Single-Molecule, Motion-Based DNA Sequencing Using RNA Polymerase.
Authors: William J. Greenleaf and Steven M. Block.
Science 113, 801 (2006).

Sequence-Resolved Detection of Pausing by Single RNA Polymerase Molecules.
Authors: K. M. Herbert, A. La Porta, B. J. Wong, R. A. Mooney, K. C. Neuman, R. Landick and S. M. Block.
Cell 125, 1083-1094 (2006).

Recommended with a commentary by
Steve Berry, University of Chicago

Two closely related papers by a group at Stanford have shown how it is possible now to sequence a single molecule of DNA. The terse presentation is the first one above. The fuller version is the second.

The essence of the approach is this: one attaches two tethers, as one normally does in single-molecule manipulation experiments. In this case, one tether is on the upstream or distal end of the nucleic acid, and the other, instead of being on that molecule, is on the RNA polymerase that reads the nucleic acid and builds a protein, one amino acid at a time. The RNA polymerase moves to the next downstream site when it has read a codon of the DNA and has attached the corresponding amino acid to the growing protein chain. The result is a somewhat discontinuous, "jerky" motion. The key to using this tool to determine the sequence is to carry out the process with just one amino acid at a time in short supply. Whenever the DNA calls for that residue, the polymerase has to wait until that correct but rare molecule comes into close proximity. Hence the "jerky" motion has an extra long pause when that scarce amino acid is designated by the DNA code. One can do this with each of the 20 naturally-occurring amino acids in turn.

The method is currently limited to determining a few thousand base pairs along the DNA because the ribonuclease can only process templates of that magnitude. The process is reasonably fast; most of the "long" pauses last only 1 to 5 seconds and occur roughly once in 100 bases, but occasionally, about once in 1000 bases, the delay can be over 20 seconds. The authors attribute these longer delays to "mis-incorporation errors."

The authors recognize that the method can be improved and made more efficient. For example, they suggest using multiple records, each from a single molecule, and carrying out a true deconvolution of the histogram records of the locations of the ribonuclease along the DNA strand. This method appears to be quite general, hence distinctly more powerful than either of the methods

previously suggested for single-molecule sequencing. One is fluorescence, and is based on the fluorescent property of those few amino acids that do fluoresce. The other uses penetration of specific DNA structures through nanopores.