A Proteomics Jurassic Park:  
The isolation of proteins from microorganisms encapsulated in amber from the Oligo-Miocene epoch 30-40 million years ago  

Gary B. Smekal1,3, Georgia O. Poinar Jr.2, Feixia Chu3 and Pier Giorgio Righetti4  
1 Harvard Catalyst, Harvard Clinical and Translational Science Center, Laboratory for Innovative Translational Technologies, Boston, MA, USA.  
2 Oregon State University, Department of Zoology, Corvallis, OR, USA.  
3 University of New Hampshire, Hubbard Center for Genome Studies, Durham, NH, USA.  
4 Politecnico di Milano, Department of Chemistry, Materials, and Chemical Engineering, Milan, Italy.  

ABSTRACT: Recent reports of peptides isolated from the mineralized skeletal elements and alleged “soft tissues” from *Tyrannosaurus rex* have invited controversy over whether proteins can endure geological time spans. However, soft tissues are not replaced by minerals in amber, which otherwise occurs during lithification. Organisms engulfed in terpenes within minutes or seconds were rapidly dehydrated, a prerequisite to preservation, thus forbidding any chemical reaction requiring water. Dominican Republic amber from the Oligo-Miocene epoch 30-40 million years ago, was interrogated for residual proteins. Exclusion of the protein isolates from polyacrylamide gel indicated molecular masses in the multi-million Dalton range, and failure of these aggregates to penetrate these gels proved to be an effective means for further concentrating trace proteins from paleontological samples. Tandem mass spectrometry (LTQ-Orbitrap) of trypsin digests led to the identification of 86 peptides from 20 saprophyte proteins. The most compelling evidence that these peptides are actually of prehistoric origin rather than the result of a contemporary experiment is the extraordinary high degree of crosslinking of these proteins.

1. INTRODUCTION

In 1994, Woodward et al. [1] reported the isolation of DNA fragments from a Late Cretaceous dinosaur bone exhumed from blumculinstitute.  
The following year, the cloning and sequencing of six putative dinosaur DNA fragments derived from a Cretaceous dinosaur egg fossil found in China was later disclosed as the recovered sequences were found to be more closely related to fungi rather than to reptiles or birds [2]. More recently, Asara et al. [3] identified peptides with sequence homology to avian collagen from the mineralized skeletal elements of *Tyrannosaurus rex*.  
Kaye et al. [4] challenged the finding showing that microbial biofils formed “endocasts” in which three-dimensional structure was preserved with morphoic detail, but in which the original organic material has been totally replaced by minerals.

The most promising circumstance enabling the preservation of biomolecules over millions of years comes from amber, the fossilized resin of leguminous trees. The unfossilized resins are comprised largely of terpenoids, labdanoids, and phenolics which rapidly dehydrate the included specimen, a prerequisite for fossilization, as well as possess anti-bacterial, anti-fungal, and anti-inflammatory properties that intervene with usual decomposition. When the specimen was completely engulfed in resin, dehydration having to occur within seconds, it resulted in unprecedented preservation later observed in amber fossils.

Michael Crichton’s novel Jurassic Park proposed the recovery of dinosaur DNA from the alimentary tracts of hemipterous insects preserved for millions of years in amber. Though Crichton's work was purely fictional, amber inclusions have shown remarkable preservation of organs at the tissue and cellular levels, and reptilian blood cells of terpenoids, labdanoids, and phenolics which rapidly dehydrate the includes specimen, a prerequisite for preservation, as well as possess anti-bacterial, anti-fungal, and anti-inflammatory properties that intervene with usual decomposition. When the specimen was completely engulfed in resin, dehydration having to occur within seconds, it resulted in unprecedented preservation later observed in amber fossils.

This report describes the isolation of high molecular mass protein aggregates from Oligo-Miocene amber from the Dominican Republic. LC-MS/MS of several proteins was homologous to *Saccharomyces cerevisiae* proteins. While mass spectra does not confirm the source of these peptides, the high degree of protein crosslinking suggests that they are not of contemporary origin.

2. METHODS

2.1 Sample Preparation

All procedures were performed under sterile conditions in a laminar flow hood. Amber pieces containing *Hymenaea protera* leaves were first scrubbed in 5% SDS with a dental brush, then heated to 90°C in 2% SDS and copiously rinsed in water. The amber was rinsed with 100% ethanol just prior to fracture. The amber was fractured and the fragments were ground to a fine triturate using in a sterile Shredder PULSE Tube with serrated metal ram insert (Pressure BioSciences, South Easton, MA). Triturates were extracted in 125 mM Tris-HCl pH 6.8 containing 2% SDS, 5 mM tributylphosphosphate, 20 mM AEDTA, 10 mM EDTA and 25 mM phenylalanylaacridinol bromide for 100 X 100 seconds at 35,000 psi maximum pressure in a Barocycler NEP 3229 (Pressure BioSciences, South Easton, MA). Samples were filtered in a Ultrafiltration column (Millipore Corporation, Danvers, MA) the filtrates were applied directly to 8% polyacrylamide gradient gels proved to be an effective means for concentrating trace proteins from paleontological samples while concomitantly removing interfering substances such as SDS prior to trypsin digestion and LC-MS/MS.

Identifying multiple proteins of *S. cerevisiae* origin was initially surprising. There were little peptide sequences of these proteins were further interrogated to verify the species assignment. For example, enolase 1 (Figure 2) sequences from *S. cerevisiae* and the rubber tree *Hevea brasiliensis* were compared. Although the eleven identified peptides hit conserved regions with high sequence similarity, none of the peptides had identical sequence between these two proteins. Blast NCBI protein database on the other four peptides in divergent regions only returned the saprophyte protein.

![Figure 1. Light microscopy showing budding yeast cells in Dominican Republic amber 20-30 million years old. Stages from single cell (1) to where the progeny cell is nearly equal in size to its parent (4) and intermediary stages (2,3) are shown. Magnification bar equals 6 um.](image)

![Figure 2. LTQ-Orbitrap mass spectra identifying peptides of sequence homology to *Saccharomyces cerevisiae* 1,3-SDS and copiously rinsed in water.](image)

Figure 2. LTQ-Orbitrap mass spectra identifying peptides of sequence homology to *Saccharomyces cerevisiae* 1,3-SDS and copiously rinsed in water.

3. RESULTS AND DISCUSSION

From mass spectra of trypsin digests, 86 peptides with sequence homology to 20 *Saccharomyces cerevisiae* proteins were identified in amber isolates (Table 1). The yeast is likely to be associated with plants and insects embedded in the amber. Experimental procedures for protein extraction, gel electrophoresis and mass spectrometric analysis were tightly controlled with blank samples. The fact that *S. cerevisiae* proteins were identified only in amber samples eliminates the possibility of contemporary contamination. More importantly, all these proteins were identified from a band at the interface of stacking and resolving gel, suggesting an extreme high degree of cross-linking for these proteins during amber formation. Mass spectrometric analysis of another gel band of amber sample did not lead to the identification of any protein. Exclusion of these proteins from 4% polyacrylamide gels indicated molecular masses of several million Daltons, and failure of the aggregates to penetrate these gels proved to be an effective means for further concentrating trace proteins from paleontological samples. Tandem mass spectrometry (LTQ-Orbitrap) of trypsin digests led to the identification of 86 peptides from 20 saprophyte proteins. The most compelling evidence that these peptides are actually of prehistoric origin rather than the result of a contemporary experiment is the extraordinary high degree of crosslinking of these proteins.

4. REFERENCES