Summary
Entamoeba histolytica, an enteric protozoan parasite causes serious intestinal amoebic colitis and extra-intestinal amoebic liver abscess in most of the developing countries. About 50 million people worldwide develop clinically invasive amoebiasis with 40,000-100,000 deaths per year.

Basic research on the parasite has received a boost with the genome project. Efforts are continuing to improve the current genome assembly and annotation of E. histolytica and complete the genome sequence of other species of Entamoeba. With detailed identification of repeat elements in the genome it becomes easier to generate accurate gene prediction and annotation. Several families of transposable elements (TEs) have been identified in the genome of E. histolytica. The most abundant of these is the non-Long Terminal Repeat (LTR) retrotransposons. Three families of autonomous non-LTR retrotransposons called EhLINEs and their non autonomous partner EhSINEs have been identified in the E. histolytica genome. EhLINE1 constitutes the largest family of TEs in this parasite. SINEs utilize the transposition functions encoded by their partner LINE for their own transposition. Although EhLINE1 is present in a few hundred copies per genome, it is not known whether a single active copy of this element exist in the E. histolytica genome since most copies have accumulated mutations including large deletions. EhLINE1 encodes two non-overlapping ORFs i.e. ORF1 and ORF2. ORF2 contains centrally located reverse transcriptase (RT) domain and C-terminal endonuclease (EN) domain which resembles the type IIS restriction enzymes. The endonuclease encoded by this EN domain was found to nick pBS supercoiled DNA and a linear DNA fragment containing insertion site of EhSINE1 in vitro. It contains the conserved PD..D motif found in type IIS enzymes. Mutation of the first Asp to Ala in this motif causes significant loss of enzyme activity in the mutant (ENM).

In the present study we have attempted to understand the detailed biochemical characteristics of this endonuclease (EN). We have artificially reconstructed a functional full-length copy of EhLINE1 and obtained E. histolytica transformants expressing both ORFs. We have demonstrated active retrotransposition in these retrotransposition-competent cells. Our main findings are described below.

1. Enzyme kinetics performed with pBS supercoiled DNA under steady state revealed a significant burst phase followed by a slower steady-state phase, indicating that release of product could be the slower step in this reaction.

2. For circular supercoiled DNA the $K_m$ was $2.6 \times 10^{-8}$ M and the $k_{cat}$ was $1.6 \times 10^2$ sec$^{-1}$. For linear E. histolytica DNA substrate the $K_m$ and $K_{cat}$ values were $1.3 \times$
10^8 M and 2.2 x 10^{-4} \text{ sec}^{-1} \text{ respectively. The } k_m \text{ values are in the range reported for restriction enzymes.}

3. The endonuclease behaved as a monomer and did not dimerize in the presence of DNA.

4. Mg^{2+} is required for endonuclease activity. Enzyme activity dropped to 60% with Mg^{2+} and no activity was detected with any other divalent metal ion. Fluorescence quenching experiments were performed by taking advantage of the three tryptophan residues in EN. Conversion of PD..D motif to PA..D did not significantly change the $K_d$ of Mg^{2+} but caused 15-fold change in $K_d$ of Mn^{2+}. This implies that the first Asp residue may play a role in the binding of Mn^{2+}. However, it may not be involved in binding to Mg^{2+} and Ca^{2+}. Similar results have been observed in PvuII restriction enzyme where first Asp residue was not found to play role in binding of Mg^{2+}.

5. Near-UV CD spectroscopy and SPR studies revealed that EN interacts with DNA. A change in CD spectrum and response unit (RU) in SPR was recorded when EN interacted with DNA. ENM did not interact with DNA due to the hydrophobic core of the protein being disturbed.

6. Recombinant ORF2 showed EN and RT activities.

7. Reconstructed full-length copy of EhLINE1 was transfected in *E. histolytica* and found to transcribe the full-length 4.8 kb transcript but no detectable polypeptide.

8. ORF1p was found to be expressed in *E. histolytica* trophozoites constitutively but ORF2p was not detected. ORF2p was efficiently expressed in *E. histolytica* transfected with a construct containing ORF2 under the control of tetracycline-inducible promoter. In the presence of tetracycline this cell-line expressed both ORFs efficiently.

9. Retrotransposition could be triggered in this cell line when both the proteins were available. A tagged-EhSINE1 copy was used to monitor retrotransposition in *vivo*. It was found to insert exactly at the same target site as used by the genomic SINE. Insertion was accompanied by target site duplication- a hallmark of retrotransposition. Thus this system exactly mimicked the retrotransposition process of *E. histolytica* and is an excellent model system to understand the detailed mechanism of retrotransposition in *E. histolytica*.