In 1953, Francis Crick and James Watson discovered the double helix of DNA, the elegant simplicity of four base pairs connected in long, twisting strands that write the genetic instructions for all life.

They didn’t know the half of it.

A collaboration of Boston University and National Institutes of Health scientists has discovered that the shape of DNA — the molecule’s width and its nooks and crannies — may be as important as the base-pair sequences when it comes to translating genetic code into people. Their findings, which appeared in the March 12 online edition of Science, could revolutionize genomics and help unravel the genetic underpinnings of disease.

“DNA is not a perfectly uniform double helix,” says Thomas Tullius, a College of Arts & Sciences professor of chemistry, who collaborated on the research with bioinformatics doctoral students Stephen Parker (GRS’09) and Loren Hansen (GRS’10), along with researchers at the NIH’s National Human Genome Research Institute led by investigator Elliott Margulies. “There are subtle differences in the shape of the molecule that are very important to the binding of proteins.” These proteins, known as transcription factors, allow DNA to produce yet more proteins and also regulate when, where, and how much of these vital compounds are created.

Tullius and his fellow researchers compared one tiny, well-studied part of the human genome with the same region from the genomes of thirty-six mammals, among them mice, chimpanzees, and rabbits. They were looking for overlaps, believing that if a portion of the genome is the same across many species, it must be important.

“Mutations occur randomly,” explains Parker. “But they’re rejected in functionally important areas.”

A few years ago, other researchers did a similar cross-species comparison and found that about 6 percent of the genome section was constrained by evolution — meaning that it’s likely to have biological significance. But that study looked only at base pairs. The BU researchers incorporated the shape of each sequence in their comparison, using an algorithm developed by Parker. When shape was taken into account, the percentage of the genome that passed evolutionary muster doubled.

According to Tullius, this indicates that a lot of DNA’s function may be determined by the molecule’s shape, even if a few base pairs have been shuffled across the generations.

The researchers also found that DNA shape alters much more as a result of genetic mutations known to cause disease or biological change versus mutations that aren’t known to impact any biological function.

“What that means is that larger structural changes correspond to biological consequences,” says Parker.

This could affect the booming field of research into associations between genetic mutations and disease. Only a fraction of mutations known to impact biological function exist among the 2 percent of the human genome that directly codes for proteins. The vast majority of them, including more than 700 mutations studied by the BU team, exist among the remaining 98 percent of base pairs, whose role is understood less well.

“They’re out there in terra incognita,” says Tullius, “where you have a very hard time figuring out what went wrong and why this particular change led to heart disease or dementia.” A focus on DNA shape offers a new way to interpret these mutations.

The next step will be to expand this analysis to the entire genome, creating a topographical DNA map that includes identification of biologically important regions through comparisons with other mammals, transcription proteins that bind to different DNA shapes, and mutations that affect those structures.

“There’s all this stuff that’s invisible to us now,” says Tullius, “but we’re getting a clue about what’s happening out there.”