Microsc. Microanal. 8 (Suppl. 2), 2002
DOI 10.1017.S1431927602100456

Linking Atomic Force Microscope Images of Proteins to Their Genetic Sequence

Steven J. Eppell and Brian A. Todd

Nanoscale Orthopedic Biomaterials Laboratory, Dept. of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106-7207

Efforts to link protein structure and biochemical function have focused mostly on primary structure, largely because of the abundance of information at this level; all of the amino acid sequences of the ~30,000 proteins in the human proteonome are known, whereas the three-dimensional crystal structures of only ~1000 proteins have been solved. Clearly, techniques that provide details concerning three-dimensional protein structure are needed to discover mechanisms governing biological function. Here, we have identified a motif in protein tertiary structure at sites of biochemical activity using scanning force microscopy (SFM). We show that sites on aggrecan, a cartilage proteoglycan, that are susceptible to catabolic enzymes are more flexible than other regions of the molecule. The results demonstrate a powerful new technique for investigating molecular scale structure-function relationships and suggest the role of flexibility in aggrecan degradation. This model system will be used to show how tip asymmetries can be accounted for using morphological processing coupled with AFM imaging simulation.

In Fig. 1A, we start with a model of a surface adsorbed aggrecan molecule. The backbone of the model is a 6 element 3rd order spline. Using an independently measured tip shape image (Fig. 1B), we then dilate the model to produce a simulated AFM image, Fig. 1C. The simulated image is then subtracted from a real AFM image, Fig. 1E, yielding an error image. This process is iterated using a method previously published to find the best fit of the model to the AFM data.[1] The result is an analytical function representing the surface adsorbed molecule.

Using intrinsic asymmetries in the models, we were able to register several such functions to obtain an average model. We looked for kinks in this model by evaluating it's second derivative (inversely proportional to the local radius of curvature). The values of these derivatives were then plotted as a function of contour length down the model backbone, Fig. 2A. Integrating previously published TEM data,[2] we mapped the backbone to the known genetic sequence[3] of the aggrecan molecule. The three highest sites of curvature overlap with three of the known aggrecanse cleavage sites.

1. W. Huyer and A. Neumaier, Global optimization by multilevel coordinate search, J. Global Optimization **14,** (1999), 331-355

2. M. Morgelin, M. Paulsson, A. Malmstrom, D. Heinegard, *J Biol Chem* **264**, (1989) 12080-90.

3. T. M. Hering, J. Kollar, T. D. Huynh, *Arch Biochem Biophys* **345**, 259-70. (1997).
We thank the NIH AR45664-02 and the Whitaker Foundation for their generous support.

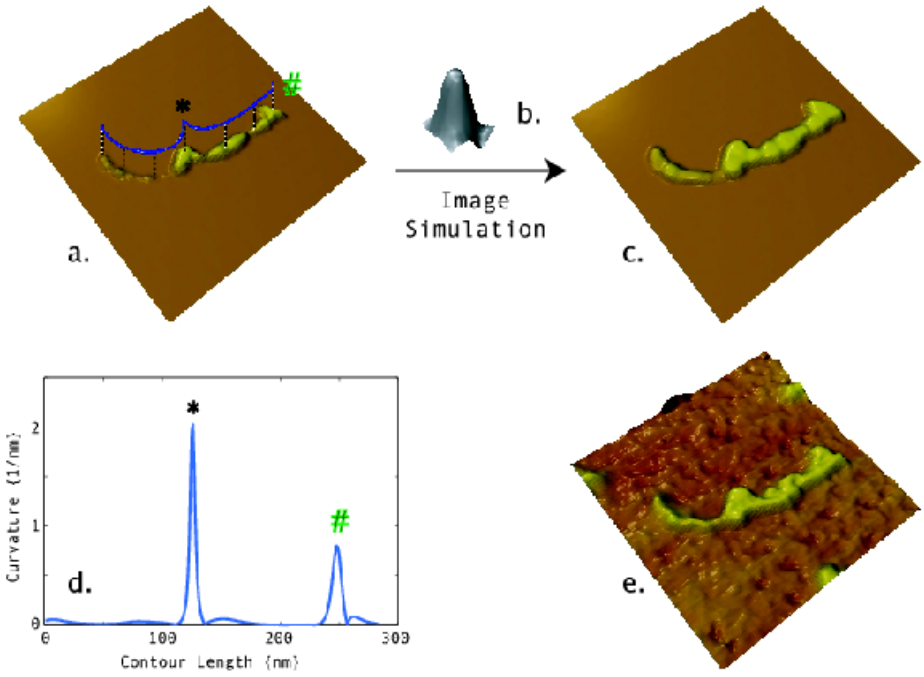

Figure 1

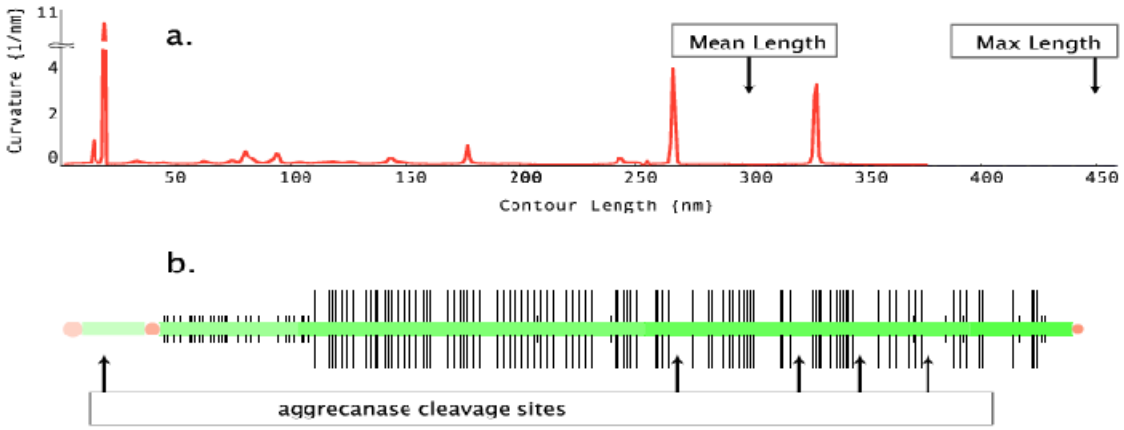

Figure 2

https://doi.org/10.1017/S1431927602100456 Published online by Cambridge University Press