Nucleic acids play a central role in biological processes. DNA is the carrier of genetic information pertaining to the sequence of proteins and RNA. It forms double helices, which bind together if their sequences are complementary. This and the fact that it conducts electricity along the helix has led to applications in nanotechnology and electrochemistry. RNA comes in many diverse types. Each of them performs specific functions in the cell, mostly connected to transcription, translation or regulating gene expression. Naturally, nucleic acids are a valuable research object for scientists at the crossroads of biology and chemistry.

Along with the continuing increase of the available computing power, computational chemistry has achieved better and better accuracy to the point that, in many cases, it is cheaper to study molecular systems using theoretical methods before synthesizing them and confirming their properties in the laboratory than to perform a vast compound screening by using experimental methods only. There is a range of methods in computational chemistry with different speed, accuracy and applications, from the very accurate but slow methods of quantum chemistry, where all electrons in a system studied are carefully accounted for, to the much faster but less accurate molecular mechanics where the result is only an approximation of forces acting on atoms or energy of the molecule. For the quantum chemical methods, the required computing time increases fast with the size of the system, which puts a limit on their applicability. Thus, very large molecules with important biological functions are usually studied by using molecular mechanics, i.e., empirical force fields. Molecular mechanics is so fast that, even for very large systems, the forces acting on atoms can be calculated very quickly, which makes simulating the motion of the molecules possible; this method is called molecular dynamics. Usually in molecular mechanics all atoms in the system can interact with each other – these kinds of molecular mechanics methods are called all-atom force fields. Using all-atom force fields for molecular dynamics has made it possible to simulate processes that take several microseconds.

Some important biochemical processes, e.g., those occurring in cell take more time. When a protein is first synthesized, it is in an unstable conformation and, therefore, it converts to other conformations until it finds a stable one which, in almost all cases, also is the bioactive conformation. This process is called protein folding and usually takes from 1 millisecond to 1 second, a timescale beyond the reach of modern all-atom molecular dynamics. This is why coarse-grained force fields have been invented. In this approach, a group of atoms is represented by a single object. This way the number of interactions and motions to be calculated is lower, which can speed up the calculations by several orders of magnitude. This acceleration comes at the expense of detailed information of the all-atom structure, which is sometimes necessary to fully understand the nature of interactions underlying the studied process. Therefore, all-atom reconstruction algorithms were developed for many coarse-grained models in order to produce atomistic-detailed structures at the end of the calculations.

NARES-2P is a coarse-grained force field for the simulations of nucleic acids. It has been able to simulate the formation of DNA double helix at the correct temperatures for different sequences. NARES-2P simulations of telomeric DNA undergoing mechanical stress have brought remarkable insight as to what is the mechanism underlying telomeres' unique stability and why their sequences have been selected for.

The aim of this project is to develop a fast, reliable and accurate all-atom reconstruction algorithm for the NARES-2P coarse-grained model. This will enable us to gain information about the all-atom structure of nucleic acids from fast coarse-grained simulations. We will study the reconstructed all-atom structures of stretched telomeric DNA which will undoubtedly shed a new light on the telomeres' resistance to stress. We will also perform new simulations of small nuclear RNA, to investigate how it attains its structure and reconstruct the all-atom geometry of important intermediate structures. Our algorithm paves the road for other coarse-grained models with few beads.