Flexible and unstructured regions of protein molecules introduce a source of conformational heterogeneity that is fundamental for their biological function. Here we study this heterogeneity with microsecond-scale molecular dynamics simulations using the NAMD 2.9 platform. Previous studies have shown that the dipeptide make-up of proteins delimits loop regions that are unstructured and flexible, driving the specificity and stereochemistry of the genetic code. A previous Blue Waters allocation confirmed that unstructured regions collapse quickly, but unexpectedly revealed increased intra-chain dynamics of short fragments. Our research provided unprecedented atomistic details of the dynamics of 74 loop regions of protein domains sampled along a timeline of domain history, studying collapse propensities and intrinsic fluctuations of each loop structure. Our study provides insight into how loop flexibility and disorder are linked to the genetic code and primes protein function in evolution.

METHODS & RESULTS
In previous work, we discovered that the speed of folding, which correlates with flexibility, is enhanced during the evolution of protein domains [5]. We also discovered that protein structures enriched with flexible loops appeared with the evolutionary unfolding of the genetic code [6]. Since structural flexibility is a conserved feature in the assembly of protein complexes [7], flexibility must be regarded as an emergent property of molecular evolution. Following an exploratory Blue Waters proof-of-concept allocation [8], we used the power of Blue Waters to study the dynamics of proteins loops in 74 protein domains of aminoacyl-tRNA synthetases, the enzymes responsible for the specificity of the genetic code. Advanced tools of molecular dynamics (MD) simulations using the NAMD 2.9 platform allowed gathering global parameters (the genetic code) and the evolutionary drivers of the structure of proteins.

Our results revealed that unstructured loop regions flex and unfold quickly, but unexpectedly revealed increased intra-chain dynamics of short fragments. Our research provided unprecedented atomistic details of the dynamics of 74 loop regions of protein domains sampled along a timeline of domain history, studying collapse propensities and intrinsic fluctuations of each loop structure. Our study provides insight into how loop flexibility and disorder are linked to the genetic code and primes protein function in evolution.

WHY BLUE WATERS?
The results of our initial exploration and benchmarking exercise provided a foundation for a Blue Waters-enabled high-throughput MD simulation study of the dynamics of a massive number of protein loops, which will be indexed with evolutionary, structural, and functional information. The study will yield unprecedented atomistic details of structural and functional evolutionary constraints that are responsible for structuring both proteins and the genetic code.

PUBLICATIONS