross, D. P. Terwilliger, and L. C. Smith. 2005. Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin: identification of unexpected immune diversity in an invertebrate. *Physiol. Genomics* 22: 33–47.
- Rast, J. P., Z. Pancer, and E. H. Davidson. 2000. New approaches towards an understanding of deuterostome immunity. *Curr. Top. Microbiol. Immunol.* 248: 3–16.
- Smith, L. C., L. Chang, R. J. Britten, and E. H. Davidson. 1996. Sea urchin genes expressed in activated coelomocytes are identified by expressed sequence tags: complement homologues and other putative immune response genes suggest immune system homology within the deuterostomes. *J. Immunol.* 156: 593–602.
- Terwilliger, D. P., K. M. Buckley, V. Brockton, N. J. Ritter, and L. C. Smith. 2007. Distinctive expression patterns of 185/333 genes in the purple sea urchin, *Strongylocentrotus purpuratus*: an unexpectedly diverse family of transcripts in response to LPS,  $\beta$ -1,3-glucan, and dsRNA. *BMC Mol. Biol.* 8: 16.
- Terwilliger, D. P., K. M. Buckley, D. Mehta, P. G. Moorjani, and L. C. Smith. 2006. Unexpected diversity displayed in cDNAs expressed by the immune cells of the purple sea urchin, *Strongylocentrotus purpuratus*. *Physiol. Genomics* 26: 134–144.
- Buckley, K. M., and L. C. Smith. 2007. Extraordinary diversity among members of the large gene family, 185/333, from the purple sea urchin, *Strongylocentrotus purpuratus*. *BMC Mol. Biol.* 8: 68.
- Buckley, K. M., S. Munshaw, T. B. Kepler, and L. C. Smith. 2008. The 185/333 gene family is a rapidly diversifying host-defense gene cluster in the purple sea urchin, *Strongylocentrotus purpuratus*. *J. Mol. Biol.* 379: 912–928.
- Buckley, K. M., D. P. Terwilliger, and L. C. Smith. Sequence variations in 185/333 messages from the purple sea urchin suggest posttranscriptional modifications to increase diversity. *J. Immunol.* 181: 8585–8594.
- Tsai, H., and L. A. Bobek. 1998. Human salivary histatins: promising anti-fungal therapeutic agents. *Crit. Rev. Oral Biol. Med.* 9: 480–497.
- Xu, T., E. Telsler, R. F. Troxler, and F. G. Oppenheim. 1990. Primary structure and anticandidal activity of the major histatin from parotid secretion of the sub-human primate, *Macaca fascicularis*. *J. Dent. Res.* 69: 1717–1723.
- Brockton, V., J. H. Henson, D. A. Raftos, A. J. Majeske, Y. O. Kim, and L. C. Smith. 2008. Localization and diversity of 185/333 proteins from the purple sea urchin – unexpected protein-size range and protein expression in a new coelomocyte type. *J. Cell Sci.* 121: 339–348.
- Clow, L. A., P. S. Gross, C. S. Shih, and L. C. Smith. 2000. Expression of SpC3, the sea urchin complement component, in response to lipopolysaccharide. *Immunogenetics* 51: 1021–1033.
- Gross, P. S., W. Z. Al-Sharif, L. A. Clow, and L. C. Smith. 1999. Echinoderm immunity and the evolution of the complement system. *Dev. Comp. Immunol.* 23: 429–442.
- Clow, L. A., D. A. Raftos, P. S. Gross, and L. C. Smith. 2004. The sea urchin complement homologue, SpC3, functions as an opsonin. *J. Exp. Biol.* 207: 2147–2155.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.
- Andon, N. L., S. Hollingworth, A. Koller, A. J. Greenland, J. R. Yates, and P. A. Haynes. 2002. Proteomic characterization of wheat amyloplasts using identification of proteins by tandem mass spectrometry. *Proteomics* 2: 1156–1168.

24. Green, P., A. Luty, S. Nair, J. Radford, and D. Raftos. 2006. A second form of collagenous lectin from the tunicate, *Styela plicata*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 144: 343–350.
25. Green, P. L., S. V. Nair, and D. A. Raftos. 2003. Secretion of a collectin-like protein in tunicates is enhanced during inflammatory responses. *Dev. Comp. Immunol.* 27: 3–9.
26. Miller, I., J. Crawford, and E. Gianazza. 2006. Protein stains for proteomic applications: which, when, why? *Proteomics* 6: 5385–5408.
27. Kurtz, J., and K. Franz. 2003. Innate defence: evidence for memory in invertebrate immunity. *Nature* 425: 37–38.
28. Kurtz, J., and S. A. Armitage. 2006. Alternative adaptive immunity in invertebrates. *Trends Immunol.* 27: 493–496.
29. Sadd, B. M., and P. Schmid-Hempel. 2006. Insect immunity shows specificity in protection upon secondary pathogen exposure. *Curr. Biol.* 16: 1206–1210.
30. Litman, G. W., J. P. Cannon, and J. P. Rast. 2005. New insights into alternative mechanisms of immune receptor diversification. *Adv. Immunol.* 87: 209–236.
31. Litman, G. W., N. A. Hawke, and J. A. Yoder. 2001. Novel immune-type receptor genes. *Immunol. Rev.* 181: 250–259.
32. Pancer, Z., C. T. Amemiya, G. R. Ehrhardt, J. Ceitlin, G. L. Gartland, and M. D. Cooper. 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* 430: 174–180.
33. Pancer, Z., W. E. Mayer, J. Klein, and M. D. Cooper. 2004. Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. *Proc. Natl. Acad. Sci. USA* 101: 13273–13278.
34. Pancer, Z., N. R. Saha, J. Kasamatsu, T. Suzuki, C. T. Amemiya, M. Kasahara, and M. D. Cooper. 2005. Variable lymphocyte receptors in hagfish. *Proc. Natl. Acad. Sci. USA* 102: 9224–9229.
35. Cannon, J. P., R. N. Haire, and G. W. Litman. 2002. Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. *Nat. Immunol.* 3: 1200–1207.
36. Cannon, J. P., R. N. Haire, J. P. Rast, and G. W. Litman. 2004. The phylogenetic origins of the antigen-binding receptors and somatic diversification mechanisms. *Immunol. Rev.* 200: 12–22.
37. Dong, Y., H. E. Taylor, and G. Dimopoulos. 2006. AgDscam, a hypervariable immunoglobulin domain-containing receptor of the *Anopheles gambiae* innate immune system. *PLoS Biol.* 4: e229.
38. Watson, F. L., R. Puttmann-Holgado, F. Thomas, D. L. Lamar, M. Hughes, M. Kondo, V. I. Rebel, and D. Schmucker. 2005. Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* 309: 1874–1878.
39. Adema, C. M., L. A. Hertel, R. D. Miller, and E. S. Loker. 1997. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. *Proc. Natl. Acad. Sci. USA* 94: 8691–8696.
40. Leonard, P. M., C. M. Adema, S. M. Zhang, and E. S. Loker. 2001. Structure of two FREP genes that combine IgSF and fibrinogen domains, with comments on diversity of the FREP gene family in the snail *Biomphalaria glabrata*. *Gene* 269: 155–165.
41. Zhang, S. M., and E. S. Loker. 2004. Representation of an immune responsive gene family encoding fibrinogen-related proteins in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*. *Gene* 341: 255–266.
42. Zhang, S. M., and E. S. Loker. 2003. The FREP gene family in the snail *Biomphalaria glabrata*: additional members, and evidence consistent with alternative splicing and FREP retrosequences: fibrinogen-related proteins. *Dev. Comp. Immunol.* 27: 175–187.
43. Fenyo, D., and R. C. Beavis. 2003. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal. Chem.* 75: 768–774.