A New Gene Family (ariel) Encodes Asparagine-Rich
Entamoeba histolytica Antigens, Which Resemble the Amoebic Vaccine Candidate Serine-Rich E. histolytica Protein

ZHIMING MAI AND JOHN SAMUELSON*

Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts 02125

Received 19 May 1997/Returned for modification 23 July 1997/Accepted 14 October 1997

Entamoeba histolytica is a protozoan parasite that causes tens of millions of cases of amebic colitis and liver abscess each year in the developing world (21). With the goal of preventing amebic infection, human antiamebic sera have been used to identify vaccine candidates on the parasite surface. One such candidate is the serine-rich E. histolytica protein (SREHP), also known as the K2 protein, which is located on the parasite surface and is composed of an amino-terminal leader sequence and a hydrophobic carboxy-terminal anchor surrounding a series of tetrapeptide and octapeptide repeats (12, 25). SREHP repeats, which are recognized by most human antiamebic sera, are composed of hydrophilic and acidic amino acids (20). SREHP repeats immunize gerbils and SCID mice against liver challenge with E. histolytica trophozoites, while SREHP gene polymorphisms may be used to distinguish axenic strains and clinical isolates of E. histolytica (4, 22, 30).

The E. histolytica chitinase, encoded by a single-copy chit genes, contains a series of degenerate heptapeptide repeats (D/E-I/S/T/V-K-H/P-D/E-S-S) which are located between an amino-terminal leader sequence and a carboxy-half catalytic domain (6). Although the amino acid composition of the chitinase repeats is almost identical to that of the SREHP repeats, the amino acid sequence and codon usage of the chitinase repeats are different from those of SREHP, indicating their independent genetic origin. The amebic chitinase genes are polymorphic from one clinical isolate to the next, making chit genes (in conjunction with srehp genes) useful for molecular epidemiological studies (22).

To discover another single-copy, polymorphic repeat gene from E. histolytica, the coding region of the SREHP gene was used to identify a gray (weakly hybridizing) plaque in an E. histolytica HM-1 genomic DNA (gDNA) library (13, 25). The plaque encoded a novel repeat protein which was called Ariel, for asparagine-rich E. histolytica protein (Fig. 1). Four more unique ariel genes were amplified from E. histolytica gDNA by using PCR with sense (GAGTCTATTCATGAAAAT) and antisense (CCACAAATAATAGCAAGGAAG) primers flanking the repeats in the ariel gDNA library clone (14). Three unique ariel genes and one ariel gene matching a PCR product were found by screening an E. histolytica cDNA library and by using 5′ and 3′ rapid amplification of cDNA ends (RACE) with the same antisense and sense ariel primers, respectively (8). The longest ariel gene, which encodes a 23-kDa protein with 16 octapeptide repeats, was labeled ariel-1, while the shortest, which encodes an 11-kDa protein with 2 octapeptide repeats, was labeled ariel-8.

Multicopy ariel genes are genetically related to the srehp gene. Eight ariel genes encode predicted proteins identical over the first 58 amino acids (aa) at the amino termini and the last 29 aa at the carboxy termini, which show 55% and 76% positional identities, respectively, with those of SREHP (Fig. 1) (25). The ariel genes, then, belong to the same family of genes as srehp and sredp, which encodes the serine-rich protein of Entamoeba dispar (formerly known as nonpathogenic E. histolytica) (7, 12). Conserved in the amino termini of the Aerial proteins and SREHP are a putative signal sequence and a series of charged amino acids prior to the repeats (Fig. 1) (29). Conserved at the carboxy termini of the Aerial proteins and SREHP are putative hydrophobic anchors (11, 25). Whether Aerial proteins, like SREHP, are phosphorylated and acetylated and have O-linked terminal N-acetylglucosamine residues also remains to be determined (26).

The hydrophilic and acidic octapeptide repeats of the Aerial proteins resemble those of SREHP in amino acid composition, sequence, and codon usage (Fig. 1 and data not shown) (12, 25). However, the asparagine residues frequent in Aerial repeats are absent in the SREHP repeats, and the tetrapeptide repeats present in SREHP are absent in the Aerial proteins. The Aerial octapeptide repeats (D/N/S-E-S-S-D/N-K/P) are identical in four or five positions to those of SREHP (E-A-S/S-D/T-D/N-K/P), while the last 4 amino acids of the Aerial octapeptide repeats are identical in three or four positions to the SREHP tetrapeptide (D-N-K-P). The same D-N-K-P tetrapeptide is tandemly repeated in protein A of Staphylococcus aureus (28). These results suggest that the Aerial repeats may contain some antigens in common with and some distinct from those of SREHP.

Constitutive expression of ariel genes. Four ariel genes are expressed by cultured trophozoites, because their cDNAs were obtained from the cDNA library or by RACE (8). Multiple ariel genes were expressed simultaneously by trophozoites cloned on soft agar, including ariel-1 (major 440-bp product) and numerous smaller (faint) ariel genes (Fig. 2) (19). Because...
ariel reverse transcription (RT)-PCR products from each amebic clone look similar, it appears that the ariel genes are constitutively rather than variantly expressed.

Antigenicity of Ariel repeats. A prediction of the Hoppe and Woods plot is that the hydrophilic repeats of the Ariel proteins are immunogenic, as has been shown for SREHP (11, 20, 25, 30). Pooled sera from patients with amebic liver abscesses reacted on Western blots with the Ariel repeats expressed as a glutathione S-transferase (GST) fusion protein (Fig. 3) (10, 23). Although it is possible that the human antibodies were raised against another amebic immunogen such as SREHP, it appears that the Ariel repeats are antigenic. Antiamebic human sera also recognized GST fusion proteins containing *E. histolytica* chitinase repeats (weakly) or alcohol dehydrogenase 1 at the carboxy termini but did not react with GST alone (6, 13). Whether an *E. histolytica* Ariel protein, like SREHP, may be

**FIG. 1.** Alignment of amino acids, in single letter code, of eight predicted *E. histolytica* Ariel proteins (Ariel-1 to Ariel-8) with those of SREHP and SREDP. Because the Ariel proteins differ only in the number and sequence of the octapeptide repeats (shaded boxes), single amino- and carboxy-terminal Ariel sequences are shown. Sequences identical to Ariel-1 are marked with dashes, gaps are marked with dots, and COOH-terminal stop codons are marked with asterisks. Arrows indicate the locations of primers used for PCR and RACE. Tetrapeptide and octapeptide repeats of SREHP and SREDP are marked with unshaded boxes. Ariel-1, Ariel-2, Ariel-4, and Ariel-5 are derived from a cDNA library or RACE, while Ariel-3, Ariel-6, Ariel-7, and Ariel-8 are derived from a gDNA library or PCR products.

**FIG. 2.** RT-PCR products from mRNAs extracted from *E. histolytica* parasites cloned on soft agar with sense and antisense ariel primers (20). Lanes M, 100-bp ladders; lanes 1 to 5, different clones; lane 6, uncloned parent strain; lane 7, gDNA PCR product from the uncloned strain. Controls with a single sense or antisense ariel primer did not produce RT-PCR products (data not shown). The prominent 440-bp band corresponds to the ariel-1 gene, while lighter bands at lower molecular weights include those predicted by other ariel gene clones. Control RT-PCR with srehp, chit, and amebapore primers each produced a single product (data not shown) (6, 15, 25).

**FIG. 3.** Western blot of recombinant GST-Ariel with human antiamebic sera. Pooled human antiamebic sera reacted strongly with GST-Ariel repeats (Ar) and GST-*E. histolytica* chitinase repeats (Ch), and not at all with control GST (Co). Lane M, protein standards.
useful as a serodiagnostic reagent for amebic infection or as a component of an antiamoebic vaccine remains to be determined (11, 20, 25, 30). Whether antiamoebic sera recognize antigens with few repeats (e.g., Ariel-8) also remains to be determined.

**New conclusions about antigenic diversity in *E. histolytica***

(i) The same ancestral gene may lead to two different repeat antigens (Ariel and SREHP) with common amino and carboxy termini surrounding distinctive repeats. This is similar to what has been described for merozoite surface protein 1 (MSP-1) and S-antigen of *Plasmodium falciparum* (1, 2, 16).

(ii) Amebic repeat antigens may be encoded by multiple *ariel* genes, making these genes useless for molecular epidemiological studies. In contrast, polymorphic *chit* and *srehp* genes are present in one or two copies and so are useful for epidemiology (4, 6, 12, 22). Single-copy polymorphic repeat genes of *P. falciparum* for epidemiological studies, encode MSP-1, MSP-2, S-antigens, and CSP (1, 2, 5, 9, 16).

(iii) At least four *ariel* genes are constitutively expressed, so antigenic diversity may exist within a single parasite isolate rather than between separate parasite isolates (as is the case for SREHP and amebic chitinase) (4, 22). It is unclear whether amebae have any variably expressed antigens, such as the *var* proteins of *P. falciparum*, the variable surface proteins of *Giardia lamblia*, and the variable surface glycoproteins of *Trypanosoma brucei* (3, 18, 27).

This work was supported in part by a grant from the NIH.

Thanks to Sam Stanley and colleagues at Washington University for extensive characterization of SREHP and to Egbert Tannich of the Nocht Institute in Hamburg for the *E. histolytica* cDNA library.

**REFERENCES**


Editor: S. H. E. Kaufmann

Downloaded from http://iai.asm.org/ on September 21, 2017 by guest