Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
It has recently been called to our attention that the word 'loa', used for centuries, first by Africans and later by parasitologists, to refer to the 'eye worm', also appears in the terminology of voodoo or Vodun where it refers to a large pantheon of deities that may possess one's soul or being. The word 'loa' may be derived from the Yoruba word 'lawo' meaning 'mystery' but according to anthropological one remains a matter of etymological source for the other. Whether the helminthological loa predated the anthropological one remains a matter of conjecture.

Near the turn of the century a 29-year-old physician sought medical care for returning to the United States, he had resided in Panama for almost six years. Sir William Osler examined the patient and noted that:

His general condition was very good, considering that he had had severe dysentery and an irregular fever for more than two months. The liver was slightly enlarged anteriorly but not specially sensitive. Posteriorly there seemed to be a very distinct extension of the duodenal [sic] upward. He had six or eight mucoid stools with traces of blood daily. I saw him subsequently on four occasions and the symptoms remained practically the same. The temperature rose each day to about 103°... The suspicion entertained at first that he had abscess of the liver was gradually confirmed, and on March 22nd Dr Tiffany aspirated, and then incised and drained two large abscess cavities in the right lobe of the liver. The pus was thick, of creamy consistence, in color, in places slightly bile-stained, but it has not the reddish-brown and anchovy-sauce-like appearance presented by the pus in many cases of hepatic abscesses. I made an examination of the pus at the Biological Laboratory, within three-quarters of an hour of its withdrawal and found in it, in large numbers, the amoebae which Karulis had described.

After the operation the dysenteric symptoms did not abate in the slightest; he continued to have from eight to sixteen movements daily. On each day there were found in these stools many characteristic examples of the amoebae... The general character of the amoebae corresponded in every particular with those found in the liver.

Osler concluded that:

It is impossible to speak as yet with any certainty as to the relation of these organisms to the disease. The subject is deserving of extended study.

The patient died on 5 April. It was one year later, in 1891, that Osler's colleagues Councilman and Lafleur published the key paper 'Amebic Dysentery' in the Johns Hopkins Hospital Report. In it they described the pathology of amebic colitis and liver abscess, including the invasion through the intestinal epithelium by the amoebae, and emphasized the lack of contribution of intestinal bacteria to the colitis and liver abscess. The occasion of the 100th anniversary of this landmark monograph is an opportune time to review recent advances in the study of this protozoon that remains a worldwide cause of human suffering. We will arbitrarily limit our discussion to the five surface proteins whose sequences have been reported, and not cover other potentially important but less well-characterized surface molecules that have been described, such as the fibronectin receptor, neuraminidase, phospholipase A2, surface sialoglycoproteins, the 112 kilodalton (kDa) adhesin, the chitotriose lectin and the 30 kDa and 90 kDa surface antigens.

Serine-rich E. histolytica Protein

A serine-rich protein has been identified, serine, and by Stanley and colleagues by differential hybridization screening of an E. histolytica complementary (cDNA) library. The library was constructed from E. histolytica strain HM1: IMSS and screened with radiolabeled cDNA from HM1: IMSS and the non-E histolytica Laredo strain. Several clones that hy-
bridged with HM1:IMSS and not Laredo cDNA were identified and one 0.7 kilobase (kb) cDNA sequenced. This E. histolytica-specific cDNA contains two open reading frames (ORF). Antibodies produced to trpE fusion proteins using the first but not the second ORF recognized E. histolytica proteins on western blots. Human immune sera and rabbit anti-E. histolytica antiserum recognized the fusion protein using ORF1 and not ORF2, which suggests that ORF2 is not translated in the trophozoites. Surprisingly, antibodies raised to the ORF1 fusion protein identify E. histolytica proteins, which at 46 and 52 kDa are twice the size predicted from the cDNA sequence (25 kDa). Post-translational modifications have been speculated to be involved but have not been shown to account for the difference in molecular weight. The proteins recognized by the ORF1 fusion protein antiserum were present in a 100,000 g pellet of sonicated trophozoites and therefore are likely to be associated with membranes or vesicles.

The structure of the serine-rich protein encoded by ORF1 shares similarities with the malaria circumsporozoite proteins. The amino terminus contains a 13 amino acid putative signal sequence that is followed by a region of charged amino acids. The center of the protein contains tandem octapeptide and decapeptide repeats, with a 13 amino acid hydrophobic carboxy-terminus domain. There are no potential sites for N-linked glycosylation.

A function for the serine-rich protein in amebic adhesion has been proposed based on the finding that a 1:100 dilution of antiserum to the ORF1-trpE fusion protein inhibits adherence to Chinese hamster ovary (CHO) 1021 cells by 70%. The same investigators have shown unique Southern blot patterns for the 125 kDa antigen in pathogenic and nonpathogenic E. histolytica have been compared. The 125 kDa antigen sequence of the pathogenic HM1:IMSS strains reported by Edman et al. and Tannich et al. in this region differ by five amino acid substitutions (1.3%). However, six amino acid substitutions (12.9%) have been seen when the sequence from the pathogen is compared to that of the nonpathogen.

The two 125 kDa antigen sequences from nonpathogenic strains are thus more similar to one another than to the sequences from the pathogenic strain HM1:IMSS. Tannich et al. had earlier shown unique Southern blot patterns for the 125 kDa antigen in pathogenic and nonpathogenic strains for nine different E. histolytica strains. They recently extended these studies to an additional 48 strains, and only one of the pathogenic strain had a nonpathogenic restriction pattern. These analyses of the 125 kDa antigen gene support the concept that pathogenic and nonpathogenic E. histolytica strains are genetically distinct.

### Galactose Lectin

The galactose lectin is an interesting molecule to study by virtue of its central role in adhesion and contact-dependent cytolysis. There is considerable evidence to support a critical role for a galactose lectin in pathogenesis. Ravdin and Guerrant have shown that...
adherence to, and killing of, CHO cells by E. histolytica are completely blocked by galactose or N-acetyl-D-galactosamine. Phagocytosis of bacteria and erythrocytes, and binding to human colonic mucins can also be blocked by galactose. Finally, CHO cell mutants lacking terminal N-acetyl lactosamine residues have been found to be resistant to amebic adherence and killing. Purified galactose lectin binds to CHO cells and competitively inhibits amebic attachment to CHO cells in a galactose-sensitive manner. Moreover, antiserum raised against the purified galactose lectin inhibits amebic adherence to CHO cells by 100% (Ref. 22) and blocks the binding of human colonic mucin glycoproteins to amebic trophozoites.

The galactose lectin is a 260 kDa heterodimer consisting of a 170 kDa subunit and a 35 kDa subunit linked by disulfide bonds. Polyclonal antisera and monoclonal antibodies raised against the native or denatured lectin recognize only the 170 kDa subunit and it appears that the 35 kDa subunit is not immunogenic in humans, mice, rabbits or gerbils. The light subunit has been shown to bind fibronectin on western blots (B.J. Mann and W.A. Petri, Jr, unpublished), and may be the fibronectin receptor identified by Talamas-Rohana and Meza. However, owing to the lack of specific antibody against the 35 kDa subunit, this has not yet been shown for intact trophozoites.

Monoclonal antibodies specific for the 170 kDa subunit block or enhance amebic adherence to CHO cells and colonic mucins, suggesting that the adherence domain resides within the heavy subunit. The ability of anti-lectin antibodies to dramatically increase the galactose-binding activity of the lectin indicates that its activity could be subject to conformational control. These same antibodies distinguish pathogenic from nonpathogenic strains of E. histolytica. All six epitopes defined by the murine monoclonal antibodies were present on the 16 pathogenic strains tested but only two of these six epitopes were present on the nonpathogenic strains. The presence of the galactose lectin in nonpathogenic strains may reflect its requirement for amebic colonization of the large bowel, because of its function as the receptor for human colonic mucins.

The Galactose Lectin as a Protective Antigen

Primates and humans are the only known reservoirs of E. histolytica. An anti-amebic vaccine that prevents large bowel colonization could theoretically eliminate E. histolytica as a cause of human disease. The galactose lectin has several properties that make it an attractive subunit vaccine candidate: (1) amino acid differences in hg12 compared to hg11. Variant amino acids are depicted by vertical lines at the residue number in hg11. Conservative amino acid changes are indicated by short vertical lines. Nonconservative changes are depicted by tall vertical lines. Insertions in hg12 with respect to hg11 (closed circles) and deletions in hg12 with respect to hg11 (closed triangles) are indicated. The number to the right of the symbols indicates the number of residues inserted or deleted. Glycosylation sites in both hg11 and hg12 (closed squares) and only in hg11 (open squares) are shown. (b) A hydrophobicity plot of the derived amino acid sequence of hg11. Residue numbers refer to hg11. (c) The putative structural domains of hg11 and hg12. Cysteine- and tryptophan-rich (C-W-rich); cysteine-free (C-free); cysteine-rich (C-rich); transmembrane domain (TM).

The molecular cloning and sequencing of the 170 kDa subunit of the galactose lectin from pathogenic E. histolytica has recently been reported by our laboratory (B.J. Mann et al., abstract* (Ref. 28) and Tannich et al. A comparison of the amino acid sequences described by these two groups revealed that they are only 87.6% homologous (Fig. 1a), suggesting the presence of at least two genes encoding the galactose lectin gene family.
170 kDa subunit. The presence of more than one gene encoding the 170 kDa subunit has been suggested by the observed microheterogeneity in the amino acid sequences obtained by sequential Edman degradation of the amino terminus and cyanogen bromide (CNBr) fragments, and by the complex restriction fragment patterns seen on Southern blots probed with the cDNA for the 170 kDa subunit. The copy number of the genes encoding the 170 kDa subunit and their genomic organization have yet to be determined. The sequences reported in Fig. 1, which represent two different genes encoding the 170 kDa subunit of the galactose lectin, are proposed to be called hgl1 (Ref. 28) and hgl2 (Ref. 29) for heavy subunit galactose lectin, hgl1 is a cDNA clone isolated from the pathogenic strain HML-IMSS. The DNA encoding the first 482 amino acids of hgl1 was isolated from the HML-IMSS genomic DNA with the remainder of the gene sequence determined from a cDNA clone of pathogenic strain H-302.NIH. A combination of sequencing and polymerase chain reaction (PCR) experiments has revealed that hgl1 does not contain any introns.

The calculated relative molecular masses of hgl1 and hgl2 are 143 241 Da and 143 780 Da, respectively, and their individual amino acid composition and overall structure are also very similar. The nonuniform distribution of hydrophobic amino acids of hgl1 and hgl2 suggests the existence of five structural domains (Fig. 1b, 1c): (1) At the amino terminus of the mature protein is a 187 amino acid cysteine (3.2 mol %) - and tryptophan (2.1 mole %) - rich amino-terminal domain; (2) the second cysteine-free domain, corresponding to hgl1 residues 188-378, contains alternating hydrophobic and hydrophilic stretches of amino acids and is the most variable between the two proteins. A substantial nonconservative change is the addition of eight amino acids at amino acid 285 in hgl1; (3) the third structural domain is the cysteine-rich domain, which contains 10.8 mol % of cysteine. The number and position of the cysteine residues have been conserved with the exception of amino acid 804, which is a serine residue in hgl2. It has been shown that the cysteine residues play an important role in the observed protease resistance of the 170 kDa subunit and may be important for parasite survival in the gut. A putative transmembrane domain contains only two conservative changes between hgl1 and hgl2 and (5) The carboxy-terminal domain is a putative cytoplasmic tail that is also highly conserved. There are a total of 11 threonine, serine and tyrosine residues that are potential sites for phosphorylation in the cytoplasmic domain of hgl2 and only nine such residues in hgl1. Tyrosine residue 1261, present in both hgl1 and hgl2, is surrounded by an amino acid sequence that shares identity with the auto phosphorylation site of the epidermal growth factor receptor. It is interesting to speculate that phosphorylation of the carboxy-terminal domain may be a mechanism of activation of the lectin, and that the degree of phosphorylation could functionally distinguish the different forms of the 170 kDa subunit. The high degree of conservation of the cytoplasmic domain also suggests that the lectin may interact with a cytoplasmic protein in another route of signal transduction.

The 170 kDa subunit has been shown to be glycosylated and the treatment of amebae with tunicamycin results in almost a complete loss of adherence to CHO cells, showing that there is a critical role for N-linked glycosylation in amebic adherence. Sixteen potential glycosylation sites have been found in hgl1 (Fig. 1c) and hgl2 has only nine sites, although all of these sites are conserved in the molecule. The number and position of potential glycosylation sites could also indicate that the different members of the lectin gene family are functionally distinct.

The carbohydrate-binding portion of the molecule has not yet been identified and the 170 kDa sequence does not share any significant amino acid identities with the conserved regions of the carbohydrate-binding domain of C- or S-type lectins, the E. coli α-1,4-galactose-binding plus or plant lectins that have been sequenced thus far. Although the evidence suggests that the carbohydrate-binding activity resides in the 170 kDa subunit, it is possible that it is contained within the 35 kDa subunit or is formed by the heterodimer of the two subunits.

### Future Investigations

The recent discovery of two different genes encoding the 170 kDa subunit of the galactose lectin opens up new avenues of investigation. The number of family members, physical organisation of the gene copies, functional differences and possible transcriptional differences can be examined. A comparison of the similarities and differences between different members of the gene family should provide valuable information regarding the adherence, cytolytic and cell-signaling functions of the lectin. The contributions of the light subunit to the functional activities of the galactose lectin remain unresolved as do the biological functions of many of the other surface antigens identified to date. Continued characterization of the galactose lectin and other surface proteins should provide greater understanding of the immunology, pathogenicity and molecular biology of this globally important parasite.